University of Siena – Department of Medical Biotechnologies

Doctorate in Genetics, Oncology and Clinical Medicine (GenOMeC)

XXXIII cycle (2017-2020)

Coordinator: Prof. Francesca Ariani

CRISPR/Cpf1 and suicide gene as personalized approach for patients with *TP53*-mutated Chronic Lymphocytic Leukemia (CLL)

Scientific disciplinary sector: MED/03 – Medical Genetics

Tutor

Prof. Francesca Mari

PhD Candidate

Flaminia Clelia Lorenzetti

FLAMINIA CLELIA
LORENZETTI
11.03.2021
14:51:39 UTC

Academic Year 2019/2020

University of Siena – Department of Medical Biotechnologies

Doctorate in Genetics, Oncology and Clinical Medicine (GenOMeC)

XXXIII cycle (2017-2020)

Coordinator: Prof. Francesca Ariani

CRISPR/Cpf1 and suicide gene as personalized approach for patients with *TP53*-mutated Chronic Lymphocytic Leukemia (CLL)

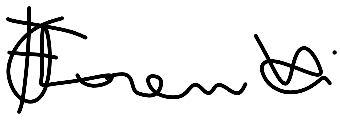
Scientific disciplinary sector: MED/03 – Medical Genetics

Tutor

Prof. Francesca Mari

PhD Candidate

Flaminia Clelia Lorenzetti



Academic Year 2019/2020

“To change the world

It starts with one step

However small

First step is hardest of all”

Dave Matthews Band, You Might Die Trying

Acknowledgments

First of all I should thank all the Medical Genetic Unit of Siena, especially Prof. Alessandra Renieri and my tutor Prof. Francesca Mari. It has been tough but motivating working with all of them. Of course, I need to thank Filomena Papa, she was always available to discuss, to help me with the experiments and to answer to all my questions, even the most silly ones! A big and warm thanks goes to all my lab crew with who I spent most of my time in the past 3 years.
I bring with me a lot of funny memories, thank you!

I would like to especially thank Prof. Rik Gijsbers. He is an incredible professor, researcher and most of all he is a really great person. It was inspiring discussing with him and his team about projects and experimental ideas. I am so grateful to have worked in his lab and I am very happy to keep collaborating with him.

Lastly, I want to thank my family and friends to have always be there to listen to me and to support me in the good and especially in the bad times of this important journey.

Table of Contents

Abbreviations

CAR: Chimeric Antigen Receptor

CIT: Chemoimmunotherapy

CLL: Chronic Lymphocytic Leukemia

CR: Complete Response

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats DSB: Double Strand Break

eGFP: Enhance Green Fluorescent Protein

FACS: Fluorescence activated cell sorting

GCV: Ganciclovir

gRNA: Guide RNA

HDR: Homology-Directed Repair

HSCT: Hematopoietic Stem Cells Transplantation

HSV1: Herpes Simplex Virus type 1

IDLV: Integration-deficient Lentiviral

Indel: Insertion and Deletion

KO: Knock-Out

LTR: Long Terminal Repeats

LV: Lentiviral

MMEJ: microhomology-mediated end joining

MRD: Minimal Residual Disease

NGS: Next-Generation Sequencing

NHEJ: Non-Homologous End-Joining

ORR: Overall Response Rate

PAM: Protospacer Adjacent Motif

PCR: Polymerase Chain Reaction

PFS: Progression Free Survival

R/R: Relapse/Refractory

TK: Thymidine Kinase

VSV-G: Vesicular Stomatitis Virus Glycoprotein

Abstract

Chronic Lymphocytic Leukemia (CLL) is a very heterogenous disease caused by alterations in both chromosomes and genes, such as deletion of the 13q14, 11q22-23, 17p12, and trisomy of chromosome 12 or genetic mutations in the *TP53, ATM, BRIC3, NOTCH*. Within genetic abnormalities, *TP53* mutations are detected in a small percentage of leukemia patients (about 10%). Currently adopted therapeutic choices (chemoimmunotherapy, targeted therapy, hematopoietic stem cell transplantation) are not effective due to the acquired abilities of *TP53*-mutated clones to escape the control systems. For this reason CLL patients harboring a mutated *TP53* gene (point mutation, insertion or deletions) are grouped in the highest risk category according to the international prognostic index for chronic lymphocytic leukemia (CLL-IPI). Hence, one possible solution for these patients would be the development of a personalized therapy.

In my study I designed and developed a CRISPR-Cpf1 system which has proved to be an efficacious technology to get rid of only mutated cancer cells. The system called CRISPR_LV_TK+, recently patented (WO2020/079574), is based on the locus specific delivery of the Herpes Simplex Virus – Thymidine Kinase (HSV-TK) suicide gene in cells bearing the target *TP53* mutation detected in a CLL patient (p.Ser183*) in follow-up in our Unit. The approach was positively tested in HEK293T cell lines engineered for the target mutation. Following administration of ganciclovir, I was able to detect a high percentage of cell death only in the samples that have properly integrated the HSV-TK. In conclusion, the results show the high efficiency and specificity of the CRISPR_LV_TK+ system, and opens avenues to be applied as personalized therapy not only for CLL but for different kinds of cancers caused by specific mutations in distinct genes which are the driver of therapy resistance.

Introduction

Chronic Lymphocytic Leukemia (CLL)

Chronic Lymphocytic Leukemia (CLL) is one of the most common types of leukemia in the western countries[1,2]. It affects mostly adults of >60 years of age who remain asymptomatic for many years but they suddenly show lymphocytosis which is the only symptom CLL patients show at diagnosis [3]. Patients post diagnosis can follow two separated ways, they can start specific therapy such as chemotherapy or they could never require any treatment[3,4]. Indeed, CLL is a heterogeneous disease characterized by several different causative events which may influence patients' treatment choice[5,6]. CLL is characterized by an uncontrolled B cells proliferation due to disruption of important cellular pathways such as the apoptotic pathway, the nuclear factor-κB (NF-κB) signaling pathway, NOTCH signaling pathway all involved in controlling cell growth and proliferation[7]. Chromosomal alterations like deletion of the 13q14, 11q22-23, 17p12, and trisomy of chromosome 12 or genetic mutations in the *TP53, ATM, BRIC3, NOTCH* genes are the main causes of the disease[7,8]. The immunoglobulin heavy variable (*IGHV*) gene mutational status is another crucial characteristic associated with CLL and patients can harbor two different forms of the genes, mutated and unmutated, which leads to a completely different outcome[7]. It has been reported that CLL patients with mutated *IGHV* genes sequence have higher overall survival (OS) due to a better response to chemotherapeutic agents with respect to patients harboring the un-mutated *IGHV* genes[4].

Patients show a different survival rate and response to treatment depending on the type of chromosomal or genetic alterations they harbor[4,9]. According to Dohner, CLL patients can be classified into 5 risk groups associated with different prognosis where CLL cases harboring the 17p chromosomal deletion or a mutant p53 are grouped in the worst prognostic class and patients with 13q deletion belong to the lowest risk group[10]. The ranking has been recently modified by an international consortium which was able to establish the international prognostic index for chronic lymphocytic leukemia (CLL-IPI) which helps clinicians to better predict the OS and the time-to-first treatment (TTFT) for specific patients harboring different genetic alterations[6]. In the CLL-IPI,

patients are divided into 4 risks groups: low-, intermediate-, high-, and very-high-risk group taking as prognostic biomarkers the clinical stage, the age and the genetic alterations (*TP53* mutation, *IGHV* status and serum β $_2$-microglobulin concentration) detected in patients[2,6]. Specific scores are assigned at each prognostic factor in order to develop a specific grading system essential to group the patients into one of the 4 categories[11]. As for the Dohner ranking, also in the CLL-IPI patients with *TP53* mutations are grouped in the very-high-risk category.

TP53 mutations or deletion of the 17p chromosome are detected in a small fraction (around 10%) of CLL patients but those harboring either or both *TP53* mutation and 17p deletion show chemo-refractoriness and a short OS (median OS of 79.6 months and of 104.2 months, respectively)[12]. Indeed, *TP53* is probably one of the most important and essential genes since its role in governing cell fate. It encodes for the tumor suppressor protein p53 (53kDa) which is a transcription regulator activated during cellular stress, hypoxia, DNA damage or cellular metabolic dysfunctions[8,13]. Low levels of p53 are present in normal and healthy cellular environments due to its downregulation performed by MDM2 which is an oncoprotein responsible for the ubiquitination and subsequent proteasome degradation of p53[14]. However, particular cellular stresses (Figure 01) activate p53 which as transcription factor triggers the expression of further genes leading to cell cycle arrest or apoptosis[9,13,14]. Mutations in the *TP53*, occurring specifically in the DNA binding domain (DBD), leads to a loss of function of p53 transcription factor and they are the main causes of human cancers development[1,14]. Moreover, cancer patients harboring *TP53* mutations show chemo-refractoriness since chemotherapy

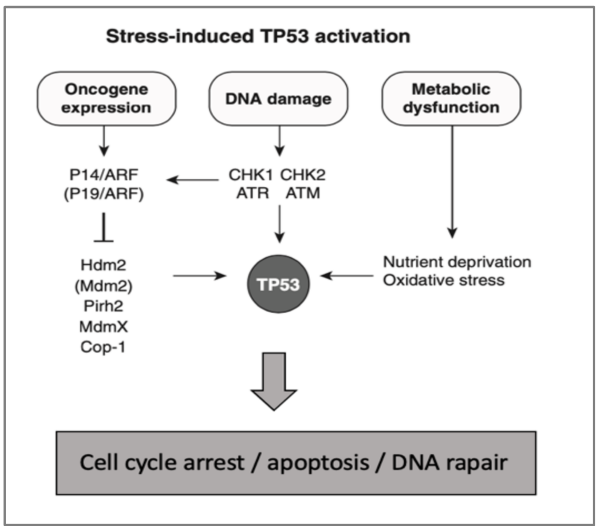


Figure 01. TP53 gene activation. Cellular stress, hypoxia, DNA damage or cellular metabolic disfunctions can trigger the activation of TP53 gene encoding for the p53 transcription factor. p53 will activate downstream genes that encode proteins responsible for cell cycle arrest, apoptosis or DNA repair. (Figure adapted from B. J. Aubrey et al., Cold Spring Harb. Perspect. Med, 2016)

induces a toxic effect on cells which should activate apoptotic pathways in order to induce specific cell death[5,12]. However, in case of *TP53* mutated cells, apoptotic pathway activation is impaired and cancer cells will keep proliferating and occasionally acquiring new genetic alterations due to the toxicity induced by chemotherapy[1].

Hence, mutations in *TP53* or deletions in the 17p chromosome need to be carefully detected at diagnosis to start the correct treatment for CLL patients. In order to accurately detect abnormalities in the *TP53* gene the European Research Initiative on Chronic Lymphocytic Leukemia (ERIC) suggested specific guidelines that should be followed[15]. First of all, the employment of the next generation sequencing (NGS) technology in order to detect genomic variants carried by the patients at low frequency which could have been hidden using Sanger sequencing[6,9,15]which is less sensitive technique than NGS[16]. Second, ERIC suggested to sequence *TP53* exons 4 to 10 which embedded the DBD region where most mutations occur[15]. Third, it is important to carefully analyze the data acquired after DNA sequencing using specific bioinformatic tools to avoid skipping any significant variants[15]. Finally, they give tips about the quality of the genomic DNA that needs to be sequenced, the databases which can be consulted (IARC or COSMIC) for the identified variants and they also give a list for clinical significance of exonic and intronic variants which are commonly detected[15,17,18].

Mutations characterized by a low variant allele frequency (VAF) need to be carefully identified in order to avoid their clonal expansion which can be deleterious for the patients[15,19]. As described by Landaou et al., two types of driver mutations may be present in a CLL patient, predominantly clonal and sub-clonal[3,20]. The former is detected at diagnosis whereas the latter and most critical one shows later on during disease course due to its expansion[20]. Indeed, the starting of chemotherapy could help the sub-clone to spread due to the toxicity of the therapy targeting the predominant clone[9,20]. Hence, with no competitors the sub-clone is able to proliferate and expand, ultimately acquiring resistance for the cancer to the therapy which lead the patients to chemo-refractoriness[3,21]. Moreover, *TP53* mutations are good markers for chemotherapy resistance in CLL patients because it has been shown that *TP53* sub-clones can be present at a very low frequency at

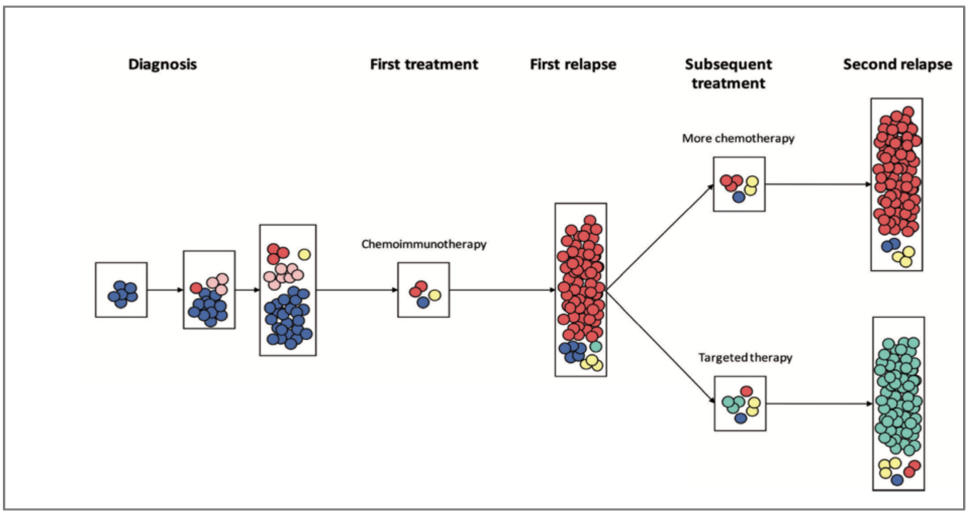


Figure 02. TP53 clonal expansion. TP53 mutated sub-clones can be missed at diagnosis because they are present at a low frequency at early stage of the disease. Following the first line of treatments mutated sub-clones can be selected leading to a second relapse. Different colors show the expansion of different mutated clones. (Figure from Campo et al. Hematologica 2018)

the start of the disease, and can be missed during the diagnosis leading to their clonal expansion[9,12,21] (Figure 02).

Deep sequencing by NGS is recommended in order to both identify possible sub-clones, which can be selected by chemotherapy, and to help clinicians to start the proper treatment avoiding unwanted clonal selection which may leads to patients relapse[19,21].

Current treatments and clinical trials

A specific therapy for CLL patients does not exist due to the high heterogeneity of the disease, thus treatment choices are different depending on the biomarkers identified in each patients[22]. Many CLL patients can live a normal life without the need of any medications, thus the first clinical approach is a "watch and wait" strategy where the patient is regularly monitored[8,23]. As soon as a patient starts developing new symptoms such as splenomegaly, anemia, thrombocytopenia the starting of a therapeutic plan is strongly suggested[24]. Several are the treatment options available for CLL patients: chemotherapy, antibody-based therapy, targeted agents, CAR-T therapy and finally the hematopoietic stem cells (HSC) transplantation.

Cytotoxic activity of chemotherapy using alkylating agents such as chlorambucil or purine analogs such as fludarabine, cyclophosphamide and bendamustine, induces DNA damages which lead to the activation of downstream molecules involved in the apoptotic pathway. Although both

alkylating agents and purine analogs showed positive response in treated patients with an adequate overall response rate (ORR) and progression free survival (PFS), they are not suitable for all CLL patients[24]. Indeed, for CLL patients characterized by an aberrant *TP53* gene the apoptotic pathway is impaired leading to accumulation of cytotoxicity which may result in the accumulation of novel mutations[1]. However, chemoimmunotherapy (CIT), thus the combination of chemotherapy and antibody-based therapy is now widely applied as a treatment of choice for almost all CLL patients. Rituximab is the first developed type of anti-CD20 monoclonal antibody (mAb) which show promising results in CLL patients[23]. Upon binding to CD20 antigen which is expressed on both neoplastic and healthy B cells, Rituximab will activate different cellular pathways in order to eliminate cancer cells, such as antibody-dependent cellular cytotoxicity (ADCC), complement-dependent cytotoxicity (CDC), programmed cell death (PCD) or adaptive cellular immunity[25]. During ADCC, immune effector cell macrophages and natural killer (NK) cells recognize and kill CD20 target cells by releasing cytotoxic enzymes[26,27]. When the CDC pathways is activated, cancer cells are lysed by cytotoxic enzymes released by the complement system[25]. The biding of Rituximab to CD20 can also trigger the adaptive immunity, thus dendritic cells (DC) will present tumor antigens to T cells which will differentiate into cytotoxic T cells able to eliminate target cells[25,26]. Lastly, Rituximab can also lead cancer cells to apoptosis by activation of the caspase 3[27].

CIT treatment with fludarabine, cyclophosphamide and rituximab (FCR) have demonstrated increase in ORR, complete response (CR) and a long PFS[28]. Rituximab combined with bendamustine (BR) is another possible CIT regime[23]. However, when compared with FCR, BR showed less toxicity but a decrease in PFS, whereas the ORR was identical in both the CITs[23]. Unfortunately, CIT did not show significant increase in the positive response rate in CLL patients harboring a *TP53* mutation or a del17p[6,9,24].

Targeted agents acting on the apoptotic pathway regardless of the *TP53* status have shown great positive outcomes in several clinal trials. The first FDA approved targeted therapy in 2016 was ibrutinib which acts as inhibitor of the Bruton's tyrosine kinase (BTK) involved in the B cell receptor

(BCR) signaling pathway[1]. Another powerful kinase inhibitor is idelalisib targeting the phosphoinositide 3-kinase δ (PI3Kδ)[23,24]. The FDA allowed its use together with ibrutinib but due to the high toxicity detected in a group of treated patients, idelalisib should not be considered as first therapeutic choice[23]. Venetoclax is a third type of targeted agent able to induce cellular apoptosis by inhibiting the anti-apoptotic protein BCL-2[1]. It was approved by the FDA and it is considered the first line therapeutic agent for CLL due to the outstanding positive results detected in the clinical trials[23]. When combined with ibrutinib, venetoclax showed an increase in ORR, PFS, CR and the total absence of minimal residual disease (MRD) or toxicity compared to all the other type of available therapies[1]. Moreover, in the MURANO Phase III study venetoclax combined with rituximab revealed a deep decrease in MRD and prolonged PFS when compared with bendamustine and rituximab therapy[29]. Interestingly, ibrutinib, idelalisib and venetoclax all demonstrated a strong performance also in relapse/refractory (R/R) patients harboring an aberrant *TP53* gene[1,23,29].

The newly developed chimeric antigen receptor (CAR) therapy was recently tested in CLL patients. CAR therapy is an *ex vivo* gene- and cell-therapy in which patients T cells are isolated, genetically modified and re-infused into the patients[30]. The resulting modified T cells express on their surface receptors that specifically recognize cancer-specific antigens present on neoplastic cells[31]. Patients receiving CAR treatment may develop cytokine release syndrome or neurotoxicity as adverse events[1,30]. However, in the next few years results from the ongoing clinical trials will give a more general safety assessment for CLL patients treatments (NCT03960840, NCT02935257, NCT01853631).

Allogenic hematopoietic stem cells (HSC) transplantation (HSCT) is considered the ultimate treatment for CLL especially for R/R patients bearing *TP53* abnormalities but it is not recommended for elderly patients[1]. HSC transplantation should be considered only when the other therapeutic options failed or when patient biomarkers are particularly favorable for the transplant[32]. About 50% of patients receiving allogenic HSCT showed stable negative MRD during follow up[32]. However, HSCT is a risky procedure which could possibly end up with the development of graft versus host

disease (GvHD)[32,33]. The group of M. Hahn stated that pre- or post-therapeutic strategies should be also explored in order to overcome the GvHD in relapse CLL patients who undergo HSCT[34]. However, *TP53* mutated CLL patients still gain little benefits from the aforementioned therapies showing relapse, thus a personalized, cancer cell specific therapy could be an important option to consider in order to help these patients.

Genome editing technology

It was a woman, named Barbara McClintock, the pioneer in genome editing. In the 1950 she studied corn (*Zea mays*) and discovered particular genomic sequences called transposons or mobile genetic elements able to move from one region of the genome to another leading to the insertion of random mutations[35]. Later on, in the 1980s, homologous recombination (HR) was described by Capecchi, Smithies and Evans. When foreign double stranded DNA (dsDNA) having homologies with a target sequence was added into target cells, the exogenous molecule was inserted at that particular place[36]. The era of the first generation of genome editing technologies ends in the late 90ies when RNA interference (RNAi) was first reported as a method to knock-down specific RNA sequences, thus controlling gene expression[37]. At the beginning of the 21st century the second generation of genome editing technologies started to be explored. Zinc Finger Nuclease (ZFN) and TAL effector nuclease (TALEN) were the first tools to be employed in the targeting of specific genomic sequences[36]. They consist of a DNA-binding moiety that defines the specificity and is coupled to one half of the bacterial FokI endonuclease[38,39]. Thus, two ZFNs or two TALENs (one left and one right of the target site) dimerize at the target sequence, and produce a double strand break (DSB) at the desired locus[40–42]. The protein domains that bind the target sites on the genomic DNA (ZnF or TALE) can be modulated in order to recognize different nucleotides with the possibility of editing several different genomic sequences[43,44]. Unfortunately, it is very laborious and expensive to modify TALE or ZnF protein structures. In 2005 the discovery of the clustered regularly interspaced short palindromic repeats (CRISPR) was still to completely revolutionize the research in life science.

CRISPR was first discovered in bacteria species which developed a system to fight against invading viruses, bacteriophages[45,46]. Indeed, after bacteriophage infection part of its genome is copied and pasted within specific repeated sequences located in the bacterial genome[45,47]. Hence, the bacteria evolved a kind of immune system which helps them to promptly recognize and get rid of a virus following a second infection. So far six CRISPR systems have been identified in different bacterial species and they can be divided into two groups: Group 1 encompassed CRISPR type I, III and IV, whereas type II, V and VI made the Group 2[48].

The well-known type II CRISPR system (SpCRISPR-Cas9) derived from *Streptococcus pyogenes* bacteria strain has been extensively studied and employed for different purposes of genome targeting and editing (Figure 03 A), not only in bacteria, but also in other species, such as mammalian cells (rat, mouse, human). The endonuclease responsible for the proper cutting of the target site is the Cas9 which is composed of two nuclease domains, HNH and RuvC domains able to cut the complementary and the non-complementary DNA strand leading to DSB generation[49]. However, Cas9 alone is not enough to perform the DNA cleavage, a small guide RNA sequence (coined sgRNA) needs to be present into target cells in order to drive the nuclease toward the desired gDNA locus[45]. The sgRNA is a sequence of 20 bp which will base pair at its 5'end with the complementary DNA strand whereas its 3'end will bind the Cas9[50]. Once the

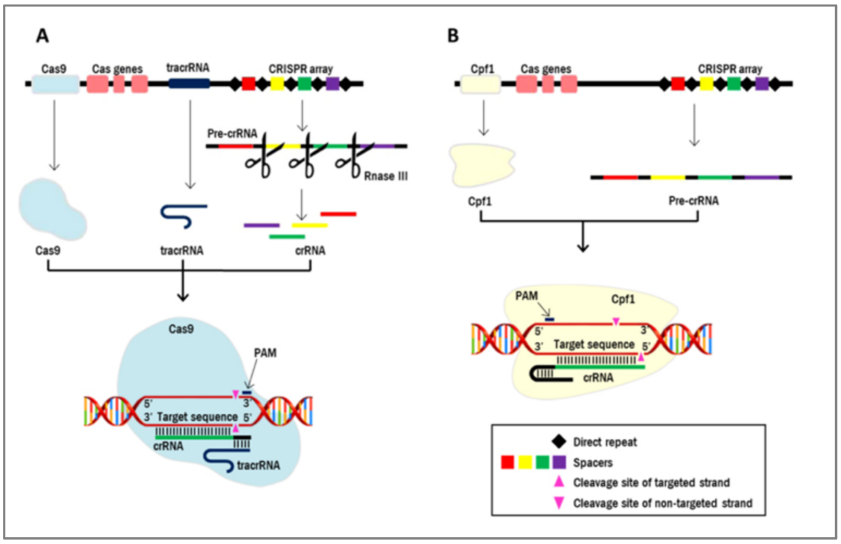


Figure 03. Differences in CRISPR-Cpf1 and CRISPR-Cas9 system. A) CRISPR-Cas9 needs two RNA sequences, tracrRNA and crRNA, which were fused together to form a sgRNA to drive the Cas9 to the target site. The Cas9 will recognize a G-rich PAM site located at the 3'end of the target site and it will generate DSB with blunt ends. B) CRISPR-Cpf1 employs a crRNA able to drive the Cpf1 toward the target site. The Cpf1 will recognize a T-rich PAM site located at the 5' end of the target sequence and it will generate sticky ends DSB. (Figure from Swiat et al. Nucleic Acids Research 2017).

sgRNA-Cas9 complex reach the target site an additional important prerequisite for a correct genome editing is the presence of a particular sequence located at the 3'end of the genomic locus to be edited, called protospacer adjacent motif (PAM) site, which is recognized by the nuclease[51]. PAM sites differ in different bacterial species, for example SpCas9 is able to recognize 5'-NGG-3' PAM[50].

In 2015, the type V CRISPR (FnCas12a or FnCpf1) system was discovered in *Francisella novicida* bacterial species[48,52–54] (Figure 03 B). It works similarly to CRISPR-Cas9 but it shows some differences. First of all, the nuclease Cpf1 recognizes an alternative T-rich PAM site (5'-TTN-3') located at the 5'end of a target sequence[48]. The Cpf1 is characterized by only a single nuclease domain, the RuvC domain[54].

Finally, CRISPR-Cpf1 will generate DSB with sticky ends rather than blunt ends produced after Cas9 cleavage, which is an advantage for knock-in experiments[55]. Following DSBs generation, cells can choose between two different repairing mechanisms (Figure 04): non-homologous end-joining (NHEJ) and homology-directed repair (HDR)[56]. NHEJ, which is basically the restitching of the ends that are generated can be activated in any phase of cell cycle and it will occasionally produce insertion or deletion (indels) at the edited site (as long as these are not a multiple of three), thus it is preferred for knock-out experiment [57,58]. On the other hand, HDR is an error-free repairing mechanism which employs a template sequence or the un-cleaved DNA strand in order to repair the DSB[59]. However, HDR occurs only during the S and G2 phase of the cell cycle[60]. Finally,

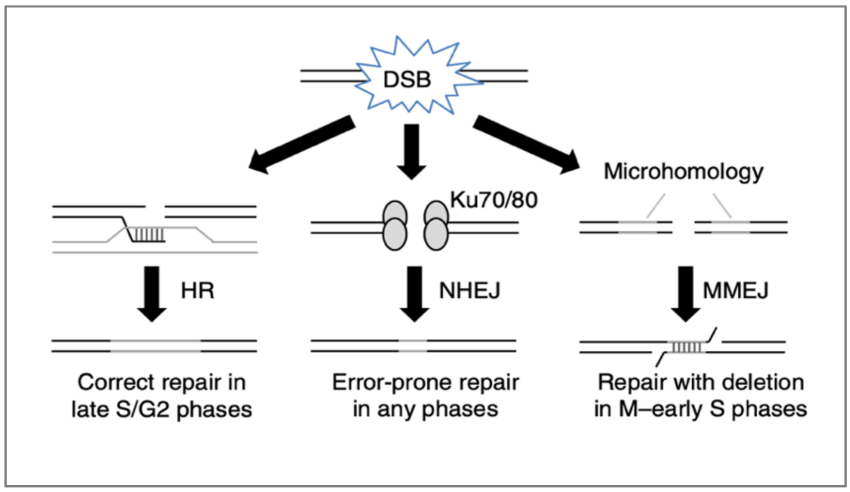


Figure 04. DSB repairing mechanisms. Homologous recombination (HR): it is an error-free repairing mechanism occurring in the S and G2 phase only of the cell cycle. Non-homologous end-joining (NHEJ): it occurs in any phase of the cell cycle and it will generate insertion or deletion (indels) at the edited site. Microhomology-mediated end joining (MMEJ): it occurs in the M, G1 and early S phase of the cells cycle. Once the MMEJ mechanism is activated, the homology arms (5-20 bp) are joined together to repair the DSB. (Sakuma et al. Nature Protocols 2016).

a third type of repairing mechanism called the microhomology-mediated end joining (MMEJ) is described (Figure 04), which occurs in the M, G1 and early S phase of the cells cycle when HR is not active. For MMEJ to occur, 5 to 20 bp of homology arms need to be present at the edited site which is fixed by the annealing of the microhomologies[61–63].

The PAM sequence in the end limits the applicability of CRISPR. However, Cas orthologs have been identified in various bacterial strains able to recognize different PAM sites, thus allowing the targeting of several unique loci for particular experimental aims[64]. CRISPR-Cas system is largely employed to study gene functions by knocking-out specific genes and it is also used for knock-in experiments aimed at correcting detrimental mutations in order to restore a wild type and healthy cellular state. Hence, CRISPR applications described so far have been applied in different areas of research from plant biotechnologies to biofuels production and also in the biomedical fields with the development of powerful gene and cellular therapies tested for a wide range of diseases.

CRISPR based therapies

Several studies are testing CRISPR editing technology in different genetic disorders such as Duchenne Muscular Dystrophy (DMD), ß-thalassemia and sickle cell disease[65]. DMD is caused by a nonsense mutation on the dystrophin gene responsible for maintaining the correct integrity of muscle fibers. Mutations occurs on the exon 23 leading to the complete loss of protein expression. Hence, the excision of exon 23 by CRISPR-Cas9 tool could theoretically restore the expression of a functional dystrophin protein. Systemic delivery of the CRISPR system showed great efficiency *in vivo* studies performed in both mice and dogs[66,67]. Moreover, positive results were achieved *in vitro* and *in vivo* on animal models for the hemoglobinopathy ß-thalassemia for which a clinical trial is ongoing (NCT03655678). ß-thalassemia is caused by mutations in the ß-globin gene which inhibit the translation of the ß-globin protein. The therapy is based on the *in vitro* editing of autologous hematopoietic stem cells (HSC) by CRISPR in order to knock out the BCL11A a repressor of the γ-globin allowing its expression which can restore a healthy condition in affected patients who will

receive the edited cells back by transfusion[68]. This strategy could also be applied for sickle cell disease characterized by the synthesis of an aberrant ß-globin due to mutation in the ß-globin gene[65,68]. The first human clinical trial using CRISPR technology as an *in vivo* approach started in 2019 and aims at restoring a healthy condition for people with Leber Congenital Amaurosis 10 (LCA10). The trial sponsored by Editas Medicine and Allergan is based on the research held by the group of H. Jiang et al.[69] who developed EDIT-101, a CRISPR Adeno Associated Virus (AAV)-based therapy that uses two gRNAs to target and excise an intronic mutation (c.2991+1655A>G) within the CEP290 (NCT03872479). Indeed, the endonuclease Cas9 and two sgRNAs are delivered into patients by subretinal injection targeting two sites adjacent to the mutation. Once Cas9 cleaves both target sites the mutated intron gets removed and correct splicing will occur leading to the translation of a functional CEP290 protein[69]. Finally, CRISPR has found its way also in the fight against cancer with the development of T cell therapies: CAR and T cell receptor (TCR). Indeed, CRISPR is employed to edit T cells isolated from patients in order to specifically drive the engineered T cells toward cancer cells. Clinical trials testing the safety and feasibility of T therapies are ongoing for both hematologic and solid tumors. The company CRISPR Therapeutics AG tried for the first time to overcome a huge obstacle related to CAR therapies by engineering allogenic CAR-T cells (http://www.crisprtx.com). One of the main issues related with CAR generation is the availability of enough autologous T cells to be edited[70]. CRISPR Therapeutics AG is testing the CTX110 which applies CRISPR-Cas9 edited allogenic CAR-T expressing an anti-CD19 receptor (http://www.crisprtx.com; NCT04035434). The phase I clinical trial started July 2019 and is still recruiting patients. Currently, it is running in the US and in Germany enrolling up to 95 participants with relapsed or refractory B-Cell malignancies (NCT04035434). Moreover, a group from the University of Pennsylvania developed an innovative type of TCR therapy which entered the first clinical trial on January 2018 for patients with multiple myeloma, melanoma, synovial sarcoma and myxoid/round cell liposarcoma (NCT03399448). The group of C.H. June[71] was able to perform the knock-out of 3 genes in autologous T cells by employing a multiplex CRISPR-Cas9 system carrying 3 different gRNAs directed toward PD1 gene, responsible

for regulating immune system cell response, and toward TCRα and TCRβ genes essential for TCR synthesis[71]. Hence, deletion of these genes will increase persistence of engineered T cells expressing the mutated TCR[71]. However, T cell therapies still present some issues related to time and cost needed for production, long persistence of the engineered T cells once infused back into the patients and the development of possible side effects such as cytokine release syndrome into treated patients[30].

Delivery by viral vector: lentiviral (LV) vector

Physical, chemical and biological delivery methods can be selected to introduce the CRISPR-Cas system into target cells. Electroporation, lipofectamine or polyethylenimine (PEI) are commonly used physical and chemical delivery techniques for transient *in vitro* experiments[72,73]. However, viral vectors are employed to efficiently deliver CRISPR tool *in vivo*. Adeno-associated virus (AAV) and lentivirus (LV) are the most widely adopted biological delivery tools. They show high transduction efficiency in various dividing and non-dividing cell types and tissues[74]. The disadvantages of viral vectors are the low packaging capacity and the possible immune reaction[72].

Lentiviral vectors (LV) are derived from lentiviruses, which belong to the *Retroviridae* viral family and they are characterized by a single stranded RNA (ssRNA) genome which is reverse transcribed once lentivirus infects a target cell[75]. The newly transcribed viral DNA flanked by specific viral sequences called long terminal repeats (LTR) is then integrated into the host genome leading to viral gene expression[76]. LV are able to efficiently transduce both dividing and non-dividing cells as described for the first time in 1996 by Naldini et al., who employed a modified version of the HIV-1 vector for gene therapy purposes[77] . Indeed, HIV-1 based vectors are the most widely used LV vectors in research. However, some concerns arise due to the ability of LV to integrate in the genome of target cells. Indeed, following lentiviral integration insertion of random mutations which may results in the activation or repression of genes important in regulating a healthy cellular state may lead to cancer development[75,78]. Hence, for safety reasons first generation LV vectors were re-designed in order to remove genes naturally present in HIV-1 virus which can be critical once integrated in a

target genome[75,76]. Second generation LV is characterized by the lack of vif, vpr, vpu and nef accessory genes and tat gene important for viral replication and pathogenicity[75,76]. Further safety improvements were applied for the production of the third generation of LV also referred to as self-inactivated (SIN) LV vectors[79]. The transcriptional role of the LTR was edited by introducing a mutation in the U3 region of 3'LTR sequence, thus decreasing the chance of generating replication competent lentivirus (RCL)[75]. The production of LV relies on the expression of essential genes located in the transfer plasmid and in the helper plasmids. Indeed, interaction between LTR, RNA packaging signal (ψ), Rev-response element (RRE), woodchuck-hepatitis virus post-transcriptional regulatory element (WPRE) and gag, pol and env is important for regulating RNA reverse transcription, RNA stabilization and export from the nucleus and for viral envelope construction[78,80].

Lentiviral vectors based therapies

LV can be pseudotyped by changing the env gene naturally present in the HIV-1 viral genome with one originating from another family of viruses in order to allow the proper transduction of different target cells. The most employed env gene is the VSV-G from the vesicular stomatitis virus glycoprotein. VSVG binds the low density lipoprotein receptor (LDLR) and shows a great tropism for the majority of mammalian cells[75]. However, the hemagglutinin (H) and the fusion (F) glycoproteins from the measles virus (MV) were used by two different groups which both showed the high efficiency of MV glycoproteins compared to VSV-G for the transduction of primary B lymphocytes[81,82]. The group of Funke added on the H domain either epidermal grow factor (EGF) receptor or an anti-CD20 antigen in order to make the B cells transduction even more specific[81]. However, they stimulate B cells with a cocktail of cytokines prior transduction. The group of Verhoeyen showed the ability of H/F-LV to transduce even unstimulated B lymphocytes with a greater efficiency than VSV-G[82]. Moreover, the envelope from the baboon endogenous retrovirus (BaEV-LV) deriving from the ß-retroviruses was also employed to efficiently pseudotype LV vectors[83]. A relevant quality characterizing BaEV-LV is its ability to escape the human complement

system decreasing the risk of starting an immune reaction in the host organism. Girard-Gagnepain et al. tested BaEV-LV into hematopoietic stem cells (HSC) which are important therapeutic targets for several disorders. They show the great specificity of BaEV-LV to transduced HSC without the need of cellular stimulation which may alter HSC phenotype[84].

Due to their broad tropism and the stable integration in the host cell genome ensuring persisting expression in daughter cells, lentiviral vectors are largely employed for the treatment of a variety of disorders such as cancer, immunodeficiency or hematological disorders. CAR-T are developed by *ex vivo* transduction of patient-derived autologous T cells using lentiviral vectors (or MLV-based retroviral vectors) carrying cDNA encoding for the CAR directed against the desired cancer-specific antigens[30]. Lentiviral vectors have been used to carry out an *ex vivo* gene therapy aimed at modifying HSC isolated from patients with Wiskott-Aldrich Syndrome (WAS) caused by mutation in *WAS* gene[85]. Indeed, autologous HSCs are in vitro transduced with LV harboring the wild-type ORF sequence of *WAS* and later on re-injected into the patient restoring a functional WAS protein[86]. A similar strategy was adopted in the phase I/II clinical trial to modify autologous HSC for the treatment of Fanconi anemia caused by mutation in the *FANCA* gene (NCT03157804). Indeed, LV were selected as vector to drive a wild-type *FANCA* gene into the HSC isolated from patients[87]. All the aforementioned trials are based on the *ex vivo* delivery of LV which will integrate into the host genome leading to the expression of the desired cDNA. Even though several studies are testing the injection of LV into animal models, *in vivo* LV delivery in humans has not been tested yet. Hence, next step in gene therapy would be the development of safe LV based vaccines to be tested *in vivo* in order to treat patients with life threatening disorders which cannot benefit enough from ex-vivo strategies.

Aim of the project

The presented study aimed at specifically targeting cancer cells holding mutation in one of the genes responsible for cancer development such as *TP53* or *KRAS*. We focused on a *TP53* stop mutation (NM_000546.5(TP53):c.548C>G(p.Ser183*)) reported as pathogenic in ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/variation/634701/) and detected in a patient with Chronic Lymphocytic Leukemia (CLL) who is in follow up at our unit. As stated in the introduction CLL patients bearing *TP53* alterations, either genomic mutations or chromosomal deletions of chromosome 17p, belong to the very-high-risk category according to CLL-IPI ranking[6] showing short overall survival and high chance of relapse post treatments. Hence, we aimed at developing a precise gene therapy system for these patients by designing an all-in-one lentiviral vector (LV) carrying the CRISPR-AsCpf1 components and the sequence of the suicide gene HSV-TK.

Results

Strategy of the system

A lentiviral vector harboring the CRISPR_LV_TK+ system, composed of both the CRISPR-AsCpf1 and the suicide gene sequence TK deriving from the herpes simplex virus type 1 (HSV1), is employed to edit and kill only mutated cells once the antiviral prodrug Ganciclovir (GCV) is provided to the cells. Indeed, GCV is phosphorylated and converted into a toxic compound by HSV-TK. The resulting GCV triphosphate will act as inhibitor of DNA elongation disrupting DNA synthesis and leading to cell death.

Two sgRNAs (Table SI) are designed which are able to recognize different target sites: the target mutation on the cell genome and a sequence on the viral vector itself (Figure 05).

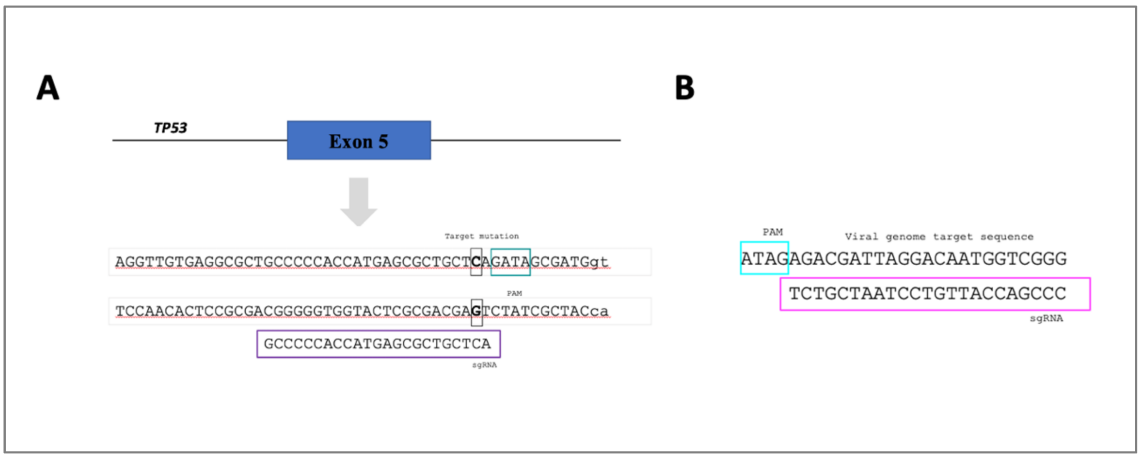


Figure 05. Target mutation and sgRNAs sequences. A) Represented are: the target mutation (c.548C>G; p.Ser183) located on the exon 5 of TP53 gene and the sequence of the corresponding designed sgRNA and the PAM site located on the target site. B) The sequence of the second sgRNA targeting the viral DNA and the PAM site found on the viral genome are shown.*

Once the LV is delivered to the cells, it binds to the cellular membrane and releases the viral RNA which is reverse transcribed to generate viral dsDNA (figure 06 A-C). The viral DNA is imported into the nucleus where both the sgRNAs and the AsCpf1 complex together and move toward the target sites leading to the production of DSB at the target locus on the cell genome and to the removal of the eGFP-TK sequence from the viral DNA itself (figure 06 D-E). The removed eGFP-TK construct holding homologous arms to the edited target genomic sequence will be properly integrated in the cell genome by MMEJ repairing mechanism, thus the edited cells will encode for both eGFP and HSV-thymidine kinase (figure 06 F). The latter will phosphorylate the antiviral prodrug GCV, which will in turn remain inside the cells and is able to inhibit DNA replication leading cells to death (figure 06 G-I).

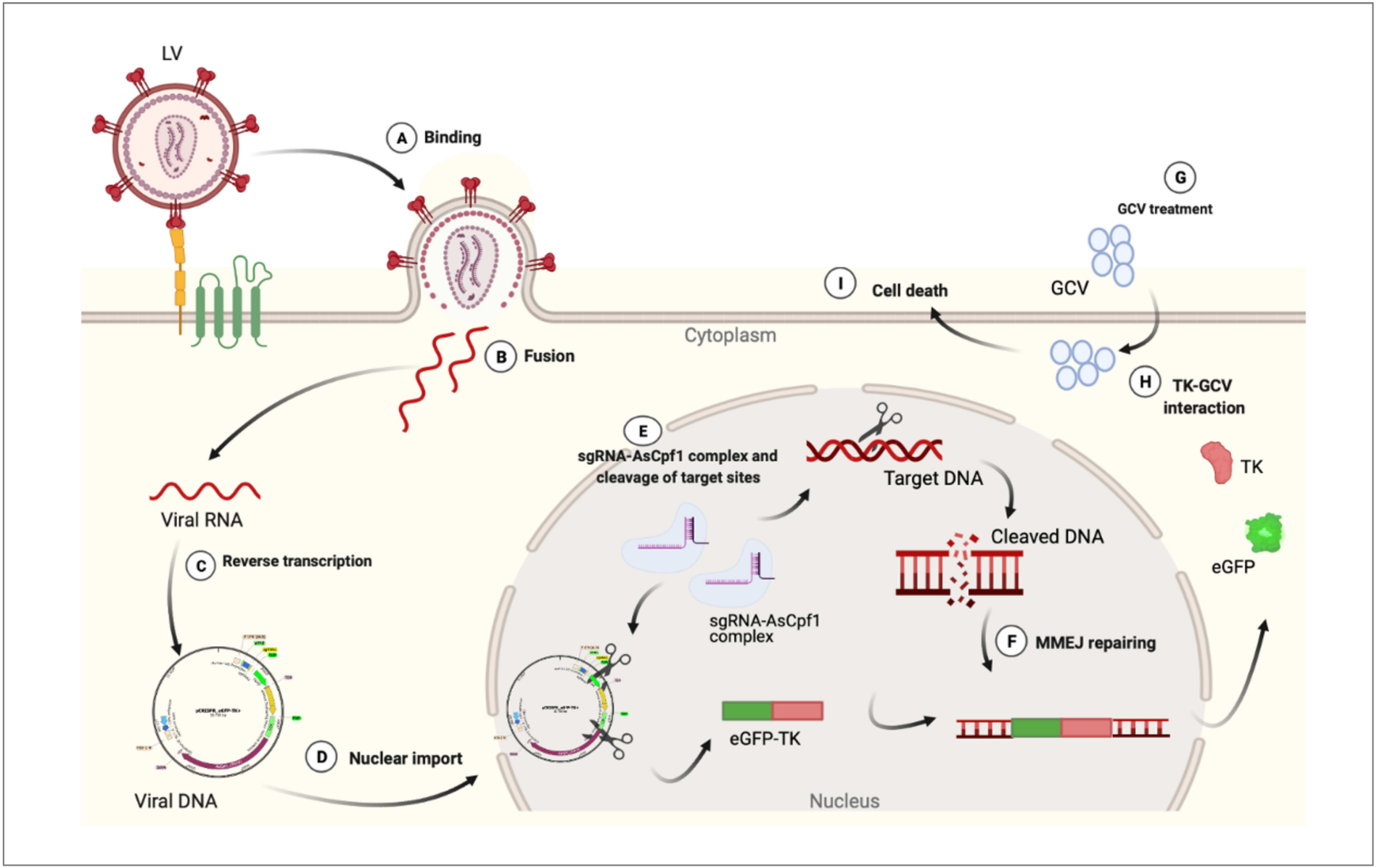

LV
A Binding
B Fusion
Cytoplasm
I Cell death
G
GCV treatment
GCV
H TK-GCV interaction
TK
eGFP
Viral RNA
C Reverse transcription
E
sgRNA-AsCpf1 complex and cleavage of target sites
Target DNA
Cleaved DNA
sgRNA-AsCpf1 complex
F MMEJ repairing
D Nuclear import
Viral DNA
eGFP-TK
Nucleus

Figure 06. Strategy of CRISPR_LV_TK+ system employed for the editing of TP53-mutated target cells. A-B) LV vector harboring the CRISPR_LV_TK+ system within its genome, binds and releases viral RNA into target mutated cells. C) Reverse transcription is initiated to convert viral RNA into viral DNA. D) The viral DNA is imported into the nucleus. E) sgRNAs and AsCpf1 complex together and move toward the target sites (viral DNA and target genomic DNA) to perform site specific cleavage. DSBs are generated at the target DNA and eGFP-TK sequence is cut out from the viral genome. F) Cleaved DNA is repaired by MMEJ resulting in the integration of the eGFP-TK sequence at the target site. G) Ganciclovir is added to the edited cells. H-I) GCV-TK interaction leads target mutated cells to apoptosis.

Lentiviral vector transfer plasmid design

Three different LV transfer plasmids are employed in our experiments. The pCRISPR_eGFP-TK+ is composed by the *Acidaminococcus sp* (As) Cpf1 endonuclease, two sgRNAs targeting both the target genomic site and the vector genome itself, two PAM sites mandatory for the proper recognition by the AsCpf1 and the HSV-TK gene which is linked by a 2A peptide sequence to the eGFP gene sequence. The two PAM sites are located at the edge of the eGFP-TK cDNA sequence which is cut out from the vector after AsCpf1 activity (figure 07 A). However, the AsCpf1 employed in the study is the variant produced by F. Zhang group[88]. Indeed, it is able to bind an alternative PAM site (TATV) which was the only PAM sequence available for our target mutation. Moreover, since the eGFP-TK construct lacks a promoter, both the proteins are encoded only when correctly integrated at the target mutated site on the cell genome by taking advantage of the *TP53* promoter. eGFP is used as reporter protein, thus target cells will turn green when the eGFP-TK construct is properly integrated. The second designed LV is the pCRISPR_eGFP-TK- (figure 07 B). It carries the same features as the pCRISPR_eGFP-TK+ but it lacks the eGFP-TK cassette. The last LV employed in all the performed analysis is the pCMV_eGFP-TK+ (figure 07 C). It misses the CRISPR system but it carries the eGFP-TK sequence which is constitutively expressed by the CMV promoter. The

pCMV_eGFP-TK+ works as a positive control during transfection and transduction experiments. At first the plasmids were used in transient transfection experiments, later the LV were produced.

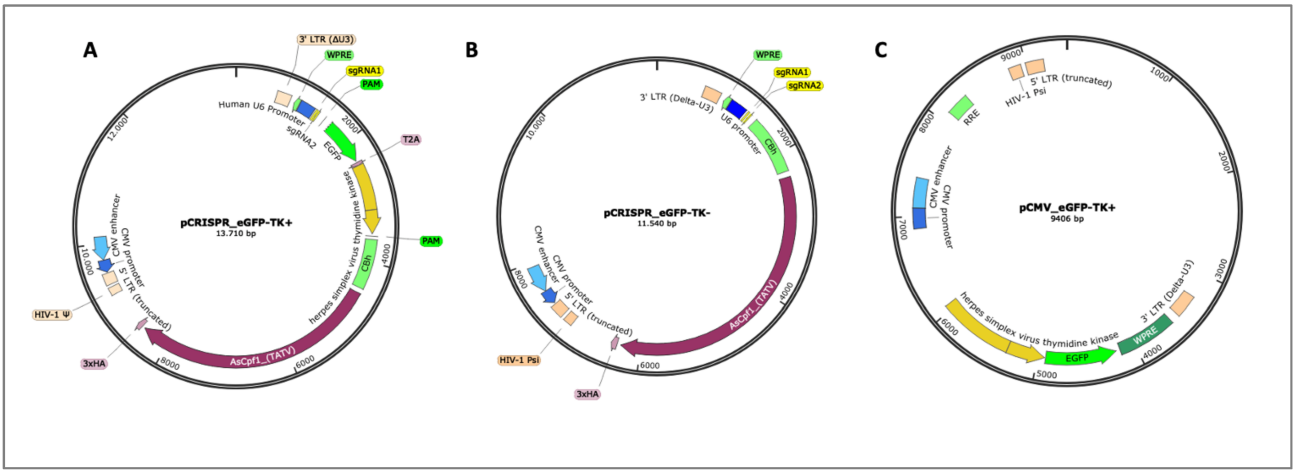


Figure 07. Cloned and employed LV transfer plasmids structures.
A) All the CRISPR components have been cloned into a LV plasmid as a one component system (pCRISPR_eGFP-TK+). The 2 sgRNAs, AsCpf1 (TATV) and the eGFP-TK construct are shown. PAM sites are also present flanking the eGFP-TK construct essential for its proper cleavage by the AsCpf1.
B,C) pCRISPR_eGFP-TK- and pCMV_eGFP-TK+ plasmids employed as negative and positive control, respectively. The former carries the CRISPR components but it lacks the eGFP-TK construct and the PAM sites flanking the construct. The latter lacks the CRISPR components and the PAM sites, but both the eGFP and TK genes are constitutively expressed due to the presence of the CMV promoter.

Validation of specific editing using transfection of reporter 293T target cells

We started testing the efficiency of our CRISPR_LV_TK+ system on a stable cell line. Indeed, HEK293T cells were engineered, as previously described[89], by randomly introduced in their genome a cassette bearing the target *TP53* mutation (p.Ser183*) within the mCherry and blasticidin resistance gene sequences (figure 08). HEK293T were stably transfected with a linear plasmid (pBML5, from Conticello's group, ISPRO, Florence, Italy) harboring the mCherry-bsr cassette.
Blasticidin resistant clones were selected and expanded. FACS analysis to test for mCherry signal was performed to validate the proper clonal selection. Clone number 20 was selected for the next described experiments since it was the one giving a highest percentage (70%) of mCherry positive

cells (Figure S01). We will refer to the engineered cells as HEK#20. Wild type HEK293T were also employed as control cell line.

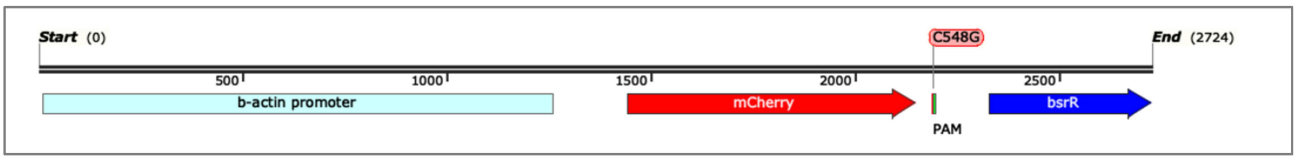


Figure 08. mCherry-bsr cassette. Showed is the cassette designed to engineered the HEK293T cells with the target mutation (c.548C>G; p.Ser183). Blasticidin S-resistance (bsrR) gene is used for cell selection. mCherry is employed as fluorescent marker to validate the proper integration of the cassette into the selected and expanded engineered clones. B-actin is used as promoter for both mCherry and bsrR. The target mutation is represented as c.C548G within the mCherry and bsrR gene sequences. The PAM site essential for the proper AsCpf1(TATV) cleavage is depicted close to the target mutation.*

In a first step, we aimed to assess the potential of the engineered system, using transient transfection of all components. Both HEK#20 and HEK293T were transfected with the three different designed plasmids (pCRISPR_eGFP-TK+; pCRISPR_eGFP-TK-; pCMV_eGFP-TK+) using Lipofectamine. On day 3 post transfection, cells transfected with pCRISPR_eGFP-TK+ were analyzed and sorted by FACS sorting (BD FACSAria) in order to collect and expand only positive clones which have properly integrated the eGFP-TK cassette. 2.5% of the cell population showing eGFP signal, were sorted and expanded for 2 weeks (figure 09 A). However, we could not sort the

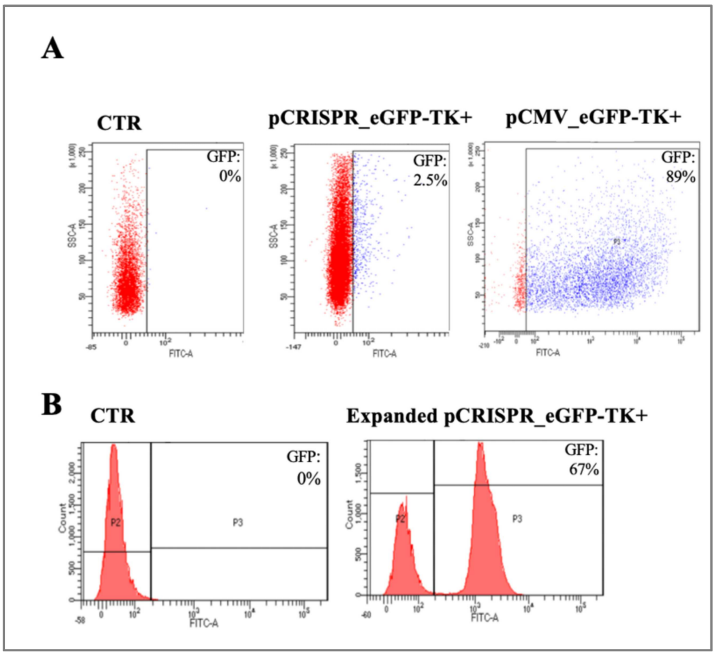


Figure 09. FACS sorter analysis and cell expansion. A) Transfected HEK#20 using pCRISPR_eGFP-TK+ were FACS sorted. eGFP expression was detected in the 2.5% of the cell population receiving pCRISPR_eGFP-TK+, positive control (pCMV_eGFP-TK+) showed a 89% of eGFP expression. In the non- transfected cells (CTR) no eGFP positive cells were detected. B) FACS analysis on the expanded HEK#20. The 67% of the cell population was eGFP positive confirming the proper integration and transmission of the eGFP-TK construct to daughter cells.

HEK#20 receiving the control pCMV_eGFP-TK+ due to the transient eGFP expression which would have been lost a few days after cell expansion. Once the sorted eGFP positive clones were expanded to confluency, a second FACS analysis was performed to check the stable insertion of our construct. We detected 67% of cells to be positive for eGFP expression after cell expansion confirming the proper stable integration of the eGFP-TK cassette which was transmitted to daughter cells and validating the high specificity of our system (figure 09 B).

We hypothesized that even if the percentage of eGFP+ cells and thus the integration of the cassette was low, indels may have been generated at the target site. Hence, a new transfection with pCRISPR_eGFP-TK+ was carried out on HEK#20 following the same procedure described above. Total DNA was extracted 3 days post transfection and amplified using specific primers designed on the target mutation. The amplicons were sequenced by Sanger sequencing and the ICE analysis tool (by Synthego) was ran following the protocol. Indeed, the ICE analysis tool is a straightforward system employed to know the amount of indels generated at the target site. It is essential just to upload the Sanger sequencing data of the edited and the control/wild-type samples together with the sgRNA sequence used in the experiment. ICE algorithms will then match the sgRNA to the wildtype sequence and to the edited samples calculating the amount of indels generated at the target locus.

As expected, 36% of indels were generated at the target site underscoring that the AsCpf1(TATV) is able to specifically cleave the target region but the integration of the cassette remains low (figure 10). This may be due to either the length of the eGFP-TK sequence (2kb) or due to the fact that cells prefer repairing DSB by NHEJ mechanism instead of HDR or MMEJ.

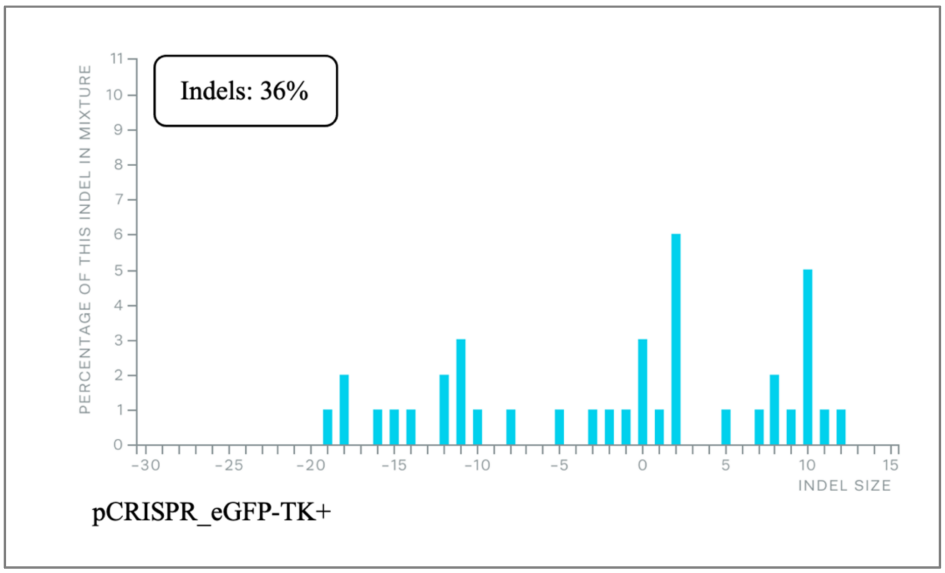


Figure 10. ICE analysis system. The ICE tool was employed to detect indels generated at the target site of HEK#20 previously transfected with pCRISPR_eGFP-TK+. 36% of indels were generated at the target site.

Optimization of Ganciclovir kill curve

Based on the study from Z.-H.Chen et al.[90], we tested different Ganciclovir (Citovirax 500mg Roche) concentrations and we read the results at different time points to find the most suitable conditions by looking at cell death using light microscope. GCV concentrations of 0.01 ug/ul ; 0.1 ug/ul; 1 ug/ul; 10 ug/ul; 100 ug/ul were tested on HEK#20 transfected with the control plasmid pCMV_eGFP-TK+. Media was changed before adding fresh GCV and results were analyzed at 24hr, 48hr and 72hr post GCV treatment. GCV concentration of 0.1ug/ul read at 72hr post treatment was selected as the best condition and applied in our further analysis.

Sorted and expanded HEK#20, HEK293T previously transfected with pCRISPR_eGFP-TK+, pCRISPR_eGFP-TK- and with pCMV_eGFP-TK+ and HEK#20 transfected with pCRISPR_eGFP-TK- and with pCMV_eGFP-TK+ , were seeded in triplicate in a 48-well plate. All samples were treated with GCV [0.1ug/ul] for 72 hrs and cells were monitored by light microscope to detect any sign of apoptosis. On day 3 post GCV treatment FACS analysis was performed to compare a decrease in cell viability between GCV treated and untreated samples. We judged cell viability by calculating

the percentage of alive and death cells gated based on FCS and SSC parameters. As shown in figure 11, HEK#20 holding the target mutation and transfected with pCRISPR_eGFP-TK+ showed a 80% decrease in cell viability when treated with GCV, whereas control HEK293T cells transfected with pCRISPR_eGFP-TK+ did not show any decrease in cell viability when comparing GCV treated and untreated samples. Moreover, a comparable apoptotic pattern was detected between HEK#20 and HEK293T transfected with pCMV_eGFP-TK+ and treated with GCV which both showed a 50% decrease in cell viability (compare panel A and B in Figure 10). As expected, we did not observe any change in cell count in our control samples. Indeed, both HEK#20 and HEK293T transfected with pCRISPR_eGFP-TK- and the un-transfected control samples (CTR) showed no differences in cell viability after GCV administration. Hence, these results demonstrate the high specificity of the system which is able to lead to cell death only mutated cells carrying the suicide gene HSV-TK.

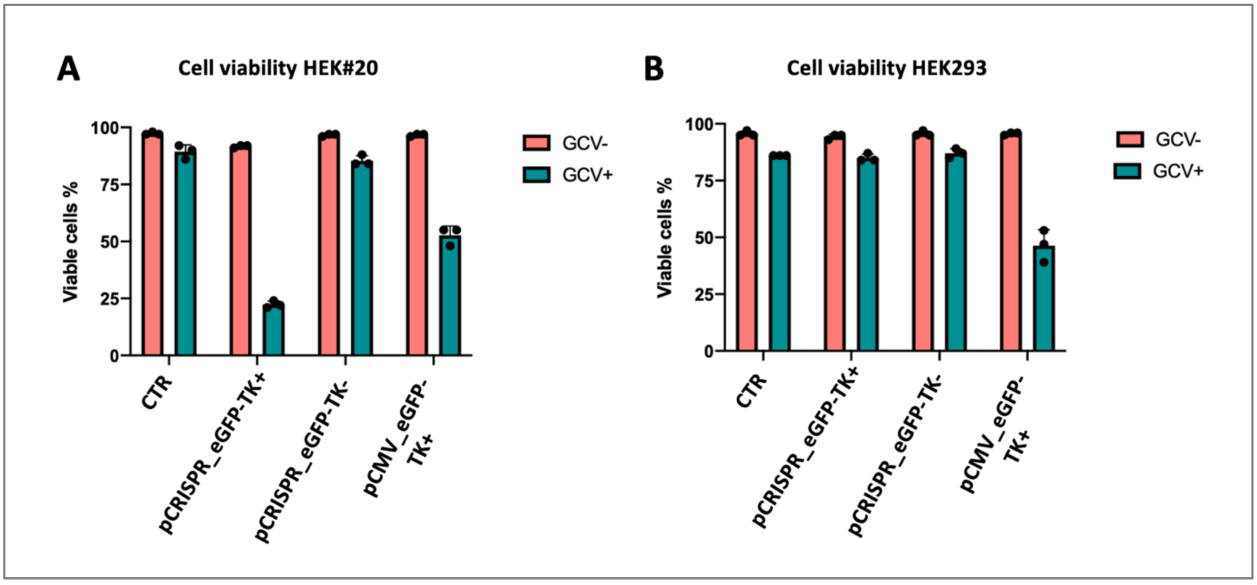


Figure 11. Cell viability assay post GCV treatment. Sorted and expanded HEK#20 (A) and HEK293T (B) transfected with pCRISPR_eGFP-TK+, pCRISPR_eGFP-TK- and pCMV_eGFP-TK+ and no transfected controls (CTR) were FACS analyzed 3 days post GCV [0.1ug/ul] treatment. A decrease in cell viability was detected in the samples of cells bearing the HSV-TK gene and treated with GCV. The HEK#20 transfected with the pCRISPR_eGFP-TK+ and with pCMV_eGFP-TK+ show a 80% and 50% of mortality, respectively. Wild type HEK293T transfected with the pCMV_eGFP-TK+ when treated with GCV showed 50% of cell mortality. No cells death was detected in the sample of wild type HEK293T transfected with the pCRISPR_eGFP-TK+ due to the lack of the target mutation. No cell death was detected in control samples (CTR and pCRISPR_eGFP-TK-) for both HEK#20 and HEK293T. Shown are experimental triplicates.

DNA analysis corroborates correct recombination of the eGFP-TK cassette

Since we only obtained a low percentage of GFP positive cells, we wanted to confirm that the recombination occurred in frame with the cassette, and not non-specific somewhere at random in the target cell genome. Therefore, we designed a specific primer set and performed a PCR on the DNA extracted from HEK#20 which were previously transfected with pCRISPR_eGFP-TK+ to check the correct integration of the eGFP-TK construct at the target site (figure 12 A). As control we used DNA extracted from HEK#20 transfected with pCRISPR_eGFP-TK-. Primers were designed in order to amplify the region of the cassette between mCherry and the integrated eGFP (figure 12 B).

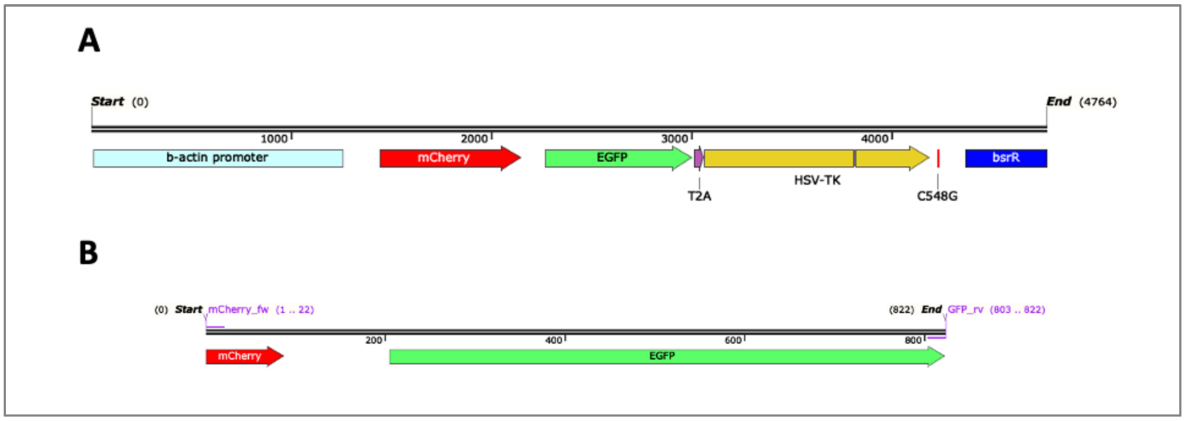


Figure 12. eGFP-TK sequence integration and validation. A) A schematic view of the eGFP-TK sequence integrated within the cassette mCherry-bsr present into HEK#20 cell line. The target mutation c.C548G is highlighted. B) Amplicon generated by the primers set: mCherry fw and GFP_rv (in pink) used to validate the proper integration of the eGFP-TK at the target mutated site.

We observed the correct band of 822bp only in the pCRISPR_eGFP-TK+ sample, whereas, no amplification was detected in the control sample (figure 13).

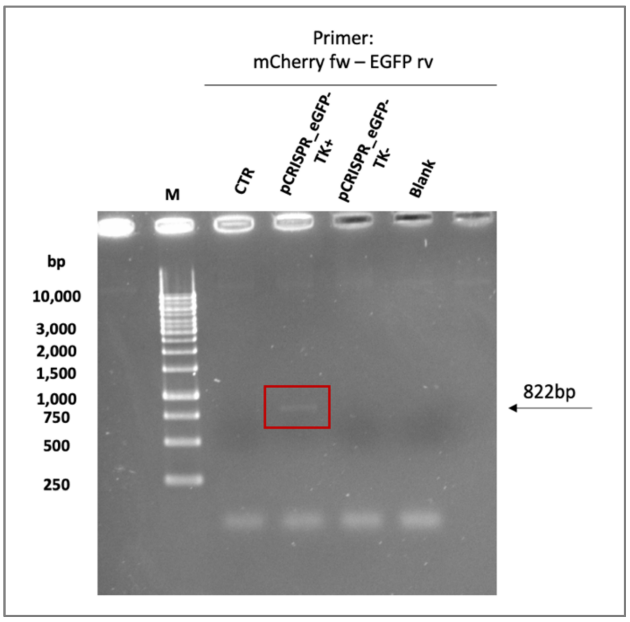


Figure 13. DNA extracted from HEK#20 previously transfected with pCRISPR_eGFP-TK+ and pCRISPR_eGFP-TK- was amplified by PCR to check the correct integration of the eGFP-TK construct at the target site. Primers were designed to amplify the region between mCherry and eGFP. The correct band of 822bp was detected only in the HEK#20 receiving the pCRISPR_eGFP-TK+. CTR: no transfected HEK#20.

Lentiviral vector (LV) production and target cell transduction

Next step of the study was the production of lentiviral vectors and testing their efficiency *in vitro*. Our designed LV vectors belong to the third generation of LV. For production, HEK293T were transfected with either transfer plasmid (pCRISPR_eGFP-TK+, pCRISPR_eGFP-TK- and pCMV_eGFP-TK+), combined with the helper plasmid (PAX) and the env plasmid (VSV-G) mandatory for an efficient vector generation. Vector production was performed by following the protocol from the Gijsbers's team. We will refer to the newly produced LV vectors as: LV_CRISPR_eGFP-TK+, LV_CRISPR_eGFP-TK- and LV_CMV_eGFP-TK+. However, following transduction of HEK#20 with the respective LV vectors, we did not observe any fluorescence for any of the vector preps, whereas we detected eGFP+ cells in the samples transduced with positive control LV_CMV_eGFP-TK+ (Figure 14 A). p24 assay was performed to assess correct vector production: 1.27x10^8, 1.78x10^7, 3x10^6, 9.89x10^7 pg p24/ml corresponding to LV_CRISPR_eGFP-TK+, LV_CRISPR_eGFP-TK- and LV_CMV_eGFP-TK+, respectively.

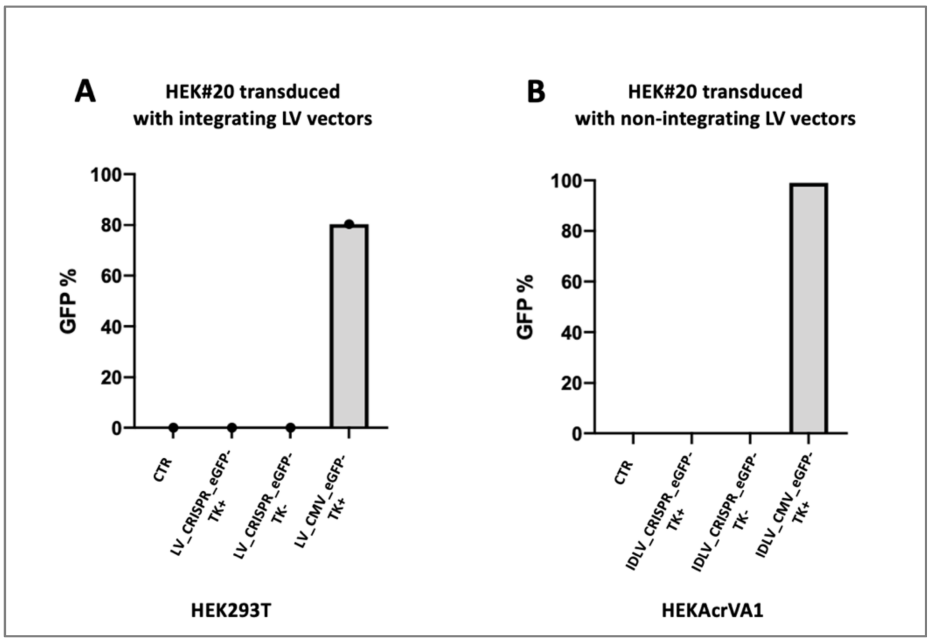


Figure 14. FACS analysis on transduced HEK#20. A) LV_CRISPR_eGFP-TK+, LV_CRISPR_ eGFP-TK- and LV_CMV_eGFP-TK+ viral vectors generated by co-transfection of PAX and VSV-G plasmids on HEK293T were transduced into HEK#20. Only the sample receiving the positive control LV_CMV_eGFP-TK+ vector show 80% of GFP expression. No eGFP expression was detected in the sample transduced with LV_CRISPR_eGFP-TK +. B) IDLV_CRISPR_eGFP-TK+, IDLV_CRISPR_eGFP-TK- and IDLV_CMV_eGFP-TK+ generated by co-transfection of psPAX2-D64V (mutant integrase) and VSV-G plasmids on HEKAcrVA1 (expressing inhibitors of Cpf1) were transduced into HEK#20. 99% of eGFP+ cells were detected in the sample transduced with IDLV_CMV_eGFP-TK+. HEK20 transduce with IDLV_CRISPR_eGFP-TK+ showed no eGFP expression. No treated HEK#20 (CTR) and HEK#20 transduced with LV_CRISPR_ eGFP-TK- or IDLV_CRISPR_eGFP-TK- show no eGFP expression in both conditions.

A possible explanation for the lack of eGFP positive transduced cells could be related to the transfer plasmid design. For safety reasons we developed a self-inactivating pCRISPR_eGFP-TK+ which limits the expression of both AsCpf1 and sgRNAs decreasing the chance of generating off-targets. Indeed, as described above, the sgRNA-AsCpf1 will complex together during vector production in 293T cells, the AsCpf1 will be able to cleave the LV transfer plasmid itself at the target site, allowing the excision of the eGFP-TK cassette and leading to the generation of an open plasmid which would be degraded by the cell, resulting to the production of an empty, no functional viral particles.

To solve this issue we started a second vector production using a specific type of HEK293T called HEKAcrVA1 developed by Doudna research group[91]. This cell line constitutively expresses

inhibitors of the Cpf1, thus our theory was that by using the HEKAcrVA1 for vector production the activity of the AsCpf1 was inhibited during LV production, leading to unaltered plasmids that should be able to generate functional virions. In addition, a mutated PAX plasmid (psPAX2-D64V) was chosen which bears the mutation in the integrase coding sequence, thus generating non-integrating lentiviral vectors, also referred to as integration-deficient LV (IDLV). The produced viral vectors IDLV_CRISPR_eGFP-TK+, IDLV_CRISPR_eGFP-TK- and IDLV_CMV_eGFP-TK+ were used to transduce HEK#20. However, also using this approach, 3 days post transduction no eGFP+ cells were detected except for the samples receiving control IDLV_CMV_eGFP-TK+ vector (Figure 14 B). We next decided to analyze AsCpf1 expression by Western blot (WB) analysis in order to understand if the lack of eGFP+ cells in the samples transduced with LV_CRISPR_eGFP-TK+ or IDLV_CRISPR_eGFP-TK+ was due to issues related to vector production, thus no AsCpf1 expression could be detected by WB, or the problem was the lack of eGFP-TK integration at the target site.

Proteins were extracted from both HEK#20 transduced with the LV_CRISPR_eGFP-TK+, LV_CRISPR_eGFP-TK- and LV_CMV_eGFP-TK+ and from HEKAcrVA1 transfected with pCRISPR_eGFP-TK+, pCRISPR_eGFP-TK-, pCMV_eGFP-TK+ plasmids. We used antibody against the 3-HA-tag linked to the AsCpf1 sequence. As shown in the figure 15, HA signal was detected only in the CTNS-3HA used as positive control and in the HEKAcrVA1 transfected cells.

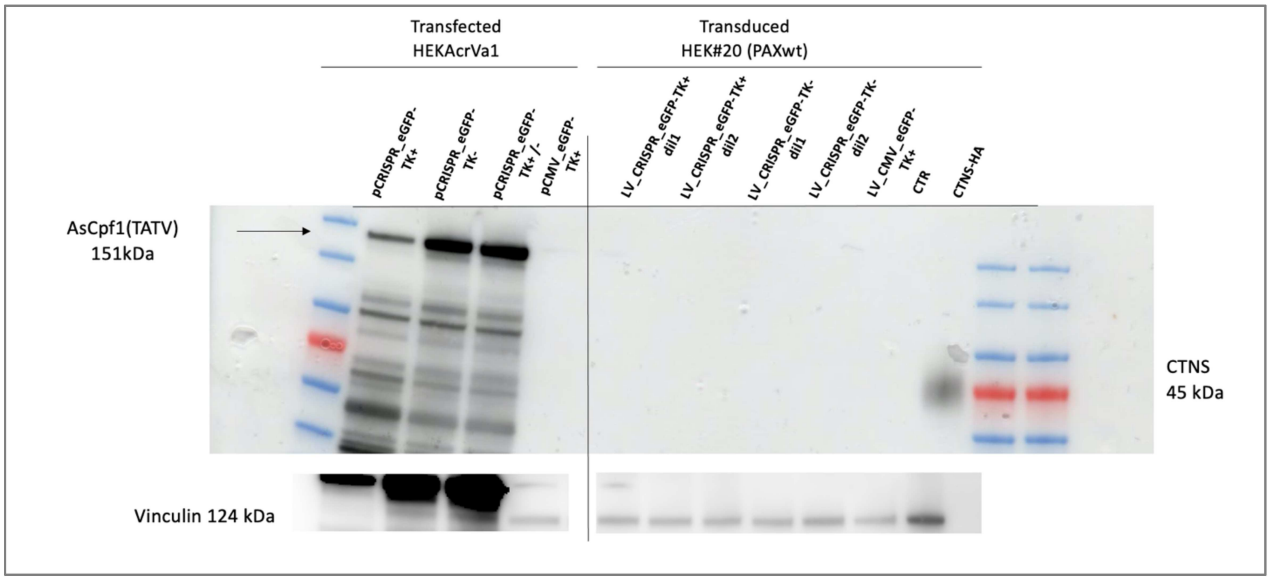


Figure 15. Western blot to test AsCpf1 expression. Protein extracted from both HEKAcrVa1 transfected with pCRISPR_eGFP-TK+, pCRISPR_eGFP-TK-, a combination of pCRISPR_eGFP-TK+ and pCRISPR_eGFP-TK- and from HEK#20 transduced with LV_CRISPR_eGFP-TK+, LV_CRISPR_ eGFP-TK- with the two highest viral vectors dilutions (dil1, dil2). Controls are: HEK#AcrVA1 and HEK#20 respectively transfected and transduced with pCMV_eGFP-TK+ / LV_CMV_eGFP-TK+ and no treated HEK#20 (CTR). The CTNS_HA-Tag is the positive control.

However, we decided to perform a qPCR analysis to assess number of integrated proviral vector copies, amplifying a region of the plasmid in common between all the employed vectors (psi region) to test that the lack of AsCpf1 expression was due to an inefficient transduction which impeded the proper integration of the vector in the host genome. DNA was extracted from wild-type HEK293T transduced with LV_CRISPR_eGFP-TK+, LV_CRISPR_eGFP-TK- and LV_CMV_eGFP-TK+. As expected, the qPCR data revealed that only the LV_CMV_eGFP-TK+ was correctly integrated into the host genomes, both the LV_CRISPR_eGFP-TK+ and LV_CRISPR_eGFP-TK- were completely absent from the cellular genome (Figure 16). These results helped us understand that the problem lied within the plasmid structure which for some reason inhibited vector production.

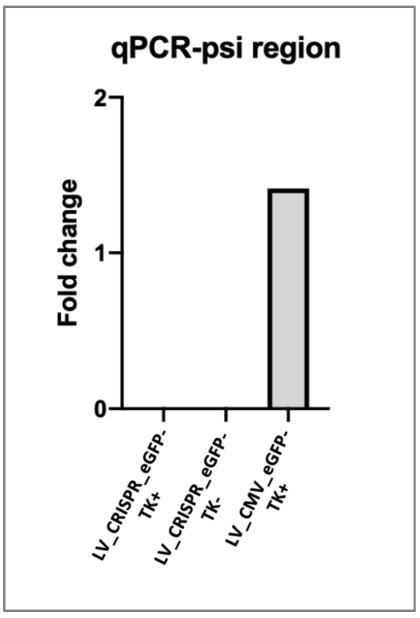


Figure 16. qPCR on the integrated psi-region. DNA extracted from HEK293T transduced with LV_CRISPR_eGFP-TK+, LV_CRISPR_eGFP-TK- and LV_CMV_eGFP-TK+ was analyzed by qPCR. Only LV_CMV_eGFP-TK+ was properly integrated into the host genome. No amplification was detected for LV_CRISPR_eGFP-TK+ and LV_CRISPR_eGFP-TK-.

Studying carefully the pCRISPR_eGFP-TK+ plasmid map we noticed two very important features. First of all, the RRE and WPRE sequences important during vector production were missing from the plasmid. Second, the Cbh promoter and the AsCpf1(TATV) sequence are in anti-sense relative to the LTRs. The latter point may compromise vector production since the mRNA sequence transcribed from the LTRs and the AsCpf1-mRNA driven by Cbh promoter are complementary to each other and could hybridize, thereby blocking the packaging signal and thus limiting the incorporation of the genomic RNA in the nascent viral vector particle. This may result in the generation of an empty and no functional viral vectors. In order to understand if the orientation of the endonuclease was the issue, we engineered a new cloning strategy aimed at cutting out from the pCRISPR_eGFP-TK+ the AsCpf1(TATV) sequence in order to test the correct production of viral vectors. By using ClaI as restriction enzyme we were able to excise the Cbh promoter and part of the AsCpf1(TATV). We will refer to the newly cloned plasmid as pCRISPR_eGFP-TK+ _AsCpf1- and to the produced LV vector as LV_CRISPR_eGFP-TK+_AsCpf1- (figure 17).

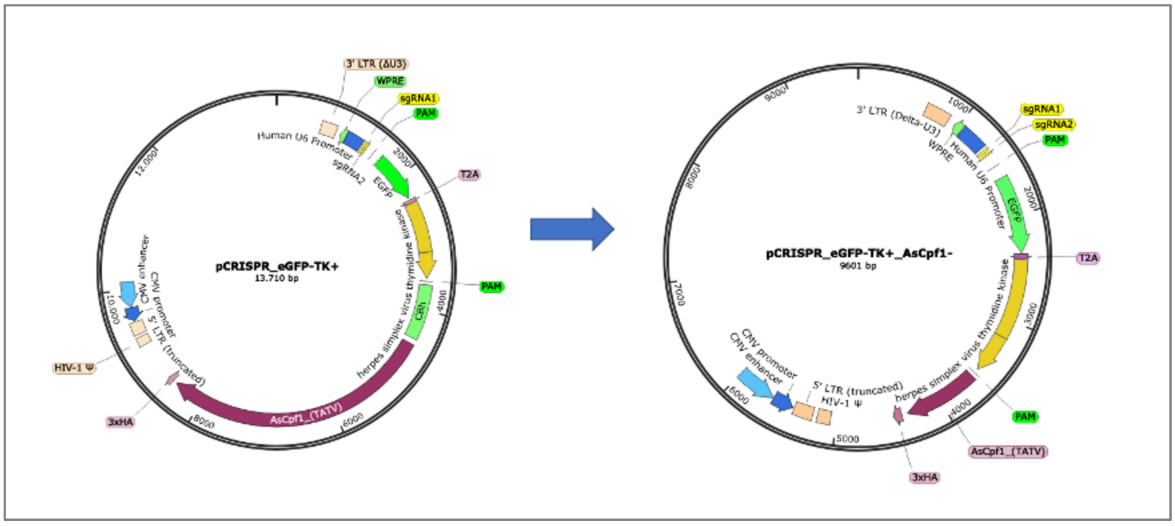


Figure 17. pCRISPR_eGFP-TK+_AsCpf1- design. ClaI restriction enzyme was used to remove the Cbh promoter and AsCpf1 (TATV) sequences from the pCRISPR_eGFP-TK+. The plasmid still carries the U6-sgRNAs sequences, the PAM flanking the EGFP-TK cassette and the cassette itself.

Afterwards, we ran a qPCR analysis to test the correct production of the vector lacking the AsCpf1(TATV). Wild-type HEK293T were transduced with LV_CRISPR_eGFP-TK+_AsCpf1- and

DNA extracted following cell expansion. qPCR revealed integrated copies for both vector productions, indicating that we were able to produce functional virions which correctly integrated into the host genome (Figure 18). However, the plasmid still lacks the aforementioned sequences, RRE and WPRE which are important during vector production to allow the nuclear export of the newly generated viral RNA sequence. From these results we concluded that the correct orientation of the AsCpf1 is mandatory for the production of functional viral vectors. Moreover, both RRE and WPRE sequences may be important to generate a high titer.

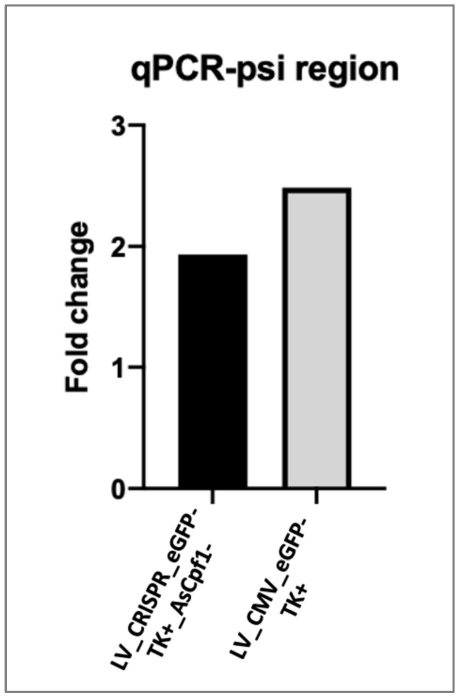


Figure 18. qPCR on the integrated psi-region. DNA extracted from HEK293T transduced with LV_CRISPR_eGFP-TK+_AsCpf1- was analyzed by qPCR. The graph showed the proper amplification of the psi-region in both the samples. Both LV_CRISPR_eGFP-TK+_AsCpf1- and LV_CMV_eGFP-TK+ (used as positive control) were correctly integrated in target HEK293T cell lines.

Even though we could use the plasmids to underscore the functionality of our set-up, we were not able to demonstrate functionality using viral vectors. We re-engineered the lentiviral vector transfer plasmid and the cloning is currently ongoing. We had to start all over again in order to reverse the orientation of the AsCpf1(TATV) which should be in frame with the LTRs and to have a final plasmid carrying both the RRE and WPRE sequences.

In vivo study: establishment of human CLL xenograft mouse model

In vitro results using plasmids showed the ability of the system to bring target mutated cells to apoptosis, thus we moved forward to an in vivo model in order to test the feasibility of the system in NOD-SCID IL2Rγnull mice. All the mice studies were performed in collaboration with Toscana Life Science (TLS). Based on previous studies[92,93] animals were intravenously (i.v.) injected with primary PBMC isolated from CLL patients' blood samples. One day before injection, mice were treated with Busulfan (25mg/Kg). On the day of injection, a vial of PBMC from a CLL patient was thawed and cells were counted to have a total of 1x10^8 cells in 100ul. A small fraction of the thawed cells was analyzed by flow cytometry to quantify the amount of T and B cells (both wt and neoplastic) present in the sample (figure 19). Indeed, it is known from the literature that B-CLL cells have better chances to engraft and proliferate in NSG mice when T cells comprised the 1-15% of the total injected PBMC[92,93].

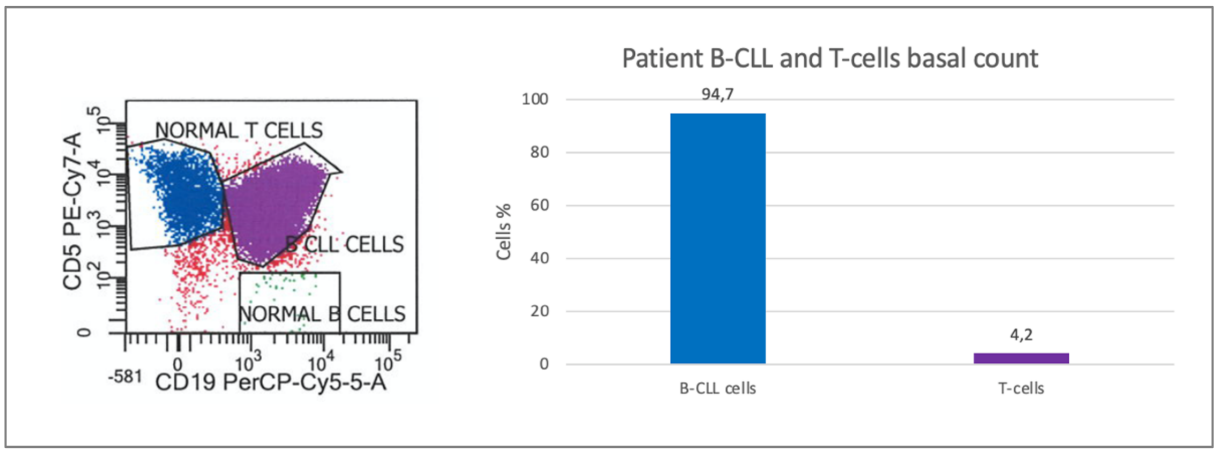


Figure 19. Flow cytometry analysis on CLL patient's PBMC used for i.v. mice injection. A fraction of PBMC injected into the mice were previously analyzed to quantify the amount of B neoplastic cells (B-CLL cells, in blue) and T-cells (in purple). 94,7% and 4,2 % of B-CLL and T-cells were present within the analyzed PBMC population.

Mice (n=2) were monitored every week for 4 weeks and blood samples were collected and analyzed by flow cytometry. In order to detect human T and B cells the following antibodies were employed during flow cytometry analysis: $CD45^+$, $CD5^+$,$CD19^-$ (T cells), $CD45^+$,$CD5^+$,$CD19^+$(B-CLL cells),

$CD45^{+}$,$CD5^{-}$,$CD19^{+}$ (B cells). On week 3 about 20% of T and B neoplastic cells were detected in mice blood samples (figure 20 A).

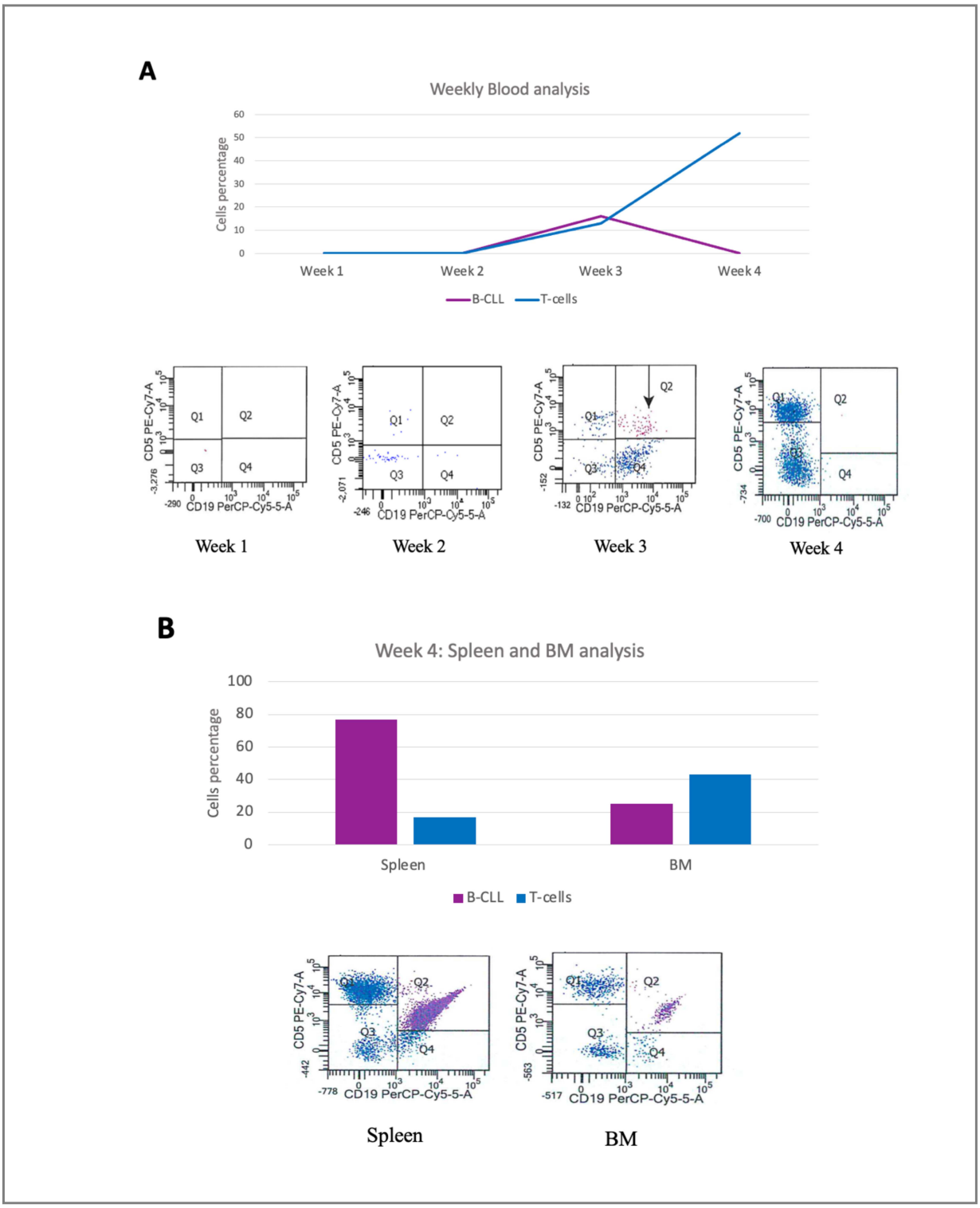


Figure 20. CLL xenograft mice analysis. A) Flow cytometry analysis on blood samples collected every week for 4 weeks from xenograft mice. B-CLL and T cells were detected on week 3. An increase in T cells and a drop in B-CLL cells count was observed on week 4. B) Flow cytometry analysis on spleen and bone marrow (BM) organs collected on week 4 from sacrificed xenograft mice. B-CLL and T cells were detected in both organs showing the successful engraftment.

However, on week 4 a drop in B-CLL cells and an increase in T cells was observed, still in line with what previously observed[93]. Moreover, on week 4 post injection mice were sacrificed and organs collected. As expected, B neoplastic lymphocytes were detected in mice lymphoid organs such as spleen and bone marrow (figure 20 B). This is also in line with the disease course in which neoplastic B cells accumulate in lymphoid organs.

Finally, the results showed that we were able to generate a CLL xenograft mouse model which will be employed to study LV vector transduction followed by GCV treatment.

Discussion

Cancer remains always a very hard player to defeat. This is mostly due to the high heterogeneity among cancer types that can originate from different kind of deleterious mutations. Specifically, CLL can be caused by either chromosomal or genetic alterations. The worst scenario is represented by patients carrying a mutation in the *TP53* gene or/and a deletion in the corresponding chromosome 17p. By now two are the most effective treatments for CLL patients holding a mutant p53 or a del17p, a combination of chemoimmunotherapy and target therapies such as ibrutinib and venetoclax or the development of CAR-T. Unfortunately, these therapies cannot be considered as universal treatments for CLL patients who still show relapse and a short overall survival when treated with CIT and target therapies or important side effects could arise in patients treated with CAR-T (e.g. cytokine release syndrome).

The aim of my research was the development of a personalized therapy which can be beneficial for CLL patients and hopefully it will be applied to other critical cancer types. The developed tested CRISPR_LV_TK+ system is an all-in-one lentiviral vector carrying the CRISPR-AsCpf1 tool and the suicide gene HSV thymidine kinase (HSV-TK). The HSV-TK is harmless by itself but it is able to lead cells to apoptosis by phosphorylating the antiviral prodrug Ganciclovir (GCV). Indeed, cells are treated with GCV which once activated by the TK will inhibit the DNA replication ultimately leading cells to death. TK was selected as suicide gene for the developed system due to the affordability and availability of GCV which is usually employed in order to treat patients with cytomegalovirus (CMV) infection. Indeed, other suicide genes are employed for gene therapy such as the inducible Caspase9 (iCasp9) described by Straathof et al. (2005)[94]. However, iCasp9 gets activated after the interaction with the synthetic drug AP1903 which is more expensive compared to Ganciclovir (AP1903 (Bio-techne): 325 euro for 5mg vs GCV (Roche): 57.67 euro for 500mg).

While I am finalizing the lentiviral vectors production, an *in vivo* model is ready to test the feasibility of the system. Indeed, we were able to properly generate a CLL xenograft mice model starting from NOD-scid IL2Rγnull mice. Experiments are on-going to investigate both viral infection and the optimal GCV dosage. To study lentiviral transduction we will first check for GFP expression in T and B cells. Indeed, xenograft mice will be i.v. injected with a control LV vector [10^8 IU/ml] encoding only GFP driven by CMV internal promoter. The transduction will take place the day after PBMC injection and mice will be monitored every week for 4 weeks when mice are finally sacrificed. Whereas, to determine the best GCV dosage, mice will receive different concentrations of the drug at different time points. However, this last experiment still needs to be fully discussed, for now the utmost importance is to investigate the infection efficiency.

Conclusion and Perspectives

With the presented study, I wanted to show the potential, strength and efficiency of our developed CRISPR_LV_TK+ system which was recently patented with the international publication number: WO2020/079574. The system based on the combination of CRISPR-AsCpf1 technology and HSV-TK gene delivery followed by ganciclovir administration revealed its potential when tested on target mutated cells. Indeed, a 80% decrease in cell viability was detected only in mutated cells carrying the HSV-TK and treated with GCV. The obtained results made me confident about the great specificity of the CRISPR_LV_TK+ system which could be applied to several cancer types characterized by different mutations in different cancer driver genes. Thanks to the advance in the NGS analysis we are now able to detect mutations in genes that can be drivers of cancer progression which can become the perfect target for the described and tested gene editing tool. The developed system is versatile and can be easily employed for multiplexing, theoretically targeting multiple mutations. It would be enough to just change a single sgRNA within the LV vector to adapt the system to target different mutations. We are now developing the system to target 3 additional *TP53* variants detected in 2 CLL patients (NM_000546.5(TP53):c.C637T:p.R213*, NM_000546.5(TP53):c.675delT:p.V225fs) and one patient with thyroid cancer (NM_000546.5(TP53):c.742C>G:p.Arg248Gly). I truly believe the CRISPR_LV_TK+ tool is a possible solution to get rid of mutated cancer cells thus bringing hopes for a possible personalized cancer therapy.

Material and Methods

Cell culture

HEK293T and HEK#20 are grown in high glucose Dulbecco's Modified Eagle Medium (DMEM), 10% Fetal Bovine Serum (FBS) and 1% L- glutamine. Engineered HEK#20 are produced by stable transfection of wild-type HEK293T with a plasmid (pBML5) harboring the target mutation, the mCherry ORF sequence and blasticidin S-resistance (bsr) gene sequences to select positive clones, as reported in a previous work[89]. PBMC are isolated from CLL patients' blood samples by Pancoll solution. Freshly isolated PBMC are either incubated at 37°C and 5% CO2 or stored in liquid nitrogen. Primary and stable cell lines were incubated at 37°C and 5% CO2. Experiments were conducted in a laminar flow hood.

Plasmids cloning

The plasmid employed in the study is cloned using a third generation (SIN) lentiviral vector backbone. Two sgRNAs are selected from MIT website (http://crispr.mit.edu) and ordered from Addgene. The plasmid pCRISPR_eGFP-TK+, derived from two plasmids: pBML4 and pAsCpf1(TATV)(BB) (pY221) purchased from Addgene. First the sgRNAs and the U6 promoter were cloned into pBML4. The AsCpf1 from the pAsCpf1(TATV)(BB) (pY221) was added into the pBML4 as well. All the sequence U6, sgRNAs, Cph and AsCpf1 were later cloned into a lentiviral plasmid from Conticello's lab (ISPRO, Florence, Italy). Finally, the construct eGFP-TK together with the PAM were cloned into the lentiviral plasmid. The main features are: U6 promoter for sgRNAs expression and Cbh promoter for AsCpf1(TATV) protein expression cloned within the 3'LTR and 5'LTR. The eGFP-TK construct does not have a promoter.

pCMV_eGFP-TK+ is a gift from Conticello's group (ISPRO, Florence, Italy).

pCRISPR_eGFP-TK+ is digested with ClaI and the homology arms are ligated back together in order to delete Cbh and AsCpf1(TATV) from the plasmid. The resulting plasmid is called: pCRISPR_eGFP-TK+_AsCpf1-.

Cell transfection and ganciclovir treatment

Both Lipofectamine2000 and Lipofectamine3000 are used to transfect HEK293T and HEK#20 with pCRISPR_eGFP-TK+, pCRISPR_eGFP-TK- and pCMV_eGFP-TK+. The transfection with Lipofectamin2000 is done by following the protocol by ThermoFisher. 8x10^4 cells are counted the day of transfection. 1ug of DNA and 1ul of Lipofectamine2000 is selected for transfection in a final Optimem volume of 50ul. The mix is incubated for 20 minutes and aliquot into each sample. Cells are plated into p24 well plate. Lipofectamine 3000 reagent protocol (by Invitrogen ThermoFisher) was followed in order to select the proper DNA concentration, Lipofectamine and P3000 amount. 4x10^5 and 1x10^4 cells were plated a day before transfection into a p6 or p96 well plate, respectively.

Ganciclovir (GCV) (Citovirax 500 mg, Roche) is added to positively transfected HEK293T, HEK#20 and control samples with a final concentration of 0.1 ug/ml. Cells receive one dosage of GCV each day for 3 days. Medium is removed each day prior GCV treatment.

Viral vector production and cell transduction

All the LV vectors used in the study are produced following the protocol from the group of Prof. R. Gijsbers (Leuven Viral Vector Core and Molecular Virology and Gene Therapy Laboratory at the KU Leuven, BE). Briefly, HEK293T are transfected with LV transfer plasmid, PAX plasmid and VSV-G plasmid using PEI as transfection method. Medium is changed 24hr post transfection. Virus is collected on day 2 and 3 post transfection using a 45uM filter. On day 3 post transfection collected virus is concentrated by Vivaspin concentration method. Viral vector production was assessed by p24 ELISA. Produced viral vector was stored at -80°C.

HEK293T and HEK#20 cells are plated in a p96 well plate a day before transduction. A 3-fold serial dilution of the LV vectors is performed. 3 days post transduction cells are FACS analyzed and the correct TU is calculated for each produced LV vector.

FACS analysis and PCR analysis

Transfected and transduced HEK293T and HEK#20 are FACS analyzed to test the expression of eGFP. Ganciclovir treated cells are FACS analyzed 3 days post treatment to evaluate the percentage of cell death. The BD FACSAria™ is used to sort eGFP positive cells. PCR is perform to evaluate the proper integration of the construct eGFP-TK into target HEK#20 after transfection with pCRISPR_eGFP-TK+. Primers are designed on the mCherry sequence and on the eGFP sequence: (FW) 5'- CTCCCACAACGAGGACTACACC-3', (RV) 5'- GACTGGGTGCTCAGGTAGTG-3'.

qPCR analysis

qPCR is performed on gDNA extracted from transduced HEK293T in order to test the correct integration of the produced viral vectors. Specific primers were designed (Table I). β-actin gene was the selected housekeeping gene. All data was later analyzed by calculating the Ct mean, the Delta Ct = Ct gene test – Ct housekeeping gene and finally the ΔΔCt = ΔCt sample – ΔCt control.

Table I: qPCR primer list used to test viral vector production.

Primer name	*Primer sequence*
Psi_fw	5'-CTGTGCGGTGGTCTTACTTT-3'
Psi_rv	5'-GGACAGCTACAACCATCCCT -3'

Western Blot analysis

Proteins are extracted from expanded HEK#20 and HEKAcrVa1 respectively transduced and transfected with p/LV_CRISPR_eGFP-TK+, p/LV_CRISPR_eGFP-TK-, both

p/LV_CRISPR_eGFP-TK+/TK- and p/LV_CMV_eGFP-TK+. Cells pellet is resuspended in 80ul of 1% SDS, left for 10 min at 95°C and later sonicated. BCA assay (ThermoFisher) is performed in order to quantify the proteins concentration and to load an equal amount of protein for immunoblotting. Antibody used: anti-HA and anti-Vinculin.

CLL Xenograft mice model

Mice (male, 8-10 weeks) NOD-scid IL2Rγnull (NSG) are purchased from CHARLES RIVER LABORATORIES ITALIA s.r.l. and kept in the animal facilities of Toscana Life Science (TLS) in Siena, Italy. Animals are kept in isolation for 2 weeks before the starting of the study. 100 ul of blood samples are collected in EDTA from the animals a day before the starting of the experiments. Mice are i.v. injected with Busulfan [25mg/Kg] (Myleran, Busulfex, 50mg) in a final volume of 500ul. 1x10^8 PBMCs from CLL patient are thaw and counted the day of injection. The final pellet is resuspended in PBS in a final volume 100ul ready to be injected in each animal. Every week for 4 weeks blood sample (100ul EDTA) is collected from the mice and quickly analyzed by flow cytometry using human antibodies: CD45+,CD5+,CD19- (T cells), CD45+,CD5+,CD19+(B-CLL cells), CD45+,CD5-,CD19+ (B cells). On week 4, mice are sacrificed and organs collected. Spleen, BM and lymph nodes are homogenized by mechanically cutting the tissues by scalpel and needles (21G and 23G) in order to isolate cells for flow cytometry analysis. Tissues are stored in optimal cutting temperature (OCT) compound at -80°C for further analysis.

References

1. Aitken, M. J. L., Lee, H. J. & Post, S. M. Emerging treatment options for patients with p53-pathway-deficient CLL. *Ther. Adv. Hematol.* **10**, 204062071989135 (2019).
2. Muñoz-Novas, C. *et al.* The International Prognostic Index for Patients with Chronic Lymphocytic Leukemia Has the Higher Value in Predicting Overall Outcome Compared with the Barcelona-Brno Biomarkers Only Prognostic Model and the MD Anderson Cancer Center Prognostic Index. *BioMed Res. Int.* **2018**, 1–8 (2018).
3. Sutton, L.-A. & Rosenquist, R. Deciphering the molecular landscape in chronic lymphocytic leukemia: time frame of disease evolution. *Haematologica* **100**, 7–16 (2015).
4. Rozovski, U., Keating, M. J. & Estrov, Z. Why Is the Immunoglobulin Heavy Chain Gene Mutation Status a Prognostic Indicator in Chronic Lymphocytic Leukemia? *Acta Haematol.* **140**, 51–54 (2018).
5. Malcikova, J. *et al.* Detailed analysis of therapy-driven clonal evolution of TP53 mutations in chronic lymphocytic leukemia. *Leukemia* **29**, 877–885 (2015).
6. Gaidano, G. & Rossi, D. The mutational landscape of chronic lymphocytic leukemia and its impact on prognosis and treatment. *Hematology* **2017**, 329–337 (2017).
7. Moia, R. *et al.* Precision Medicine Management of Chronic Lymphocytic Leukemia. *Cancers* **12**, 642 (2020).
8. Cohen, J. A. *et al.* An Updated Perspective on Current Prognostic and Predictive Biomarkers in Chronic Lymphocytic Leukemia in the Context of Chemoimmunotherapy and Novel Targeted Therapy. *Cancers* **12**, 894 (2020).
9. Campo, E. *et al. TP53* aberrations in chronic lymphocytic leukemia: an overview of the clinical implications of improved diagnostics. *Haematologica* **103**, 1956–1968 (2018).
10. Döhner, H. *et al.* Genomic Aberrations and Survival in Chronic Lymphocytic Leukemia. *N. Engl. J. Med.* **343**, 1910–1916 (2000).
11. Gentile, M. *et al.* Validation of the CLL-IPI and comparison with the MDACC prognostic index in newly diagnosed patients. *Blood* **128**, 2093–2095 (2016).
12. Rossi, D. *et al.* The Prognostic Value of TP53 Mutations in Chronic Lymphocytic Leukemia Is Independent of Del17p13: Implications for Overall Survival and Chemorefractoriness. *Clin. Cancer Res.* **15**, 995–1004 (2009).
13. Aubrey, B. J., Strasser, A. & Kelly, G. L. Tumor-Suppressor Functions of the TP53 Pathway. *Cold Spring Harb. Perspect. Med.* **6**, a026062 (2016).
14. Harris, S. L. & Levine, A. J. The p53 pathway: positive and negative feedback loops. *Oncogene* **24**, 2899–2908 (2005).
15. on behalf of the European Research Initiative on Chronic Lymphocytic Leukemia (ERIC) — TP53 network *et al.* ERIC recommendations for TP53 mutation analysis in chronic lymphocytic leukemia—update on methodological approaches and results interpretation. *Leukemia* **32**, 1070–1080 (2018).
16. Kosicki, M. *et al.* Dynamics of Indel Profiles Induced by Various CRISPR/Cas9 Delivery Methods. in *Progress in Molecular Biology and Translational Science* vol. 152 49–67 (Elsevier, 2017).
17. Petitjean, A. *et al.* Impact of mutant p53 functional properties on *TP53* mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum. Mutat.* **28**, 622–629 (2007).
18. Tate, J. G. *et al.* COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **47**, D941–D947 (2019).
19. Pinto, A. M. *et al.* Low-level *TP 53* mutational load antecedes clonal expansion in chronic lymphocytic leukaemia. *Br. J. Haematol.* **184**, 657–659 (2019).
20. Landau, D. A. *et al.* Evolution and Impact of Subclonal Mutations in Chronic Lymphocytic Leukemia. *Cell* **152**, 714–726 (2013).

21. Rossi, D. *et al.* Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. **123**, 10 (2014).
22. Stilgenbauer, S., Schnaiter, A., Paschka, P., Zenz, T. & Rossi, M. Gene mutations and treatment outcome in chronic lymphocytic leukemia: results from the CLL8 trial. **123**, 8 (2014).
23. Boddy, C. S. & Ma, S. Frontline Therapy of CLL: Evolving Treatment Paradigm. *Curr. Hematol. Malig. Rep.* **13**, 69–77 (2018).
24. A New Era is Coming up in the Treatment of Chronic Lymphocytic Leukemia. *Lymphoma Chronic Lymphocytic Leuk.* **4**, 9–19 (2014).
25. Alduaij, W. & Illidge, T. M. The future of anti-CD20 monoclonal antibodies: are we making progress? *Blood* **117**, 2993–3001 (2011).
26. Boross, P. & Leusen, J. H. W. Mechanisms of action of CD20 antibodies. 15.
27. Smith, M. R. Rituximab (monoclonal anti-CD20 antibody): mechanisms of action and resistance. *Oncogene* **22**, 7359–7368 (2003).
28. Thompson, P. A. *et al.* Fludarabine, cyclophosphamide, and rituximab treatment achieves long-term disease-free survival in IGHV-mutated chronic lymphocytic leukemia. *Blood* **127**, 303–309 (2016).
29. Kater, A. P. *et al.* Fixed Duration of Venetoclax-Rituximab in Relapsed/Refractory Chronic Lymphocytic Leukemia Eradicates Minimal Residual Disease and Prolongs Survival: Post-Treatment Follow-Up of the MURANO Phase III Study. *J. Clin. Oncol.* **37**, 269–277 (2019).
30. June, C. H., O'Connor, R. S., Kawalekar, O. U., Ghassemi, S. & Milone, M. C. CAR T cell immunotherapy for human cancer. *Science* **359**, 1361–1365 (2018).
31. Mancikova, V. *et al.* Performance of anti-CD19 chimeric antigen receptor T cells in genetically defined classes of chronic lymphocytic leukemia. *J. Immunother. Cancer* **8**, e000471 (2020).
32. Dreger, P. *et al.* Managing high-risk CLL during transition to a new treatment era: stem cell transplantation or novel agents? *Blood* **124**, 3841–3849 (2014).
33. Sorror, M. L. *et al.* Five-Year Follow-Up of Patients With Advanced Chronic Lymphocytic Leukemia Treated With Allogeneic Hematopoietic Cell Transplantation After Nonmyeloablative Conditioning. *J. Clin. Oncol.* **26**, 4912–4920 (2008).
34. Hahn, M. *et al.* Allogeneic hematopoietic stem cell transplantation for poor-risk CLL: dissecting immune-modulating strategies for disease eradication and treatment of relapse. *Bone Marrow Transplant.* **50**, 1279–1285 (2015).
35. Fedoroff, N. V. McClintock's challenge in the 21st century. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 20200–20203 (2012).
36. Fernández, A., Josa, S. & Montoliu, L. A history of genome editing in mammals. *Mamm. Genome* **28**, 237–246 (2017).
37. Unniyampurath, U., Pilankatta, R. & Krishnan, M. RNA Interference in the Age of CRISPR: Will CRISPR Interfere with RNAi? *Int. J. Mol. Sci.* **17**, 291 (2016).
38. Urnov, F. D. *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646–651 (2005).
39. Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143–148 (2011).
40. Smith, J. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res.* **28**, 3361–3369 (2000).
41. Li, T. *et al.* TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.* **39**, 359–372 (2011).
42. Mahfouz, M. M. *et al.* De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc. Natl. Acad. Sci.* **108**, 2623–2628 (2011).
43. Carroll, D. Genome Engineering With Zinc-Finger Nucleases. *Genetics* **188**, 773–782 (2011).

44. Boch, J. *et al.* Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science* **326**, 1509–1512 (2009).
45. Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).
46. Mojica, F. J. M., D□ez-Villase□or, C., Garc□a-Mart□nez, J. & Soria, E. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *J. Mol. Evol.* **60**, 174–182 (2005).
47. Barrangou, R. *et al.* CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science* **315**, 1709–1712 (2007).
48. Wright, A. V., Nuñez, J. K. & Doudna, J. A. Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. *Cell* **164**, 29–44 (2016).
49. Nishimasu, H. *et al.* Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **156**, 935–949 (2014).
50. Wu, X., Kriz, A. J. & Sharp, P. A. Target specificity of the CRISPR-Cas9 system. *Quant. Biol.* **2**, 59–70 (2014).
51. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
52. Fonfara, I., Richter, H., Bratovič, M., Le Rhun, A. & Charpentier, E. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* **532**, 517–521 (2016).
53. Świat, M. A. *et al.* FnCpf1: a novel and efficient genome editing tool for Saccharomyces cerevisiae. *Nucleic Acids Res.* **45**, 12585–12598 (2017).
54. Zetsche, B. *et al.* Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* **163**, 759–771 (2015).
55. Fagerlund, R. D., Staals, R. H. J. & Fineran, P. C. The Cpf1 CRISPR-Cas protein expands genome-editing tools. *Genome Biol.* **16**, 251 (2015).
56. Wyman, C. & Kanaar, R. DNA Double-Strand Break Repair: All's Well that Ends Well. *Annu. Rev. Genet.* **40**, 363–383 (2006).
57. Ma, Y., Zhang, L. & Huang, X. Genome modification by CRISPR/Cas9. *FEBS J.* **281**, 5186–5193 (2014).
58. Gaj, T., Gersbach, C. A. & Barbas, C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397–405 (2013).
59. Wyman, C., Ristic, D. & Kanaar, R. Homologous recombination-mediated double-strand break repair. *DNA Repair* **3**, 827–833 (2004).
60. Lee, J., Chung, J.-H., Kim, H. M., Kim, D.-W. & Kim, H. Designed nucleases for targeted genome editing. *Plant Biotechnol. J.* **14**, 448–462 (2016).
61. Taleei, R. & Nikjoo, H. Biochemical DSB-repair model for mammalian cells in G1 and early S phases of the cell cycle. *Mutat. Res. Toxicol. Environ. Mutagen.* **756**, 206–212 (2013).
62. Lin, D.-W. *et al.* Microhomology-based CRISPR tagging tools for protein tracking, purification, and depletion. *J. Biol. Chem.* **294**, 10877–10885 (2019).
63. Sakuma, T., Nakade, S., Sakane, Y., Suzuki, K.-I. T. & Yamamoto, T. MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. *Nat. Protoc.* **11**, 118–133 (2016).
64. Hsu, P. D., Lander, E. S. & Zhang, F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* **157**, 1262–1278 (2014).
65. Wu, S.-S., Li, Q.-C., Yin, C.-Q., Xue, W. & Song, C.-Q. Advances in CRISPR/Cas-based Gene Therapy in Human Genetic Diseases. *Theranostics* **10**, 4374–4382 (2020).
66. Nelson, C. E. *et al.* In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* **351**, 403–407 (2016).
67. Tabebordbar, M. *et al.* In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* **351**, 407–411 (2016).

68. Wu, Y. *et al.* Highly efficient therapeutic gene editing of human hematopoietic stem cells. *Nat. Med.* **25**, 776–783 (2019).
69. Maeder, M. L. *et al.* Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10. *Nat. Med.* **25**, 229–233 (2019).
70. Halim, L. & Maher, J. CAR T-cell immunotherapy of B-cell malignancy: the story so far. *Ther. Adv. Vaccines Immunother.* **8**, 251513552092716 (2020).
71. Stadtmauer, E. A. *et al.* CRISPR-engineered T cells in patients with refractory cancer. *Science* **367**, eaba7365 (2020).
72. Glover, D. J., Lipps, H. J. & Jans, D. A. Towards safe, non-viral therapeutic gene expression in humans. *Nat. Rev. Genet.* **6**, 299–310 (2005).
73. Gori, J. L. *et al.* Delivery and Specificity of CRISPR/Cas9 Genome Editing Technologies for Human Gene Therapy. *Hum. Gene Ther.* **26**, 443–451 (2015).
74. Giacca, M. & Zacchigna, S. Virus-mediated gene delivery for human gene therapy. *J. Controlled Release* **161**, 377–388 (2012).
75. Pauwels, K. *et al.* State-of-the-Art Lentiviral Vectors for Research Use: Risk Assessment and Biosafety Recommendations. *Curr. Gene Ther.* **9**, 459–474 (2009).
76. Merten, O.-W., Hebben, M. & Bovolenta, C. Production of lentiviral vectors. *Mol. Ther. - Methods Clin. Dev.* **3**, 16017 (2016).
77. Naldini, L. *et al.* In Vivo Gene Delivery and Stable Transduction of Nondividing Cells by a Lentiviral Vector. *Science* **272**, 263–267 (1996).
78. Vink, C. A. *et al.* Eliminating HIV-1 Packaging Sequences from Lentiviral Vector Proviruses Enhances Safety and Expedites Gene Transfer for Gene Therapy. *Mol. Ther.* **25**, 1790–1804 (2017).
79. Schmidt, F. & Grimm, D. CRISPR genome engineering and viral gene delivery: A case of mutual attraction. *Biotechnol. J.* **10**, 258–272 (2015).
80. Cockrell, A. S., van Praag, H., Santistevan, N., Ma, H. & Kafri, T. The HIV-1 Rev/RRE system is required for HIV-1 5' UTR cis elements to augment encapsidation of heterologous RNA into HIV-1 viral particles. *Retrovirology* **8**, 51 (2011).
81. Funke, S. *et al.* Targeted Cell Entry of Lentiviral Vectors. *Mol. Ther.* **16**, 1427–1436 (2008).
82. Frecha, C. *et al.* Efficient and stable transduction of resting B lymphocytes and primary chronic lymphocyte leukemia cells using measles virus gp displaying lentiviral vectors. *Blood* **114**, 3173–3180 (2009).
83. Levy, C. *et al.* Baboon envelope pseudotyped lentiviral vectors efficiently transduce human B cells and allow active factor IX B cell secretion *in vivo* in NOD/SCIDγc $^{-/-}$ mice. *J. Thromb. Haemost.* **14**, 2478–2492 (2016).
84. Girard-Gagnepain, A. *et al.* Baboon envelope pseudotyped LVs outperform VSV-G-LVs for gene transfer into early-cytokine-stimulated and resting HSCs. *Blood* **124**, 1221–1231 (2014).
85. Sereni, L. *et al.* Lentiviral gene therapy corrects platelet phenotype and function in patients with Wiskott-Aldrich syndrome. *J. Allergy Clin. Immunol.* **144**, 825–838 (2019).
86. Aiuti, A. *et al.* Lentiviral Hematopoietic Stem Cell Gene Therapy in Patients with Wiskott-Aldrich Syndrome. *Science* **341**, 1233151–1233151 (2013).
87. Diez, B. *et al.* Therapeutic gene editing in CD 34 $^{+}$ hematopoietic progenitors from Fanconi anemia patients. *EMBO Mol. Med.* **9**, 1574–1588 (2017).
88. Gao, L. *et al.* Engineered Cpf1 variants with altered PAM specificities. *Nat. Biotechnol.* **35**, 789–792 (2017).
89. Niccheri, F., Pecori, R. & Conticello, S. G. An efficient method to enrich for knock-out and knock-in cellular clones using the CRISPR/Cas9 system. *Cell. Mol. Life Sci.* **74**, 3413–3423 (2017).
90. Chen, F. *et al.* Targeted activation of diverse CRISPR-Cas systems for mammalian genome editing via proximal CRISPR targeting. *Nat. Commun.* **8**, 14958 (2017).
91. Watters, K. E., Fellmann, C., Bai, H. B., Ren, S. M. & Doudna, J. A. Systematic discovery of natural CRISPR-Cas12a inhibitors. *Science* **362**, 236–239 (2018).

92. Bagnara, D. *et al.* A novel adoptive transfer model of chronic lymphocytic leukemia suggests a key role for T lymphocytes in the disease. *Blood* **117**, 5463–5472 (2011).
93. Herman, S. E. M. *et al.* Modeling tumor–host interactions of chronic lymphocytic leukemia in xenografted mice to study tumor biology and evaluate targeted therapy. *Leukemia* **27**, 2311–2321 (2013).
94. Straathof, K. C. *et al.* An inducible caspase 9 safety switch for T-cell therapy. *Blood* **105**, 4247–4254 (2005).

Clinical trials references

ClinicalTrials.gov Identifier:NCT03960840.

CD19-specific CAR-T Cells in CLL/SLL and DLBCL (2019)

https://clinicaltrials.gov/ct2/show/NCT03960840

ClinicalTrials.gov Identifier:NCT02935257.

Immunotherapy for High Risk/Relapsed CD19+ Acute Lymphoblastic Leukaemia, B-cell Non-Hodgkin's Lymphoma (B-NHL) and Chronic Lymphocytic Leukaemia (CLL)/ Small Lymphocytic Lymphoma (SLL) Using CAR T-cells to Target CD19 (ALLCAR19).

https://clinicaltrials.gov/ct2/show/NCT02935257

ClinicalTrials.gov Identifier: NCT01853631.

Activated T-Cells Expressing 2nd or 3rd Generation CD19-Specific CAR, Advanced B-Cell NHL, ALL, and CLL (SAGAN) (SAGAN). https://clinicaltrials.gov/ct2/show/NCT01853631

ClinicalTrials.gov Identifier: NCT03655678.

A Safety and Efficacy Study Evaluating CTX001 in Subjects With Transfusion-Dependent β-Thalassemia. https://clinicaltrials.gov/ct2/show/NCT03655678

ClinicalTrials.gov Identifier: NCT03872479.

Single Ascending Dose Study in Participants With LCA10.

https://clinicaltrials.gov/ct2/show/NCT03872479

ClinicalTrials.gov Identifier:NCT04035434.

A Safety and Efficacy Study Evaluating CTX110 in Subjects With Relapsed or Refractory B-Cell Malignancies. https://clinicaltrials.gov/ct2/show/NCT04035434

ClinicalTrials.gov Identifier:NCT03399448
NY-ESO-1-redirected CRISPR (TCRendo and PD1) Edited T Cells (NYCE T Cells).
https://clinicaltrials.gov/ct2/show/NCT03399448

ClinicalTrials.gov Identifier:NCT03157804
Lentiviral-mediated Gene Therapy of Fanconi Anemia Patients Subtype A (FANCOLEN-1).
https://clinicaltrials.gov/ct2/show/NCT03157804

Supplementary information

Table SI: sgRNAs sequences.

Name	*Sequence*
sgRNA – *TP53*	5'- TcAGCAGCGCTCATGGTGGGGGC – 3'
sgRNA – universal	5' – TCTGCTAATCCTGTTACCAGCCC – 3'

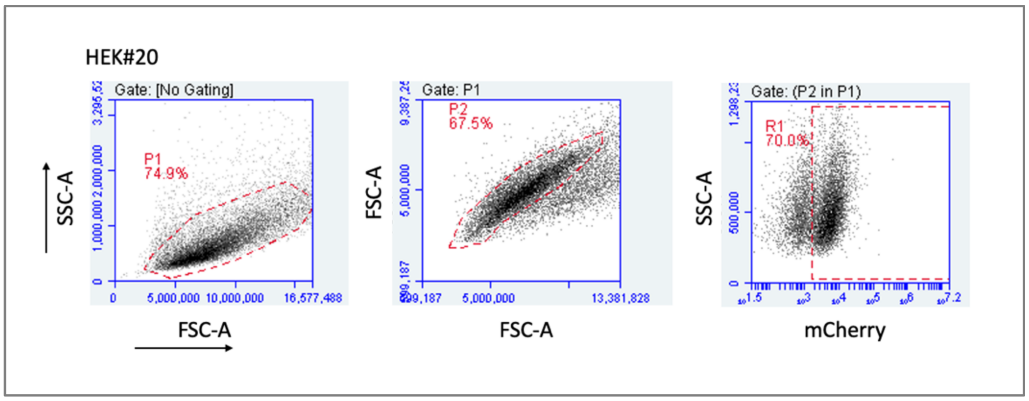


Figure S01. FACS analysis on HEK#20. Validation of selected and expanded HEK293T clone number 20 (HEK#20) by testing for mCherry positive cells. The samples derived from the stable transfection of HEK293T with the pBML5 plasmid carrying the mCherry-bsr cassette. 70% of mCherry positive cells are detected within the expanded cell population.