UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Centre hospitalier
universitaire vaudois

SCHOOL OF MEDICINE – MASTER THESIS

GENOME EDITING OF PRIMARY SYNOVIAL SARCOMA CELLS TO PRODUCE V5-TAGGED FUSION SYT-SSX AND TO ALLOW ITS CRE-MEDIATED SILENCING

Student
Radiana Ferrero

Tutor
Prof. *Ivan Stamenkovic*
Institut Universitaire de Pathologie – CHUV

Co-Tutor
Dr. *Carlo Fusco*
Institut Universitaire de Pathologie – CHUV

Expert
Prof. *Nicolo Riggi*
Institut Universitaire de Pathologie – CHUV

Lausanne, December 2017

Table of content

List of figures 3

List of tables 3

List of abbreviations 4

ABSTRACT 5

1. INTRODUCTION 6

1.1 Biology of Synovial Sarcoma 6

1.2 Genome editing strategies 7

1.3 MGH5 cell background 8

2. AIM OF THE STUDY 8

3. RESULTS 9

3.1 Design of the donor DNA template 9

3.2 Amplification and assembly of the donor DNA into the rAAV vector 12

3.3 CRISPR/Cas9 design and cloning in lentivirus plasmid 14

3.4 Lentivirus production 15

3.5 rAAV production 15

4. DISCUSSION 16

5. MATERIAL & METHODS 19

5.1 MGH5 cell-culture 19

5.2 Viral production 19

5.3 Design and cloning of CRISPR 19

5.4 Design and cloning of Donor 20

5.5 Measurement of viral copy number 20

5.6 Western blot 20

ACKNOWLEDGEMENTS 21

BIBLIOGRAPHY 22

List of figures

Figure 1 Design of gene targeting 9

Figure 2 Design of the Donor template 11

Figure 3 Cre-mediated silencing of SYT-SSX 12

Figure 4 Assembly of the donor plasmid for rAAV production 13

Figure 5 CRISPR/Cas9 design – Lentiviral vector 15

Figure 6 Viruses production 16

List of tables

Table 1 Donor restriction sites 14

List of abbreviations

bFGF	Basic fibroblast growth factor	MIT	Massachusetts's Institute of Technology
ChIP-seq	Chromatin immunoprecipitation and deep sequencing	MOI	Multiplicity of infections
CRISPR	Clustered regularly interspaced short palindromic repeats	mRNA	Messenger ribonucleic acid
DNA	Deoxyribonucleic acid	NeoR	Neomycine resistance
EDTA	Ethylenediaminetetraacetic acid	NHEJ	Non-homologous end joining
EF1α	Elongation factor one alpha	NMD	Non-sense mediated mRNA decay
EGF	Epithelial growth factor	p.	Page
Em	Emerald	PCR	Polymerase chain reaction
ER	Epigenetic regulators	Pen/Strep	Penicillin/Streptomycin
GFP	Green fluorescent protein	Pol	Polymerase
HDAC	Histone deacetylase	PTM	Post transcriptional modifications
HDR	Homologous directed repair	PuroR	Puromycin resistance
HEK293T	Human embryonic kidney 293 for transfection	q-RT-PCR	Quantitative real-time PCR
HR	Homologous recombination	qPCR	Quantitative polymerase chain reaction
IDT	Integrated DNA technologies	rAAV	Recombinant adeno-associated virus
IMDM	Iscove's Modified Dulbecco's Medium	ROI	Region of interest
Indel	Insertions and deletions	SDS-PAGE	Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis
KO	Knock out	sgRNA	Short guide RNA
MGH5	Massachusetts's General Hospital 5	SS	Synovial sarcoma

Abstract

Synovial Sarcoma (SS) is a soft tissue neoplasm characterized by the translocation [t(X;18)(p11;q11)] which produces the aberrant fusion protein SYT-SSX. SS is an aggressive tumour which develops mainly in children and young adults, has a poor prognosis, due to its tendency to metastasize, and is frequently resistant to chemo and radiotherapy. To develop new therapeutic approach is required a better understanding of the molecular mechanisms leading to the disease.

In recent years it has become clear that SYT-SSX profoundly affects the epigenetic of cells, including mesenchymal cells, which are believed to be SS' cells of origin. A very powerful approach to measure the epigenetic status of the chromatin and its relations with the transcriptome is the use of Chromatin Immunoprecipitation and deep sequencing (ChIP-seq). These techniques allow the analysis of both histone Post Transcriptional Modifications (PTM) and Epigenetic Regulators (ER) and require the use of high-affinity and low cross-reactivity antibodies targeting proteins of interest. Such antibodies are not available for SYT-SSX. To allow a ChIP-seq analysis of the latter, we started preliminary work to add to the genomic sequence of SYT-SSX, a V5 peptide tag, for which a ChIP-grade antibody is available.

Furthermore with the same genome editing we allow the Cre-mediated depletion of SYT-SSX while leaving unaffected the wild type counterparts as an important control reagent.

This study has the potential to help understanding the molecular pathogenesis behind SS occurrence and aspires to lay the basis for the development of new therapies against the malignancy.

Keywords: cancer, genome editing, primary tissue culture, Synovial sarcoma, SYT-SSX

1. Introduction

Sarcomas are a broad family of soft tissue neoplasm raised from mesenchymal cells.

Although sarcomas are relatively rare pathologies, accounting for approximately 1% of all malignant neoplasm, they are often lethal diseases and most frequently diagnosed among children and young adults (1). The actual therapeutic approach relies at first line on surgery; however, this is often associated to chemo and radiotherapy as adjuvant treatments. Nevertheless, the rate of early and late recurrences is high, and the survival after 10 years is on the order of 50% (2).

Specific non-random chromosomal translocations of diagnostic interest are found in a number of sarcomas, including Ewing's Sarcoma [*EWSR1-FLI1*, t(11;22)(q24;q12)] and Synovial Sarcoma [*SS18*(SYT)-*SSX*, t(X;18)(p11;q11)] (3). In this subset of sarcomas the chromosomal rearrangements is believed to play an important role in the pathogenesis of the malignancies. Thus, studying in detail the effect of the chimeric protein resulting from these translocations as well as its influence on cell's destiny has been extensively pursued by the scientific community.

To conceive and develop new targeted and efficient therapies, it is important to improve our understanding of sarcoma's molecular pathology.

1.1 Biology of Synovial Sarcoma

We are particularly interested in the Synovial Sarcoma (SS) [t(X;18)(p11;q11), *SS18*(SYT)-*SSX*].

The *SS18* gene encompasses over 7.6 kb of DNA and contains 17 exons that encode a 418 nucleotides of mRNA (4). Its product (SYT) is a nuclear receptor co-activator that plays a role of transcriptional co-activator (5). The *SSX1* gene encompasses over 13.3 kb of DNA and contains 9 exons that encode 188 nucleotides of mRNA (4). The role of its protein is less clear but it is believed to function as a transcriptional repressor (6).

First studies on its fusion protein have shown that neither SYT nor SSX proteins have DNA binding domains however they both possess protein-protein interaction motifs that mediate association with transcriptional regulators and chromatin remodelling complexes, contributing in an epigenetic way to the dysregulation of transcriptional processes (7–10). Most recent evidences suggest that the SYT-SSX fusion protein would play a major role in the maintenance of stemness and poor differentiation by partially activating canonical Wnt pathway in the absence of Wnt ligands (10).

Chromatin Immunoprecipitation coupled with high throughput sequencing (ChIP-seq) analysis of SYT-SSX would be useful to define how and where the protein interacts with the chromatin.

ChIP-seq is a recent laboratory technique (11) that allows to study both the genomic localization of a protein of interest, in this case SYT-SSX, and the epigenetic status of the chromatin which defines the transcription activity of promoter and enhancers (12). However, the technique relies on the use of a high-quality and specific antibody.

SYT-SSX is a protein composed of two endogenous polypeptides fused together, for this reason finding an antibody targeting the translocation rather than the corresponding wild-type protein is difficult.
A solution to this problem would be to design a monoclonal antibody that would bind the chimeric protein right across the two subunits, with one arm of the antibody binding SYT and the other arm binding SSX. This has been unsuccessfully tried in our lab with SYT-SSX, resulting in either a lack of specificity or poor affinity of the antibody.
A second solution would be to edit the sarcoma cells' genome by adding a tag to the endogenous chimeric protein, such as a V5 tag, a short peptide sequence derived from the C-terminus of the P and V proteins of Simian Virus 5 (13,14), and than using the already available highly-specific and efficient antibodies directed against this tag.

1.2 Genome editing strategies

The ability of modifying mammalians cell's genome has been available since the late eighties and has been extensively used to generate genetically modified mouse lines (15). The method takes advantage of the homologous recombination pathway (HR) to introduce a genetic modification by providing the cells with a donor sequence containing the target flanked by two homologous regions. However, because baseline HR events are extremely rare, this method works only in a limited set of cell types, including embryonic stem cells (16). Therefore the ability of modifying the genome in such way is limited to the context of mouse models, and could not be applied to primary and established cell line cultures.
Without a massive improvement of the efficiency of HR between the receiving genome and the donor template, it would have been impossible to modify the genome of cell lines. Fortunately, the technical landscape in the last few years has dramatically changed.
The first improvement has come from the use of recombinant Adeno-Associated Viruses (rAAV) as vectors for the delivery of the donor sequence (17). This is a single-stranded DNA virus with a genome long about 4.7 kb(18). Alone and following a molecular mechanism nowadays not yet well elucidated, the use of rAAV has increased of three orders of magnitude the frequency of baseline HR events (19).
The second break through is the discovery of genome editing tools including CRISPR/Cas9 technology (Clustered Regularly Interspaced Short Palindromic Repeats) (20). The latter allows to specifically damage the genome close to the targeted Region Of Interest (ROI) and therefore to trigger the Homologous Directed Repair (HDR) pathway.
Finally in 2013, a paper pointed out an inhibitor of Non-Homologous End Joining (NHEJ), Sc7 which is able to increase up to 19-fold the frequency of homologous recombination following DNA damage(21). Taken all together, it is now possible to achieve a high frequency of HR events using CRISPR/Cas9 directed DNA damage of the ROI, coupled with rAAV-mediated delivery of a donor template, while inhibiting NHEJ pathway.

In my thesis I will present the use of these strategies for the experimental design and the realization of the tools necessary to V5 tag SYT-SSX1 protein in MGH5 (Massachusetts's General Hospital 5) primary SS cells. As secondary objective, I will describe the Cre-mediated silencing of SYT-SSX1-V5 protein leaving unchanged the wild type counterparts in the same cells.

1.3 MGH5 cell background

MGH5 are SS primary cells that hold the translocation SYT-SSX and were generated in Massachusetts's General Hospital by putting in culture cells dissociated from a SS specimen surgically extracted from a patient. Culture in non-adherence conditions as spheroids allowed to maintain their stem-cells-like properties. By sequencing analysis of their genomic DNA we confirmed that these cells harbour the canonical SS break point (22), which is located between exon X of *SS18* gene, on the long arm of the 18th chromosome, and the exon VI of *SSX1* gene, on the short arm of X chromosome [Fig. 1A].

2. Aim of the study

In the present work, we wanted to modify the genome of MGH5 so that SYT-SSX1 will harbour a V5 tag at its C-terminus end (SYT-SSX1-V5). In addition, we wanted to have the possibility of inducing the depletion of SYT-SSX1-V5 leaving unchanged SYT expression. Because MGH5 are primary cells and would be very difficult to obtain clones from single cells, a high rate of HR was necessary to obtain this genetic modification in the bulk population of cells.

We addressed this problem with the concomitant use of Scr7 drug, the CRISPR-directed DNA damage of the region adjacent to SYT-SSX1 STOP codon [Fig. 1A,B] and the rAAV-mediated delivery of the template harbouring both V5 tag and two LoxP sites [Fig. 1C; 2A].

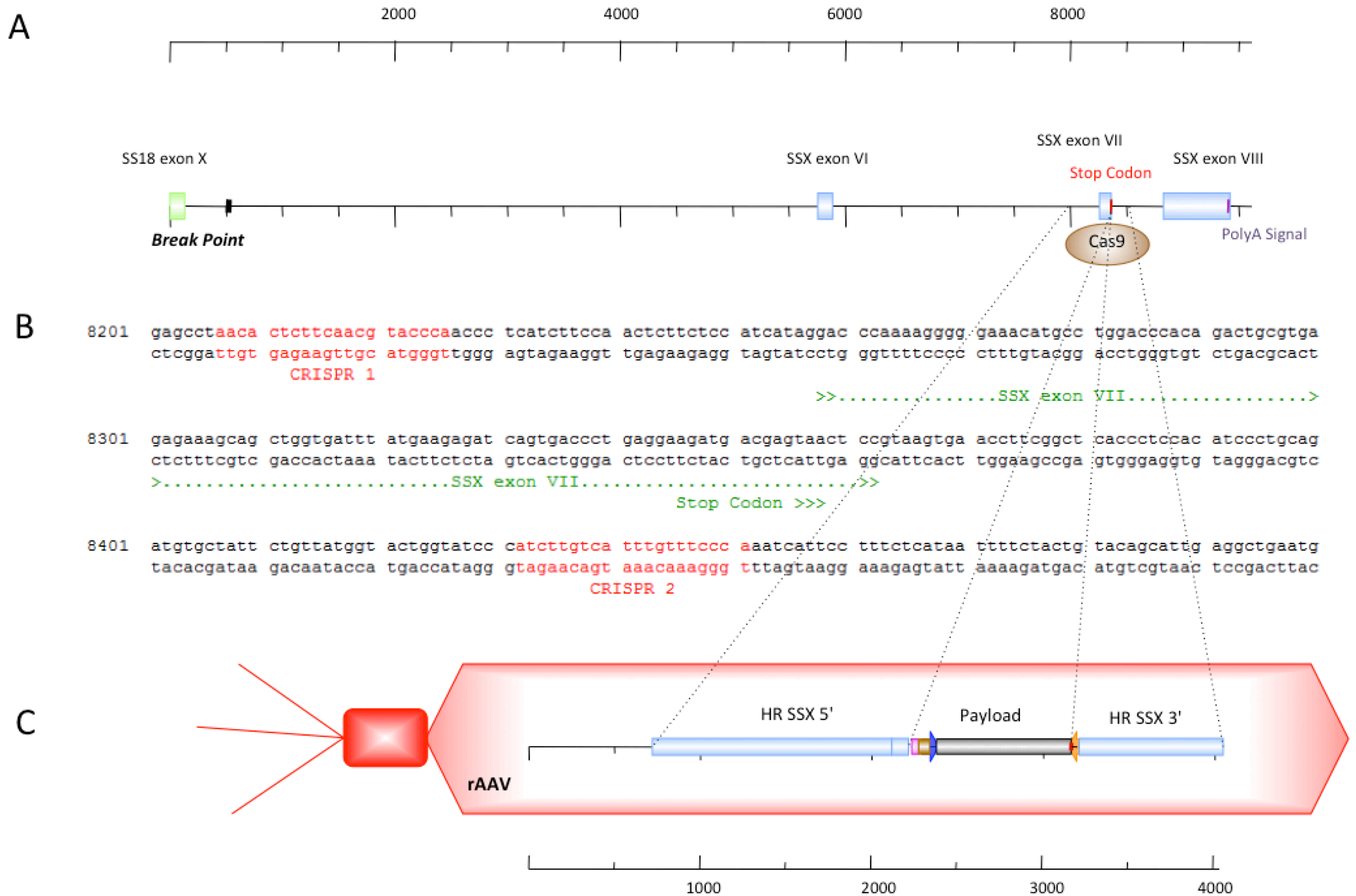

Fig. 1 – Design of gene targeting. (A) *SS18-SSX1* translocation in MGH5 primary cell line. The canonical Break Point is between *SS18* (exon X) and *SSX1* (exon VI). The STOP codon is in the exon VII of *SSX1*. The donor DNA will replace *SS18-SSX1* stop codon. **(B)** CRISPR/Cas9 is used to double strand break the DNA at the targeted site and enhance HDR. CRISPR1 and CRISPR2 will serve as a guide to Cas9 and were generated by MIT CRISPR Designer and CHOPCHOP software. **(C)** Donor DNA is transduced into MGH5 by a rAAV type 2 vector. The Payload is flanked by two SSX1 homologous regions (5'HR and 3'HR) for homologous recombination.

3. Results

3.1 Design of the donor DNA template

We designed the donor [Fig. 2A] to have three main segments. At the two ends, we placed the homologous regions (5'HR and 3'HR) flanking in the middle the payload, which contains all the relevant modifications: the V5 tag, the 2A region (23), the first LoxP sequence (LoxP66) with a 5' to 3' orientation, the neomycin resistance with the stop codon and the second LoxP sequence (LoxP71) with a 3' to 5' orientation (iLoxP71). Altogether the payload has a length of approximately 800 bps while the homologous regions are between 700 and 1100 bps[Fig. 2A].

When the homologous sequences are exchanged in a HR event with the corresponding sequences in the chromosome, the payload in MGH5 genome will be correctly positioned [Fig. 1A, C]. Because these events are rare, a positive selection becomes necessary. This goal is achieved by introducing in the donor cassette

the neomycin drug resistance gene (NeoR). By doing so, all the cells that do not go through the HR event will be sensitive to the drug treatment and will die, while the cells that translocate will survive.
Normally, this is accomplished by introducing an entire transcriptional unit, including the promoter coding for the NeoR protein, in the payload. However, this approach leads to a high number of false-positives originated by random genomic insertions (16). While the rAAV vector has been modified to minimize these events (20), we also designed a promoterless neomycine delivery system, which would further increase the frequency of on-target events (23). Due to the absence of the promoter, any random insertion of this cassette would not have a transcriptionally active NeoR gene. Only the accurate positioning by HR in the intended site will allow the SYT-SSX1 promoter to express NeoR gene.
The promotorless NeoR delivery system requires [Fig. 2A]:

1. The positioning of the V5 tag in frame with the last codon of SYT-SSX;
2. The introduction of a phage-derived 2A sequence (23) followed by the neomycine resistance gene;
3. The displacement of the STOP codon at the end of the cassette.

The 2A sequence leads to the transcription of a bi-cystronic mRNA capable of expressing two separate proteins (SYT-SXX1B-V5 and NeoR) during a unique translational event [Fig. 2C].

As described above, the HR of the designed cassette would lead to the tagging of SYT-SSX1 and at the same time the expression of the positive selection marker. However, we also want to achieve an inducible silencing of SYT-SSX1. This requires further modifications of the donor.
In recent years it has become common to obtain mRNA depletion using CRISPR-Cas9 construct which, by inducing targeted DNA damage, activates the *Non Sense Mediated mRNA Decay* (NMD)(24).
The NMD is a cascade process of mRNA degradation triggered by an early stop codon when it is placed outside of a tolerance window, between the 55 nucleotides preceding the last splicing site and the end of the last exon [Fig. 3A] (24). After the DNA damage, the Non Homologous End-Join Repair (NHEJ) leads to either insertions or deletions events (*indel*) around the break point site, and consequently to a frame shift. The end result of this process is an early termination signal outside the tolerance window and consequently the triggering of mRNA degradation.
We wanted to induce the activation of the NMD in a way that is compatible with the previously described design of our construct. We can achieve this objective by simply shifting the stop codon upstream, outside the NMD tolerance window, after inducing the rotation of the NeoR marker. The latter can be mediated by the Cre recombinase [Fig. 3C].
The Cre enzyme recognizes two sequences belonging to a family of sequence variances, collectively denominated LoxP, and delete or invert the DNA placed in between them (25). Deletion or inversion is obtained by the relative orientation of the two LoxP sites (25)[Fig. 3B]. In particular, when they face each other symmetrically the Cre enzyme induces an inversion [Fig. 3B].

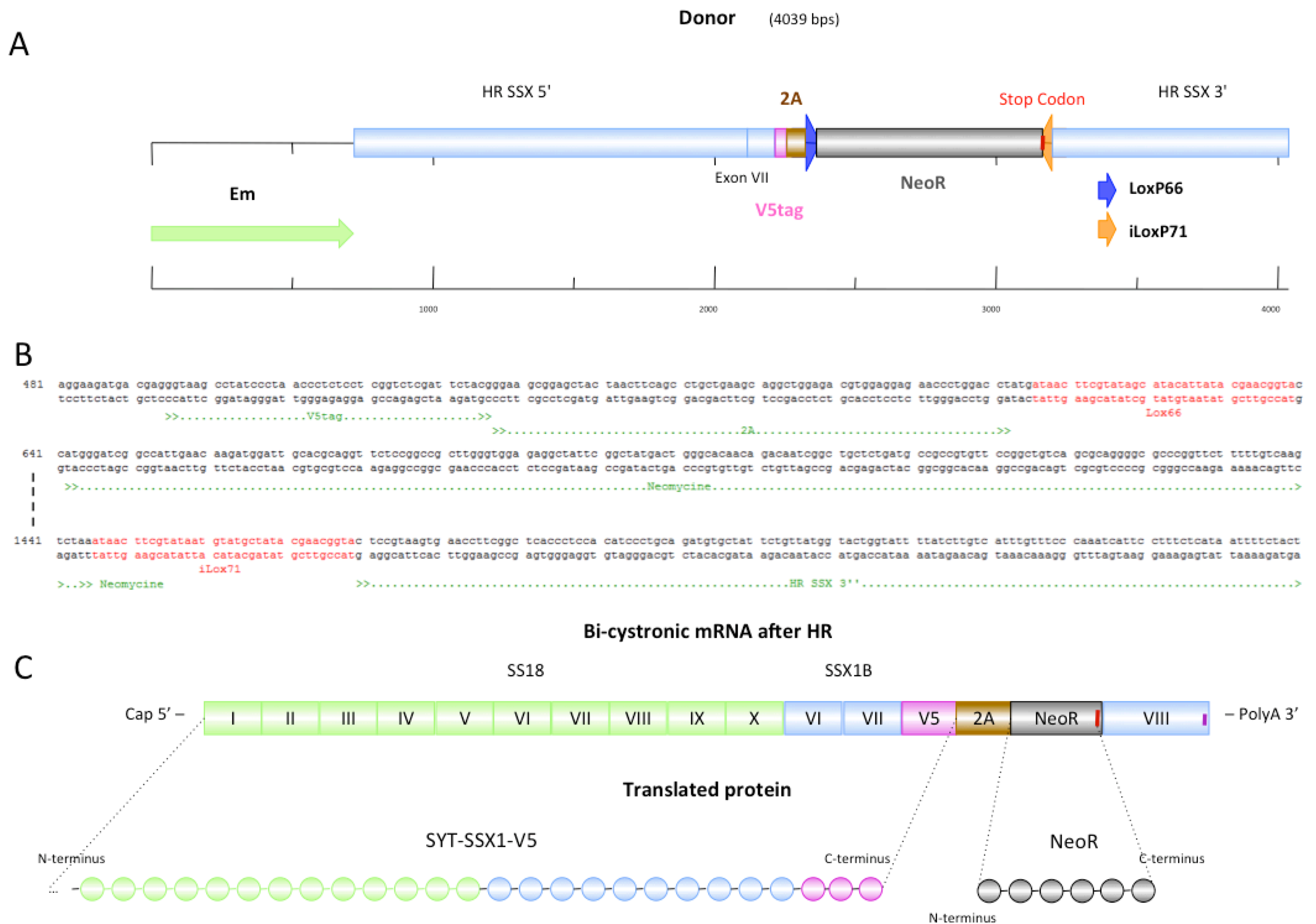

Fig. 2 – Design of the Donor template. (A) From 5' to 3', Donor consists in: opening reading frame GFP *Emerald* (Em) to detect false-positive insertions, *SSX Homologous region 5' (HR SSX 5')* including *SSX exon VII, V5 tag* that will be placed in frame after SSX1 exon VII, *2A* for a promotorless expression of neomycin resistance cassette, *Neomycin resistance* (NeoR) for positive selection, strategically placed *LoxP* sites (LoxP66 and iLoxP71) for Cre-mediated silencing of SYT-SSX1-V5 and *SSX Homologous region 3'* (HR SSX 3'). **(B)** Sequences of the two strategically placed LoxP for Cre-mediated SYT-SSX-V5 silencing. The two LoxP are placed in anti-parallel orientation to allow flip over of NeoR sequence. **(C)** Upper panel, resulting bi-cystronic mRNA after gene targeting. Lower panel, resulting polipeptides synthetized during translation. 2A is responsible for the the two separate peptides [**(C)** is not in scale].

Therefore we added to the donor sequence two LoxP sites placed strategically to either sides of neomycin resistance cassette and oriented in the opposite direction [Fig. 2A].

To prevent a chain of continuous inversion events, the two LoxP sites are not identical (LoxP66 and iLoxP71)[Fig. 2B]. This configuration leads to the destruction of both LoxP sites after the first inversion event, therefore preventing further recombinations (25).

In synthesis, this construct allows to flip over the neomycin resistance DNA fragment in presence of Cre recombinase, consequently introducing an early STOP codon which will trigger the NMD(24), resulting in the silencing of SYT-SSX1 [Fig. 3C].

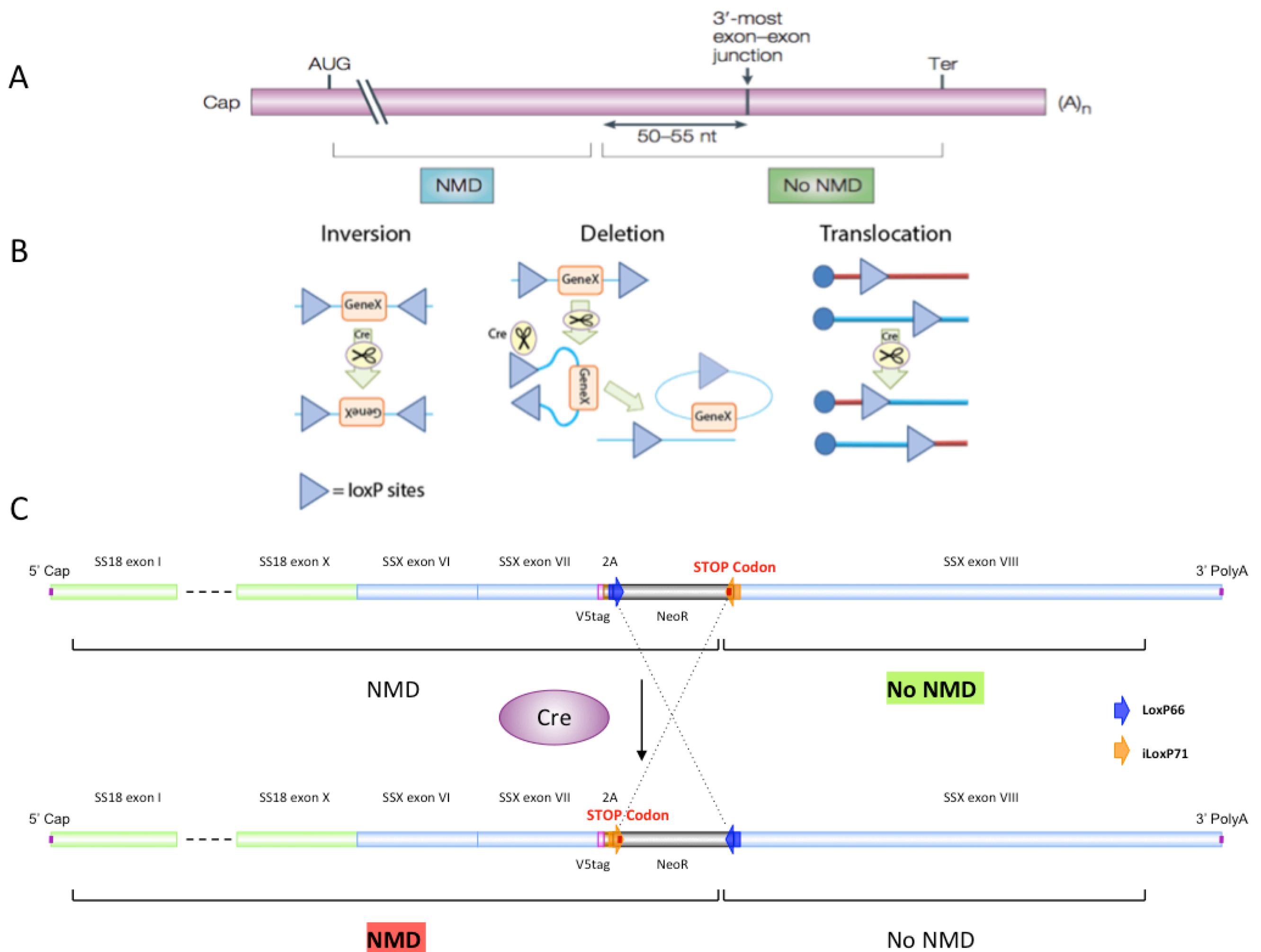

Fig. 3 – Cre-mediated silencing of SYT-SSX. (A) Mechanism of Non-sense Mediated mRNA Decay (NMD). When the STOP codon is more than 50-55 nt far from 3'-most exon-exon junction, the NMD is triggered and the mRNA is degraded. Image ref: Maquat et al. 2004 **(B)** Cre-Lox recombination system allows inversion, deletion or translocation following the relative orientation of the LoxP sites on the genome. Image ref: media.addgene.org **(C)** In the upper panel, the absence of Cre treatment, the *STOP codon* does not trigger the NMD and SYT-SSX-V5 will be translated by MGH5 cells. In the lower panel, Cre treatment allows to flip over NeoR fragment and it displaces the STOP codon outside the NMD tolerated window, hence SYT-SSX-V5 mRNA will be destroyed.

3.2 Amplification and assembly of the donor DNA into the rAAV vector

To create the donor, we divided it in 4 segments.

1. A promoterless open reading frame coding for the Emerald (Em) Green Fluorescent Protein (GFP) (26), which will work as a sensor for activation of the transcription due to random insertion in frame into a transcriptional unit of the genome.
2. 5'HR
3. A segment of DNA containing all the different elements of the construct (V5, 2A, LoxP66, NeoR cassette and iLoxP71)
4. 3'HR

The Em GFP, was amplified by a vector containing the Emerald gene (see materials and methods p.20). The two HR were amplified from genomic DNA extracted from MGH5 cells. Finally, the payload was chemically synthetized (gBlock, IDT) containing all the elements previously described [Fig. 4B].

The next step was to reassemble these four segments and clone them into the rAAV vector [Fig. 4A]. We initially tried the Gibson's assembly method, which relies on an overlapping sequence of at least 30 bps in common with each segment. Unfortunately, the assembly did not work probably because the presence of internal repetitions inhibits the process by creating secondary structures. Therefore we had to change our strategy to the traditional cloning method based on restriction enzymes cut and paste. This required to design the segments so that each would have individual, uniquely compatible, restriction sites [Tab. 1] (see materials and methods p.20).

To guarantee an ideal enzymatic restriction, each individual segment was first cloned in the pCR-Blunt vector (see materials and methods p.20), which is a plasmid designed specifically for blunt-end cloning with a low frequency of false-positives self-ligation events (27). After the cloning, each segment was checked by Sanger sequencing (28). Next, we digested with the corresponding enzyme each plasmid and gel-purified the different segments, including the rAAV backbone. Finally to join them together, we set up a ligation of all five elements, i.e. the four inserts and the backbone. After screening 12 clones by digestion analysis with *Bgl*II and *Acc*I we identified two positives clones one of which was checked by DNA sequencing.

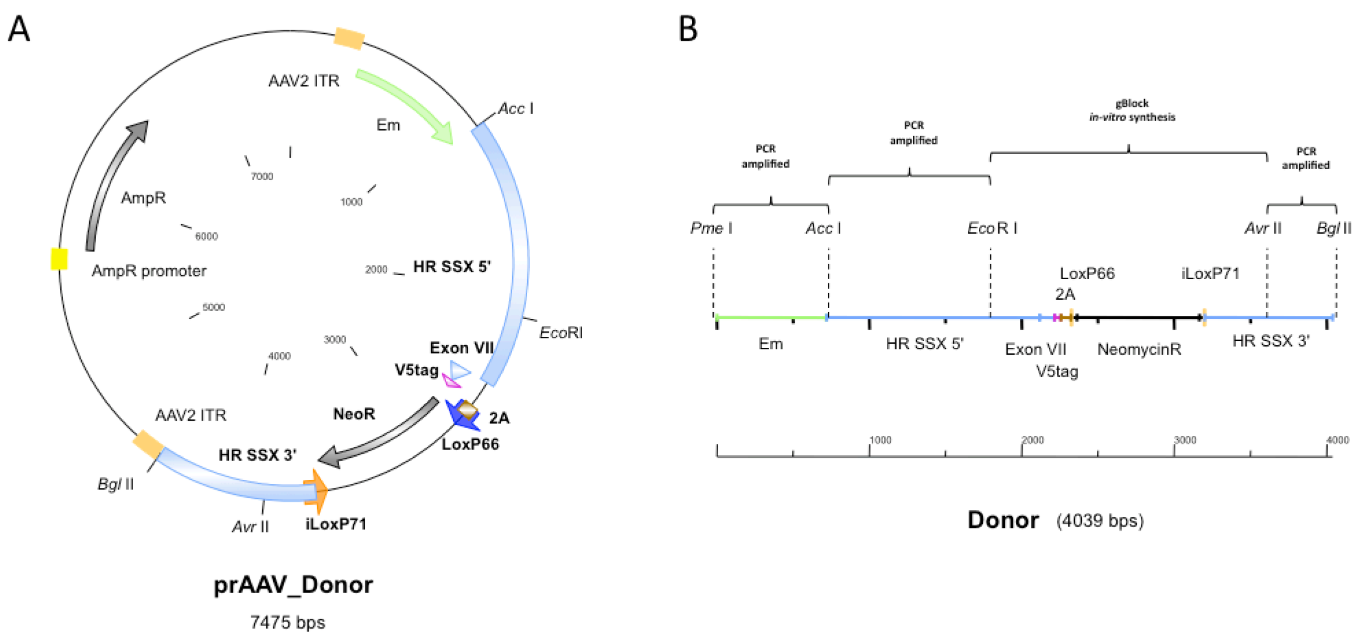

Fig. 4 – Assembly of the donor plasmid for rAAV production. (A) *prAAV–Amp*^{r} carries the donor DNA between its *Internal Terminal Repeats* (ITR). Donor DNA from 5' to 3': opening reading frame GFP *Emerald* (Em) to detect false-positive insertions, *SSX Homologous region 5' (HR SSX 5')* including *SSX exon VII, V5 tag* that will be placed in frame after SSX1 exon VII, *2A* for a promotorless expression of neomycin resistance cassette, *Neomycin resistance* (NeoR) for positive selection, *LoxP* sites (LoxP66 and iLoxP71) configured for Cre-mediated silencing of SYT-SSX1-V5 and *SSX Homologous region 3'* (HR SSX 3'). *AccI, EcoRI, AvrII and BglII* are the restriction sites used for the ligase-mediated assembling. **(B)** The Em, was amplified from a vector containing the Emerald gene (Invitrogen). The two HR were amplified from genomic DNA extracted from MGH5 cells. Finally, the payload was *in-vitro* synthetized as a gBlock (IDT) containing HR SSX 5', V5tag, 2A, LoxP66, NomyR, iLoxP71 and HR SSX 3'.

		Restriction enzymes (RE)	RE recrutement sites
Backbone *prAAV–Amp^r^*	5'	*Bgl* II	A/GATCT TCTAG/A
	3'	*Hind*III-blunt	GTTT/AAAC CAAA/TTTG
Emerald	5'	*Pme* I	
	3'	*Acc* I	GT/MKAC CAKM/TG
HR SSX 5'	5'		
	3'	*Eco*R I	G/AATTC CTTAA/G
G-Block	5'		
	3'	*Avr* II	C/CTAGG GGATC/C
HR SSX 3'	5'		
	3'	*Bgl* II	A/GATCT TCTAG/A

Tab. 1 – Donor restriction sites. We divided the entire clone in four segments, each one flanked by unique restriction sites.

3.3 CRISPR/Cas9 design and cloning in lentivirus plasmid

A targeted nuclease-directed double strand brake in the genomic region where the recombination is intended to happen increases the probability of HDR versus NHEJ when donor is delivered by rAAV (20). Cas9 is a yeast (*Saccharomyces cerevisiae*) derived nuclease which allows to cut an eukaryotic DNA in a programmable and sequence-specific manner when guided by a short-guide RNA strand (sgRNA) (29). The sgRNA will bind to its homologous DNA region over the genome, creating an heteroduplex which is recognized by the Cas9 enzyme. Therefore Cas9 is tethered over that segment where it will catalyse a double strand cut hence activating the HDR.

To deliver the CRISPR system to the MGH5 cells we used the pLentiCRISPRV2 Lentiviral vector [Fig. 4A] (30), which in its payload contains two cassettes. The first expressing the sgRNA under Pol III U6 promoter and the second a polycistronic mRNA coding for Cas9 and a Puromycine resistance (PuroR) under Pol II EF1α promoter [Fig. 5A].

This double cassette plasmid allows the delivery of both elements of the CRISPR-Cas9 system at the same time and requires only the cloning of a short oligoDNA, coding for the sgRNA (30).

To maximize the chances of a successful double strand cut we decided to design two different oligos targeting the same ROI [Fig. 4B]. We used MIT CRISPR design software and CHOPCHOP software (see materials and methods p.19) to define all the candidate sequences, of which we have chosen the two with the highest score in both programs [Fig. 5B].

We synthetize and cloned these two sequences into the pLentiCRISPERV2 (see materials and methods p.19) plasmid obtaining two CRISPR systems targeting the same region [Fig. 1A,B].

Both the Lentivirus(CRISPR) and the rAAV(Donor) are produced in similar manner. They require the co-transfection of the payload-plasmid, packaging and helper constructs into a host cell line, HEK293T, which will then produce the vectors.

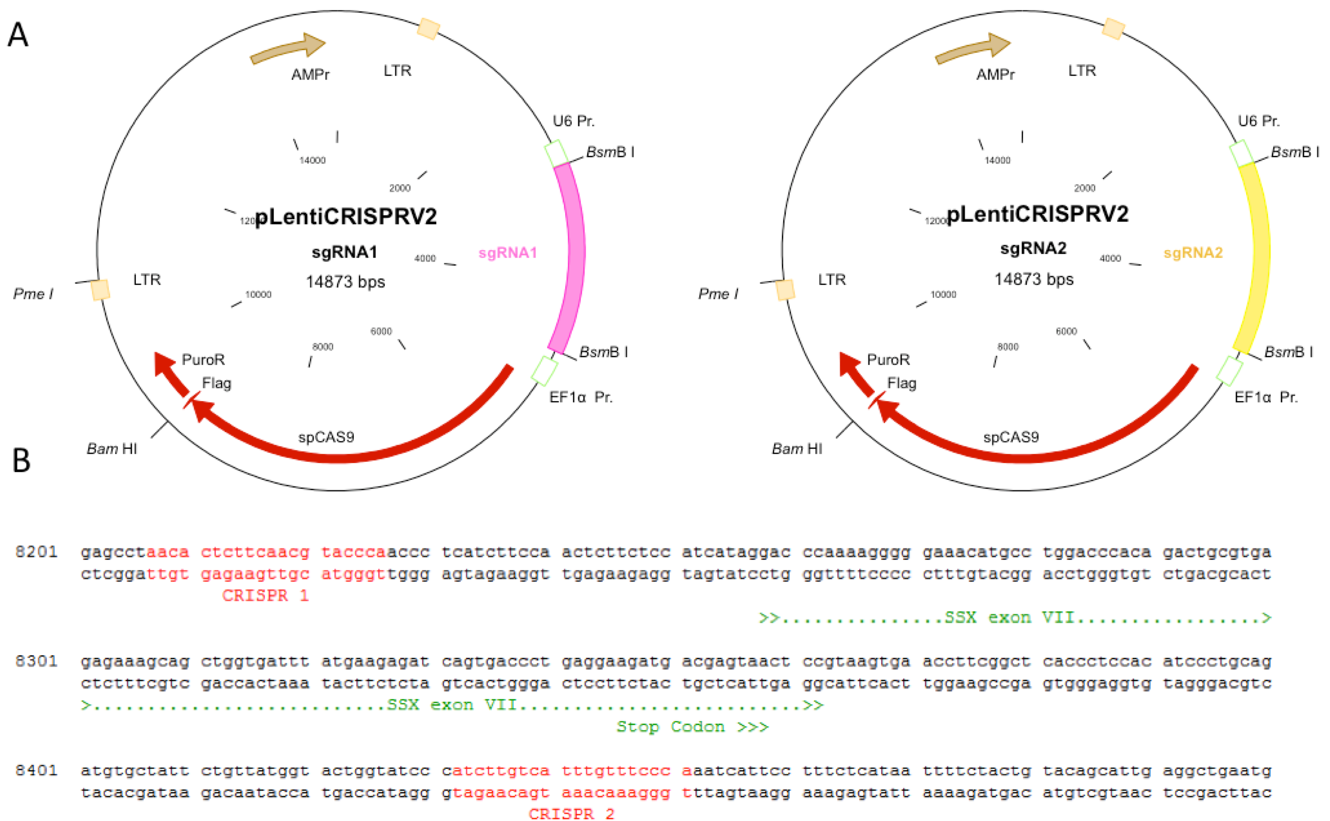

Fig. 5 – CRISPR/Cas9 design – Lentiviral vector. (A) pLentiCRISPRV2, 1 and 2 : sgRNA1 and 2 are placed under U6 promoter and Cas9 is placed under EF1α promoter **(B)** Sequences of two sgRNA have been chosen on the basis of two different softwares (MIT CRISPR Designer and CHOPCHOP). Cas9 recruited on *SS18-SSX* locus will double strand cut the DNA near the STOP codon. [Plasmids not in scale]

3.4 Lentivirus production

We produced the lentiviral vectors into HEK293T cells using a packaging mutated gene, which leads to an integration deficient lentivirus (see materials and methods p. 19). The CRISPR plasmids together with the other four support plasmids were delivered to HEK293T using the liposome based agent. After 72 hours, we purified them from the supernatant and titrated the viral particles using a qPCR-based measurement of the viral genome copy number [Fig. 6C]. This was done so that we could target the MGH5 cells at three different Multiplicity Of Infections (MOI) using the highest non-toxic dose. The expression of Cas9 was as well tested on MGH5 cells transduced with CRISPR1 and CRISPR2 after 48 hours from the initial infection by western-blot analysis using the anti-Flag anti-body [Fig. 6D].

3.5 rAAV production

Once the donor was cloned in the rAAV plasmid, we proceeded to produce the virus. Briefly, similar to the lentiviral system, the donor plasmid was cotransfected with the packaging and envelop plasmids into HEK293T cells using the liposome based Fugene kit. In the cells, the viral particles are a blend of single strand sequences randomly generated from the original double-stranded donor (31). 48 hours post-transfection, the cells were harvested and after cell lysis the virus was purified then concentrated. Finally, the number of viral genome copies was assessed by qPCR [Fig. 6A] and an optimal MOI between 20 and 10 was chosen after viability test [Fig. 6B].

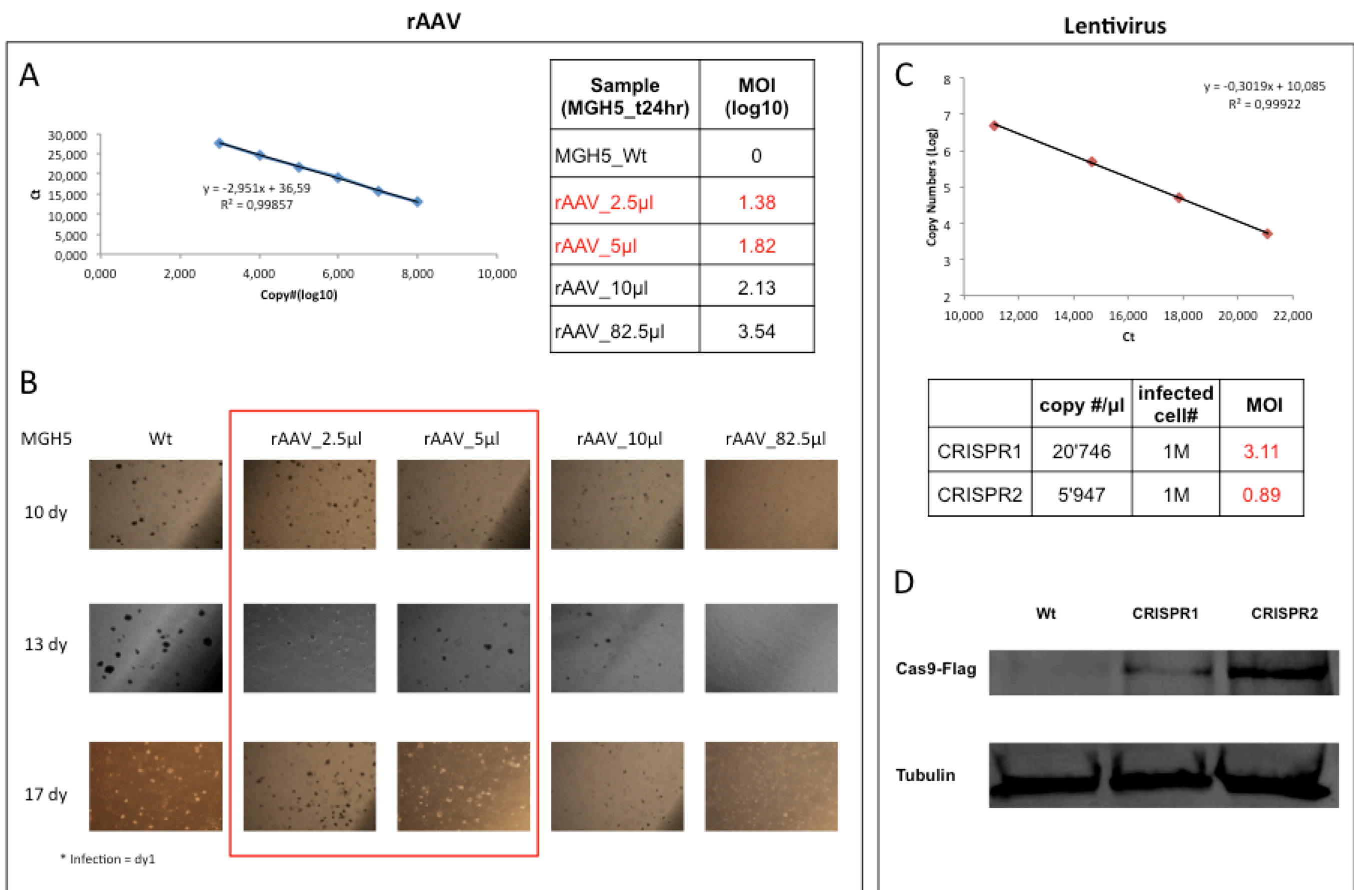

Fig. 6 – Viruses production. (A) rAAV titration by qPCR. MGH5 cells were infected with the amount indicated in the right panel. Left panel, titration curve. The resulting MOI is reported in the figure. In red, the corresponding values reported in B panel **(B)** Viability test of cells infected with the corresponding amount of rAAV. In the red box, the viable cells. **(C)** Lentivirus titration by qPCR. Top, titration curve. Bottom, reported MOI in MGH5 cells. **(D)** Western-blot of MGH5 cells transduced with CRISPR1 and CRISPR2, MOI indicated in panel C. Cells were processed 48 hr after the infection. Tubulin shows the equal loading. Negative control were the cells infected with an empty Lentivirus.

4. Discussion

In this thesis we have described the preliminary work to add a sequence tag to SYT-SSX1 genomic locus, with the intent of being able to use a high-affinity and low cross-reactivity antibody (i.e. anti-V5) capable of recognizing exclusively the fusion protein. Also, we wanted to further modify the genome of MGH5 primary SS cell line to subsequently induce the depletion of SYT-SSX1.

Genome editing of this kind requires a double homologous recombination event. Thus, the main challenge we faced was to increase the HR frequency reported so far in the literature of several orders of magnitude (15,16,32). This would give us a chance to modify a bulk population of primary SS cells, as opposed to a clonal selection. This is necessary because, in a primary cell culture, single cell cloning would require a long selection process, ranging from several weeks to few months, likely to be incompatible with the biology of primary cells. As a strategy to burst HR events, we have designed a CRISPR-Cas9 system to damage DNA near the stop codon of SYT-SSX1 locus, therefore activating the HDR pathway close to the region of interest. We have also decided to use rAAV vector, which is known to be very efficient as a delivery system

of the template in HR settings, even when provided with short homologous regions (20,31,33). Finally, the inhibition of the NHEJ has been also shown to be able to increase up to 19-fold the success rate of HR events(21).

At the present time, we have prepared the tools that will allow us to tag in primary SS spheres SYT-SSX1 fusion protein. We have designed two different sgRNA [Fig. 5B], which, as stated before, in a CRISPR system will stimulate the HDR pathway and activate the HR cell activity in the target genomic sequence. We cloned these sgRNA into a Cas9 producing Lentiviral vector, thus constituting a complete single plasmid CRISPR-Cas9 system [Fig. 5A].

Next we have designed a template with the following characteristics [Fig. 2A]:

1. A promoterless open reading frame GFP which works as a sensor for random genomic insertions in transcriptionally actives genomic regions, therefore detecting false positives events;
2. The two HR sequences flanking our payload;
3. A payload containing the V5 tag and the NeoR positive selection marker;
4. Two strategically placed LoxP sites necessary to induce the silencing of SYT-SSX1 upon Cre expression.

Finally we have merged these elements in a single rAAV vector, which as stated before should increase successful HR events (20,31,33) [Fig. 4A]. The viruses have also been produced and titrated in a preliminary experiment [Fig. 6].

The following steps remain to be done:

1. A dose-response experiment to define the highest non-toxic-titer of viruses tolerated by MGH5;
2. Assess the best dose of neomycine antibiotic and the best conditions for an efficient selection of MGH5 cells.
3. Co-transduce the three viruses with the highest possible tolerated MOI, ideally 5 to 10, which will allow the 99,3% to ~100% of cells to be successfully co-transduced;
4. Finally, inhibit the NHEJ by treating MGH5 with Scr7 drug.

Following this co-transduction, we need to place the cells under a neomycine selection for 5 days and wait 3 to 4 weeks for surviving spheres to emerge. A failure to observe growing neomycine selected spheres will require a troubleshooting step including an assessment of CRISPR activity and eventually, if we will observe a low frequency damage of DNA, designing and cloning of new sgRNAs.

Spheres of primary cell lines are difficult to culture, they might spontaneously stop growing or fail to maintain a sphere-like morphology over this long window of selection. In this case, we will need to change model and use a SS established cell line in alternative to the more pertinent primary cell culture.

After the positive selection, the neomycine-resistant cells will be expanded and checked for the presence of the V5 tag, as well as for the depletion of SYT-SSX-V5 upon Cre-recombinase treatment.

To check for the presence of the V5 tag, we will perform a western-blot of the recombinant MGH5, using the anti-V5 anti-body to detect a protein of molecular weight compatible with the size of SYT-SSX1-V5. After this test, the presence of false-positives cells will be assessed by qPCR on the genome of neomycine resistant MGH5 with two pairs of primers: one detecting the V5 tag and the other specifically SYT-SSX. The ratio between the two measurements should approach 1 if there are no false-positives in the cell population. Finally, the genomic DNA of the cells will be analysed by PCR-cloning and Sanger sequence.
To detect the efficiency of SYT-SSX-V5 depletion after Cre-recombinase treatment we will do a reverse transcription/qPCR of the MGH5 neomycine resistant cells after Cre transduction, using a pair of primers specific for SYT-SSX mRNA. An efficient depletion should show a major drop in the amount of the amplicon compared to the same recombinant MGH5 transduced with a control virus.

The established synovial sarcoma cell lines, while are a useful model, have undergone to thousands of generations of cultures in an artificial environment, therefore likely diverging from the original tumour. From this point of view, the tagging of the fusion protein SYT-SSX1 in cells freshly isolated from synovial sarcoma tissue will allow the study of the aberrant functions of this chimera in a model more relevant for the reality of the disease.
In this model, the promoter regulation and the level of expression of both the fusion protein and the wild type SS18 will not be modified, thus all the bias potentially coming from over-expression would be eliminated.
In addition, the inducible knockout of SYT-SSX-V5 can provide means for a direct comparison of chromatin state in the presence and absence of the fusion protein.
The presence of the tag will allow the use of an high-quality commercially available, ChIP-grade anti-V5 anti-body.

This opens at least three different lines of research.
First, ChIP-seq data of SYT-SSX1-V5 can be compared to those of wild type SYT protein in the presence or absence of SYT-SSX1 and uncover how the chimera affects the function of the corresponding wild type protein.
Second, recent works on synovial sarcoma (34) have shown that *SS18-SSX* affects the SWI/SNF chromatin remodelling complex by competing with SS18. Its expression results in an altered complex lacking the tumour suppressor BAF47. Other works have shown that the fusion protein interact with histone deacetylases (HDAC) (10,35), which could therefore affect also the functions of complexes such as the NuRD/Mi2/NRD and the Sin3/HDAC. In this context, the genomically tagged SS cells could be used to improve our understanding of this functional network in relationship with SYT-SSX.

Finally, the edited SS spheres will allow the identification of genome-wide DNA binding sites of complexes that include SYT-SSX1. Together with the ChIP-seq histone marks analysis of the chromatin state in these regions, this will give information concerning the transcriptional state of promoters and enhancers associated with the presence SYT-SSX1. Statistical analysis and information mining of these data, could therefore identify new promising targets for drug design.

5. Material & methods

5.1 MGH5 cell-culture

MGH5 primary SS cells were obtained by first mincing the SS specimen, incubating for 2 hours at 37° with a cocktail of collagenases and DNAses and passing the cells in a 70 µm cell-strainer in order to obtain a single cells suspension. After red blood cells lysis, cells were cultured in IMDM medium (Gibco), supplemented with 20% KO serum (Gibco), 10 ng/ml human recombinant EGF and bFGF (Invitrogene) and 1% Pen/Strep (Gibco) in ultra-low attachment flasks (Corning) at 37 ^{o}C and 5% CO_2 (36).

5.2 Viral production

The viruses were produced by transfecting HEK293T cells with pLentiCRISPRV2 (30) (a gift from Feng Zhang Addgene #52961), pCMVdeltaR8.2 (Addgene #12263) and pMD2.G (Addgene #12259) were a gift from Didier Trono for Lentivirus and prAAV2, pRC2-mi342 and pHelper (Takara Clonetech) for rAAV, using Fugene liposomes (Sigma), with the ratio of 3:1 for a total of 8 µg of DNA per a million of cells. For the Lentiviruses, the target vector is the pLentiCRISPRV2 (30). The ratio between the plasmids was 1:1:1 for rAAV and 1:1.5:3 (pLentiCRISPRV2 : pCMVdeltaR8.2 : pMD2.G) for Lentivirus.

The Lentivirus was purified by collecting 30 ml of supernatant from $3 \cdot 10^{6}$ of HEK293T cells transfected after 72 hours of incubation, with a treatment of 4 mmol caffeine after 17 hours post-transfection (37). The supernatant was sinned in an ultracentrifuge at 60000 g for 1,5 hours. The pellet was resuspended in 500 µl of medium, aliquoted and frozen at -80 ^{o}C.

The rAAV system was purchased from Takara Clonetech (catalogue 6230), and the virus was purified accorded to the manufactures protocol from $1 \cdot 10^{6}$ HEK293T cells.

5.3 Design and cloning of CRISPR

Two sgRNA (see Fig. 4 for sequences) were designed using MIT CRISPR Design (link: http://crispr.mit.edu:8079) and CHOPCHOP software (link: https://chopchop.rc.fas.harvard.edu). Each sgRNA was synthetized as a double strand oligo and annealed and cloned into the pCRISPRV2 plasmid, following the methods described in Zangh et al. (30).

5.4 Design and cloning of Donor

The GFP Em (26) opening reading frame was amplified from cells transduced with BacMam GFP transduction (Thermofisher catalogue # B10383) by PCR using the following primers:

- Fw 5'-CGCGGATCCGCGAGCTTTGTTTAAACGGCGCGCCGGATGGTGAGCAAAGGTGAGGAACT- 3' ;
- Rv 5'-GAGAAGCTGCAGCAAGTCTACTTGTACAGCTCGTCCATGC-3'.

The 5’HR and 3’HR were also amplified from 50 ng genomic MGH5 DNA extracted with DNA extraction kit (Quiagen), using respectively 5’HR primers:

- Fw 5'-ACGAGCTGTACAAGTAGACTTGCTGCAGCTTCTCCATCA-3'
- Rv 5'-CATGGGCCTTCTTTATCCAG-3'

and 3’HR primers:

- Fw 5'-ATCTTGTCATTTGTTTCCCAAATC-3'
- Rv 5' -CCCATCGATGGGGGAAGATCTTCCCACTCAGAACTGCCCTCAGTAG-3'

The remaining part was *in-vitro* synthetized and lyophilized as a double-stranded DNA (IDT) and was subsequently resuspended at the concentration of 1 ug/ul.

Each element was flanked by a pair of unique restriction sites as indicated in *Table 1* (p.14) designed for the final reassembly.

The amplicons as well as the synthetic DNA were individually cloned blunt-ended into the pCR-Blunt plasmid (Invitrogene catalog #K2700-20) and sequence-verified (Sanger reaction) using M13 forward and reverse primers, which recognize their corresponding sequence on the pCR-Blunt vector. The DNA (2 µg) extracted from each individual clone was digested with the corresponding restriction enzyme and gel-purified. Together with 2 µg of backbone (BB) prAAV digested *Hind*III and blunted using T4 polimerase (New England Biolabs) and subsequently digested with *Bgl*II, as well gel-purified. A ligation of 10 ng of each element was performed for 5 hours at 15 ^{o}C using 1U of T4 ligase (New England Biolabs). The ligation was precipitated and resuspendend in 2 µl of water and electroporated in STBL3 electrocompetent bacterial cells (Termofisher). After selection 10 randomly chosen clones were analyzed by digestion with *Acc*I and *Bgl*II of the three resulting positive candidates, one was analyzed and confirmed by Sanger sequence.

5.5 Measurement of viral copy number

An aliquot of 10 µl of purified Lentivirus were used to titrate the preparation by q-RT-PCR using the Takare Clonetech (catalog #631235) according to the manufactures specifications.

5.6 Western blot

10^{6} of MGH5 cells were lysated with 200 µl of RIPA buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 1% Triton X-100, 0.1% SDS and Protease inhibitor cocktail) (Roche). The sample was sonicated three cycles of 10 seconds at 50% power in ice. 100 µg of total proteins was loaded on a 6% SDS-PAGE gel and run at 120 V

for 2 hours. The gel was transferred to a nitrocellulose membrane at 250 mA for 2 hours. The membrane was blocked with 5% milk for 30 minutes and probed with anti-Flag anti-body (Cell Signaling catalog #2368) at 4 ^{o}C overnight.

Acknowledgements

I would like to thank Professor Stamenkovic for his great availability and kindness as well as for introducing me to the exciting world of scientific research.

A special thank goes to Dr. Carlo Fusco for his infinite patience in teaching me all the basis of laboratory work and for accompanying me trough out all the steps of my master thesis production.

A final thank goes to all the scientists working in Prof. Stamenkovic's lab for allowing me in the laboratory and make me feel part of it. In particular, many thanks to Dr. Luisa Cironi and Tugba Keskin, specifically working on the biology of synovial sarcoma, for their great disposability in answering all my questions.

Bibliography

1. Herzog CE. Overview of sarcomas in the adolescent and young adult population. J Pediatr Hematol Oncol. 2005 Apr;27(4):215–8.
2. Nielsen TO, Poulin NM, Ladanyi M. Synovial sarcoma: recent discoveries as a roadmap to new avenues for therapy. Cancer Discov. 2015 Feb;5(2):124–34.
3. Riggi N, Cironi L, Suvà M-L, Stamenkovic I. Sarcomas: genetics, signalling, and cellular origins. Part 1: The fellowship of TET. J Pathol. 2007 Sep;213(1):4–20.
4. Clark J, Rocques PJ, Crew AJ, Gill S, Shipley J, Chan AM-L, et al. Identification of novel genes, SYT and SSX, involved in the t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma. Nat Genet. 1994 Aug;7(4):502–8.
5. Iwasaki T, Koibuchi N, Chin WW. Synovial Sarcoma Translocation (SYT) Encodes a Nuclear Receptor Coactivator. Endocrinology. 2005 Sep 1;146(9):3892–9.
6. Crew AJ, Clark J, Fisher C, Gill S, Grimer R, Chand A, et al. Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with homology to the Kruppel-associated box in human synovial sarcoma. EMBO J. 1995 May 15;14(10):2333–40.
7. Ishida M, Tanaka S, Ohki M, Ohta T. Transcriptional co-activator activity of SYT is negatively regulated by BRM and Brg1. Genes Cells. 2004 May 1;9(5):419–28.
8. Kato H, Tjernberg A, Zhang W, Krutchinsky AN, An W, Takeuchi T, et al. SYT Associates with Human SNF/SWI Complexes and the C-terminal Region of Its Fusion Partner SSX1 Targets Histones. J Biol Chem. 2002 Feb 15;277(7):5498–505.
9. Nagai M, Tanaka S, Tsuda M, Endo S, Kato H, Sonobe H, et al. Analysis of transforming activity of human synovial sarcoma-associated chimeric protein SYT-SSX1 bound to chromatin remodeling factor hBRM/hSNF2α. Proc Natl Acad Sci. 2001 Mar 27;98(7):3843–8.
10. Cironi L, Petricevic T, Fernandes Vieira V, Provero P, Fusco C, Cornaz S, et al. The fusion protein SS18-SSX1 employs core Wnt pathway transcription factors to induce a partial Wnt signature in synovial sarcoma. Sci Rep. 2016 Feb 24;6:22113.
11. Johnson DS, Mortazavi A, Myers RM, Wold B. Genome-Wide Mapping of in Vivo Protein-DNA Interactions. Science. 2007 Jun 8;316(5830):1497–502.
12. Farnham PJ. Insights from genomic profiling of transcription factors. Nat Rev Genet. 2009 Sep 1;10(9):605.
13. Kolodziej KE, Pourfarzad F, de Boer E, Krpic S, Grosveld F, Strouboulis J. Optimal use of tandem biotin and V5 tags in ChIP assays. BMC Mol Biol. 2009 Feb 5;10:6.
14. Southern JA, Young DF, Heaney F, Baumgärtner WK, Randall RE. Identification of an epitope on the P and V proteins of simian virus 5 that distinguishes between two isolates with different biological characteristics. J Gen Virol. 1991 Jul;72 (Pt 7):1551–7.
15. Thomas KR, Capecchi MR. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell. 1987 Nov 6;51(3):503–12.
16. Waldman T, Kinzler KW, Vogelstein B. p21 Is Necessary for the p53-mediated G1 Arrest in Human Cancer Cells. Cancer Res. 1995 Nov 15;55(22):5187–90.
17. Chen X, Gonçalves MAFV. Engineered Viruses as Genome Editing Devices. Mol Ther J Am Soc Gene Ther. 2016 Mar;24(3):447–57.
18. Kohli M, Rago C, Lengauer C, Kinzler KW, Vogelstein B. Facile methods for generating human somatic cell gene knockouts using recombinant adeno-associated viruses. Nucleic Acids Res. 2004 Jan 1;32(1):e3–e3.
19. Russell DW, Hirata RK. Human gene targeting by viral vectors. Nat Genet. 1998 Apr 1;18(4):325.

20. Kaulich M, Dowdy SF. Combining CRISPR/Cas9 and rAAV Templates for Efficient Gene Editing. Nucleic Acid Ther. 2015 Nov 5;25(6):287–96.

21. Maruyama T, Dougan SK, Truttmann M, Bilate AM, Ingram JR, Ploegh HL. Inhibition of non-homologous end joining increases the efficiency of CRISPR/Cas9-mediated precise [TM: inserted] genome editing. Nat Biotechnol. 2015 May;33(5):538.

22. Wei Y, Sun M, Nilsson G, Dwight T, Xie Y, Wang J, et al. Characteristic sequence motifs located at the genomic breakpoints of the translocation t(X;18) in synovial sarcomas. Oncogene. 2003 Apr 10;22(14):2215.

23. Kim JH, Lee S-R, Li L-H, Park H-J, Park J-H, Lee KY, et al. High Cleavage Efficiency of a 2A Peptide Derived from Porcine Teschovirus-1 in Human Cell Lines, Zebrafish and Mice. PLOS ONE. 2011 Apr 29;6(4):e18556.

24. Maquat LE. Nonsense-Mediated mRNA Decay: Splicing, Translation And mRNP Dynamics. Nat Rev Mol Cell Biol. 2004 Feb;5(2):89–99.

25. Zhang Z, Lutz B. Cre recombinase-mediated inversion using lox66 and lox71: method to introduce conditional point mutations into the CREB-binding protein. Nucleic Acids Res. 2002 Sep 1;30(17):e90.

26. Fluorescent Protein Vectors [Internet]. [cited 2017 Sep 27]. Available from: https://www.thermofisher.com/uk/en/home/life-science/protein-biology/protein-expression/mammalian-protein-expression/fluorescent-protein-vectors.html

27. Zero Blunt PCR Cloning Kit - Thermo Fisher Scientific [Internet]. [cited 2017 Sep 27]. Available from: https://www.thermofisher.com/order/catalog/product/K270020

28. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol. 1975 May 25;94(3):441–8.

29. Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol. 2014 Dec 1;32(12):1262–7.

30. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods. 2014 Aug;11(8):783–4.

31. Hüser D, Gogol-Döring A, Chen W, Heilbronn R. Adeno-Associated Virus Type 2 Wild-Type and Vector-Mediated Genomic Integration Profiles of Human Diploid Fibroblasts Analyzed by Third-Generation PacBio DNA Sequencing. J Virol. 2014 Oct;88(19):11253.

32. Kan Y, Ruis B, Lin S, Hendrickson EA. The Mechanism of Gene Targeting in Human Somatic Cells. PLoS Genet [Internet]. 2014 Apr [cited 2017 Sep 27];10(4). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3974634/

33. Deyle DR, Russell DW. Adeno-associated virus vector integration. Curr Opin Mol Ther. 2009 Aug;11(4):442.

34. Kadoch C, Crabtree GR. Reversible Disruption of mSWI/SNF (BAF) Complexes by the SS18-SSX Oncogenic Fusion in Synovial Sarcoma. Cell. 2013 Mar 28;153(1):71.

35. Su L, Sampaio AV, Jones KB, Pacheco M, Goytain A, Lin S, et al. Deconstruction of the SS18-SSX Fusion Oncoprotein Complex: Insights into Disease Etiology and Therapeutics. Cancer Cell. 2012 Mar 20;21(3):333.

36. Suvà M-L, Riggi N, Stehle J-C, Baumer K, Tercier S, Joseph J-M, et al. Identification of Cancer Stem Cells in Ewing's Sarcoma. Cancer Res. 2009 Mar 1;69(5):1776–81.

37. Ellis BL, Potts PR, Porteus MH. Creating Higher Titer Lentivirus with Caffeine. Hum Gene Ther. 2011 Jan;22(1):93.