UNIL | Université de Lausanne

Université de Lausanne
Faculté de biologie et médecine
Ecole de médecine
Travail de Maîtrise en médecine

Creating an epigenetic toolbox to study pediatric sarcomas

par

Beatrice Rucci

sous la direction de

Prof. Ivan Stamenkovic

Année académique 2018-2019

Creating an epigenetic toolbox to study pediatric sarcomas

Abstract

Soft tissue and bone malignancies that affect children and young adults frequently have a poor prognosis and a limited set of therapeutic options. A better understanding of their molecular pathogenesis would give us a chance to improve the treatment, by making it more specific and less toxic than the generic chemo and radiotherapy. Recently, it has become clear that the chromatin status of cancer cells is essential for the development of these malignancies but the study of the epigenome remains technically difficult and frequently elusive. The purpose of this work is to help overcome this difficulty by creating a set of tools tailored to study chromatin modifications and their effect on tumor biology. In particular, the aim of our work is to adapt CRISPR-Cas9 technology, which has already revolutionized genome editing in the last decade, toward the study of epigenetic events.

Introduction

Pediatric sarcomas – there is still so much to discover

Sarcomas are a heterogeneous group of soft tissue and bone neoplasms. They are rare tumors (only 1% of all malignancies), but they affect children and young adults with higher frequency and their mortality remains excessively elevated (metastatic cases have a five-year survival rate of 20-30%) (1). Current treatments are based on aggressive surgery, chemotherapy, and radiations. The poor outcome of these cancers underlines the concrete necessity to find other treatments, targeting more specific mechanisms of the tumors. In this context, an interesting direction of research is pointing toward epigenetic modifications.

In fact, an important subset of these pediatric sarcomas, including Ewing sarcoma, synovial sarcoma, rhabdomyosarcoma and dermatofibrosarcoma, is characterized by unique chromosomal translocations, which lead to the expression of a fusion protein that in most cases functions as an aberrant transcription factor. Whereas these translocations have a wide range of effects on the transcriptome, there is an increasing body of evidence linking their activity to epigenetic modifications.
For example, in synovial sarcoma the inhibition of histone deacetylase (HDAC) can disrupt the oncoprotein complex, leading to apoptosis (2). It has also been demonstrated that synovial sarcoma translocation interacts with human BRM protein, that modulates chromatin remodeling (3). In another example, several links have been observed between the EWS-FLI-1 fusion protein and the epigenetic activity of Ewing sarcoma cells, including EWS-Fli1-mediated inhibition of p300-dependent acetylation of p53, which blocks apoptosis and inhibition of HDAC mediated deacetylation that can lead to decreased tumor growth (4,5).

In summary, better comprehension of these mechanisms could lead to novel therapeutic strategies in pediatric sarcomas that may significantly improve the outcome of patients, which is currently unsatisfactory.

CRISPR-Cas9 system – a new era for (epi)genetic manipulation

Genetic defects linked to disease and cancer have been studied for a very long time thanks to constant improvement of the available technology. In particular, the human genome project in the last 40 years has decoded the whole human genetic information and allowed the development of efficient research tools, including SNP arrays, low cost genome sequencing and high throughput screening techniques (6).

Much less understood are chemical modifications of genomic DNA that do not affect its sequence but have a major impact on its function by modifying its structure and thereby constituting a separate layer of activity. These modifications and the machinery that implements them are termed the epigenome. This epigenetic control is influenced by the environment, is hereditable, and is exploited by cancer cells; its importance therefore cannot be overstated. However, the epigenome has been an unexplored landscape largely for technical reasons, until the discovery of genome editing tools, including CRISPR and TALENs, allowed a technological leap forward.

Among the different genome editing tools, CRISPR-Cas9 technology has emerged as the most efficient and user-friendly and therefore we elected it to use it in our work (7,8).
Discovered in 1987 in Japan, RNA guided Cas9 nuclease is characterized by clustered regularly interspaced short palindromic repeats (CRISPR), and acts in the microbial adaptive immune system. Bacteria take advantage of these repeated sequences to recognize and destroy foreign DNA coming from phages and other sources (9).
CRISPR-Cas9 is composed of two essential elements: a Cas9 nuclease that cuts DNA, and a small guide RNA (gRNA) that directs the enzyme to a specific sequence of the genome (7) (Figure 1, left panel).

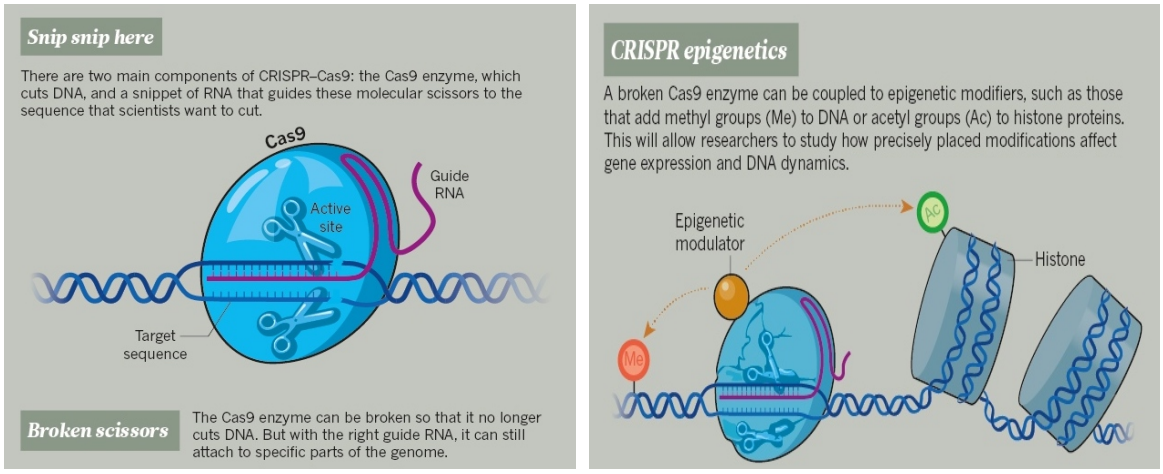


Figure 1, modified from Ledford, Nat News. 2016 (7) - CRISPR-Cas9: broken scissors and epigenetic modulator
A schematic view of the catalytically inactive Cas9 (dCas9) coupled with an epigenetic module that can activate or repress gene expression.

Despite the fact that this endonuclease was discovered in the eighties, the scientific community started to take advantage of its power to edit the genome only in the last few years.
The biggest advantage of CRISPR-Cas9 over all the other methods is the simplicity of the system. In fact, it allows targeted cleavage of DNA and can be used to induce any modification at the genomic level. Furthermore with a small modification the tool can be repurposed to accomplish the same effect at the epigenetic level (Figure 1, right panel).

Stanley Qi, from Stanford University in California, made the first modification necessary to achieve this goal: the abrogation of Cas9 nuclease activity with retention of the DNA/RNA binding properties. Subsequently, the next leap forward came in 2015, when a group from Duke University published a system to use dCas9 coupled with an epigenetic module.
In fact, the deactivated Cas9 (dCas9) can be fused with other enzymes that affect the transcriptional activity and/or the epigenetic status of DNA on a chosen sequence of the genome, with unprecedented precision and specificity. Thus, it is possible to modulate the activity of selected genes by modifying their epigenetic status (7,9,10).

Project goals – and a little hope

The aim of our project is to create a coherent and single-plasmid epigenetic toolbox using CRISPR-Cas9 technology, which will allow us to simplify and speed up the study of pediatric sarcoma pathogenesis and biology. The use of our toolbox could help the scientific community to progress in pediatric sarcoma understanding and we hope that this will push forward the development of more targeted treatments. The toolbox is composed of six of the more relevant epigenetic modules for pediatric sarcomas: VP64, P300, Tet1, DNMT3a, HDAC and KRAB (2,4,5,11–16) (Table 1).

Results

The universal vector

The first goal we wanted to accomplish was to develop a homogeneous system to deliver our epigenetic modules to the cells. To do that, we used a Lentivirus based method, which is one of the most efficient ways to transfer genetic material *in vivo*. In fact, the main advantages of VSV pseudotyped Lentiviruses are high efficiency of transduction, wide tropism, the capacity to infect cells independently of their mitotic state, and the ability to package more than one transcriptional unit in one single payload.

In particular, we wanted to base our common system on the Feng Zhang CRISPR-Cas9 second-generation vector, which is a single plasmid capable of delivering both the dCas9 module and the guide RNA (17). The plasmid contains two long terminal repeats (LTR) regions, which enclose the spCas9 (replaced with a dCas9), a spacer (used to facilitate the cloning of the gRNA) and a Puromycin selection marker (Puro) (Figure 2).

In summary, we wanted to build a unified and efficient system of all these tools, which until now, taken all together, had a heterogeneous delivery method with a non-coherent design.

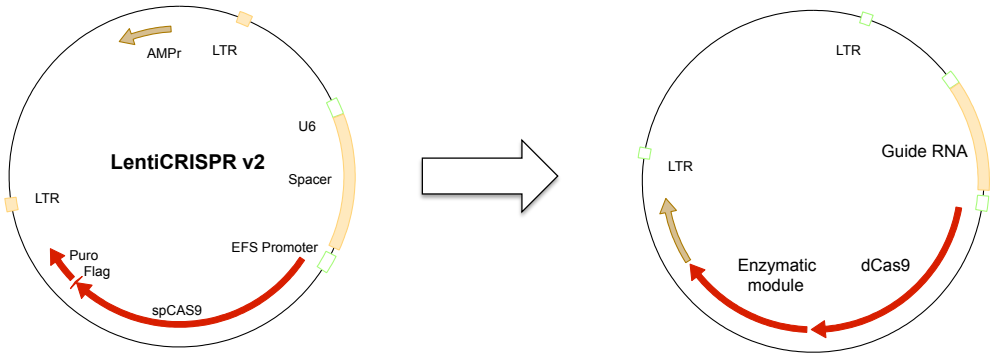


Figure 2 - The universal vector
Reengineering of LentiCRISPR second-generation plasmid (17). The spCas9 module, driven by the elongation factor 1alpha promoter (EFS), is substituted by an epigenetic-module-dCas9 fusion protein. The ability to add a guide RNA in a separate transcriptional unit (U6 promoter) remains unchanged.

Red Puro – And there was light

In a viral infection the ratio between transforming units (TU) and the number of cells (multiplicity of infection, MOI) determines both the number of infection events and how many provirus copies are incorporated within the cell genome. Thanks to the Poisson distribution equation, these variables can be calculated (18). Therefore, to have a controlled experiment for these two parameters it is necessary to measure the TU of the virus preparation. To accomplish this task one of the most convenient methods is based on the use of a fluorescent marker in a FACS analysis setting. Therefore, we exchanged the Puro cassette with the red-fluorescence-puromycin fused sequence (RedPuro), to add a fluorescent signal to the system. Protein domains are modular and tend to remain functional, even if embedded in a different amino acid sequence. In this case, it has been proved that by fusing the red fluorescent protein (RFP) in-frame to the Puro cassette, both enzymatic activities are retained (19) (Figure 3). The RedPuro is embedded in the same transcriptional unit of spCas9, rendered bi-cistronic by the 2A peptide (2AP) (20). This design reduces the payload size, which has the advantage of maintaining good packaging efficiency.

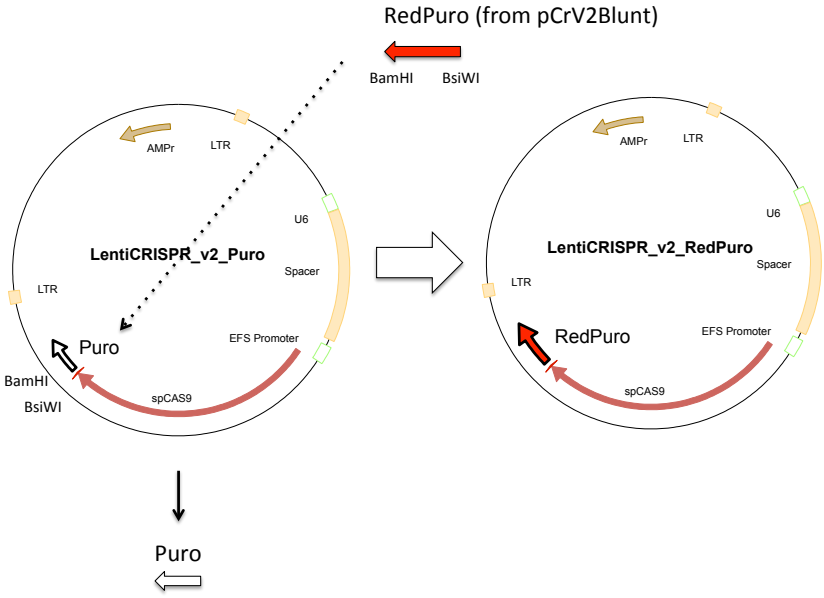


Figure 3 – LentiCRISPR second-generation RedPuro design
Design of the sticky end cloning (BamHI, BsiWI) used to modify the Puro cassette with a RedPuro. We inserted RedPuro into an intermediate vector (pCrv2Blunt, Invitrogen) before moving it into the LentiCRISPR second-generation plasmid. The dotted arrow indicates the cloning site. The continuous arrow shows the excision of the insert. The right panel represents the final clone.

To test the expression of the RedPuro cassette we transfected LentiCRISPR second-generation RedPuro plasmid into the HEK 293T cells and verified the expression of RFP fused with Puro by using an epifluorescent microscope. (Figure 4)

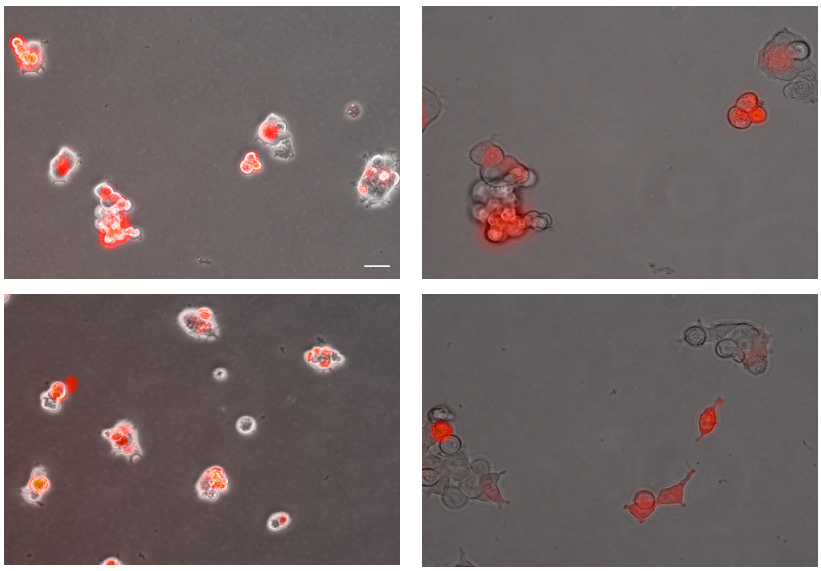


Figure 4 - HEK 293T transfected with LentiCRISPR second-generation RedPuro
The image shows four different fields acquired in bright light merged with the epifluorescent 564 nm signal. The scale unit is shown in the first panel and represents 100 µm.

Epigenetic modules – the gateway to change

To have a complete toolbox, we included six of the most relevant epigenetic modules used to regulate gene expression. We targeted DNA methylation as well as histone acetylation and methylation, in addition to generic transcriptional activation.

More specifically, we targeted DNA methylation, by including in the toolbox Tet1 and DNMT3a, which respectively demethylate and methylate DNA (11,12). We also targeted histone modifications, namely acetylation, deacetylation (P300 and HDAC) (13,14) and methylation (KRAB) (15). Finally, we added to the toolbox a transcriptional activator that acts by recruiting different cofactors (VP64) (13,16) (Table 1).

Epigenetic Module	Activity	Mechanism
VP64	Activation of transcription	Sequential recruitment of co-factors
P300	Activation of transcription	Histone acetylation
Tet1	Activation of transcription	DNA demethylation
DNMT3a	Repression of transcription	DNA methylation
HDAC	Repression of transcription	Histone deacetylation
KRAB	Repression of transcription	Histone methylation and deacetylation

Table 1 - Epigenetic modules
The table represents the list of the epigenetic toolbox components. VP64, P300 and Tet1 activate, instead DNMT3a, HDAC and KRAB repress transcription. Each tool affects the chromatin in a different way, including a direct effect on the DNA or through a modification of the histones, as well as by recruiting co-factors.

In summary, our work consists of exchanging the spCas9 with the dCas9 epigenetic modules in the LentiCRISPR second-generation RedPuro vector (Figure 3). These tools are compatible with the original strategy to clone the gRNAs between BsmBI sites, developed by Zhang et Al. (21).

In the next sections we will detail the strategies and challenges we had to face to accomplish this goal.

VP64 and P300 – the entry-level job

The cloning of VP64 and P300 had a similar design and will be detailed in the following sections.

VP64

Concerning VP64 module, our strategy was to insert VP64-dCas9 into the backbone of the LentiCRISPR second-generation RedPuro, using XbaI and BamHI restriction sites (Figure 5). However, because there is an extra BamHI site in the VP64 sequence, we first had to destroy it by mutation (A4255C), leaving the translation unchanged by taking advantage of the genetic code redundancy.

To change the nucleotide sequence, we used the PCR technique. Briefly, we opted to use as a template the VP64-dCas9 vector from Addgene (ID 61425) to obtain an amplicon. Then we designed two pairs of primers generating two different amplicons, harboring the mutation in an overlapping sequence common to both products (see Figure 5, panels A and B).
By using these two molecules as a template, we joined them together in the final mutated product by a second PCR amplification using only the external primers (see Figure 5, panel C).
We used a proofreading Taq-polymerase to avoid eventual out-of-frame mutations introduced by the Taq-polymerase propensity to add an extra dATP at the 3'end of the amplicon.

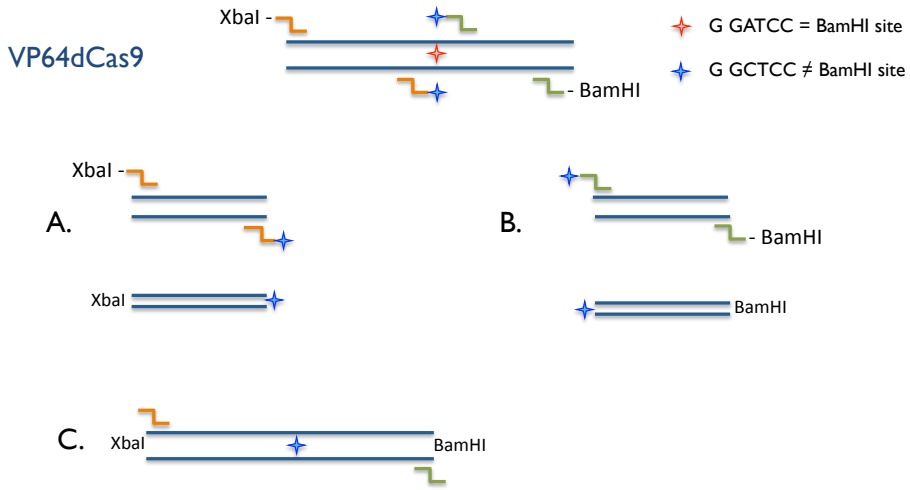


Figure 5 - PCR strategy to delete BamHI site
VP64dCas9 PCR used to mutate the adenine in position 4255 from the ATG to a cytosine (A4255C). Panel A: First PCR reaction based on 5' part of amplicon. 5' primer adds XbaI restriction site, 3' primer inserts A4255C mutation. Panel B: Second PCR reaction based on 3' part of the amplicon. 5' primer inserts A4255C mutation, 3' primer adds BamHI restriction site. Panel C: Third PCR reaction that merges the two amplicons resulting from the previous PCRs. The blue star shows that the result of the third PCR contains the desired mutation.

The external primers also harbored an extra sequence introducing the corresponding XbaI and BamHI sites. The direct PCR cloning could be problematic because the digestion of a PCR product is very inefficient, especially with large amplicons (22). Therefore, the final product was cloned into an intermediate vector (pCRv2Blunt), designed for Blunt-end cloning. In the last step we prepared both the backbone and the insert by digestion with the corresponding restriction sites and gel-purification. Finally, the two molecules were pasted together by a fast ligation reaction, following a homemade protocol (see Materials and methods).

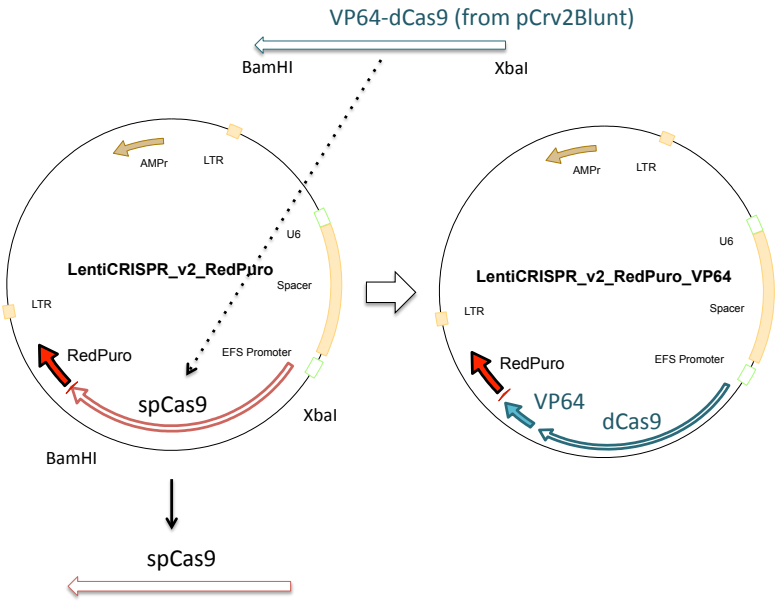


Figure 6 - Design for LentiCRISPR second-generation RedPuro VP64
To clone the VP64-dCas9 module we used as the backbone the LentiCRISPR second-generation RedPuro plasmid and as the insert the amplicon VP64-dCas9, cloned from an intermediate vector (pCrv2Blunt). The restriction sites used for ligation are indicated in the figure. The dotted arrow points the cloning site. The continuous arrow indicates the excision of the insert. The right panel represents the final clone.

P300

To clone the P300-dCas9 module, we used a strategy identical to the one outlined in the previous section, with the following differences (Figure 7):

- The template was taken from Addgene (ID 61357);
- The internal restriction site to be mutated was BglII (C4907T);
- The flanking restriction sites of the insert were BglII and SpeI, harboring a compatible overhang with respectively BamHI and XbaI; this design was necessary because there were multiple internal BamHI and XbaI sites.

The P300 sequence harbors an internal BsmBI site as well (see Figure 8). Therefore to clone the gRNA we have to modify the standard Zhang protocol (21), as outlined below:
Instead of using a pair of primers to be annealed and originate the BsmBI overhangs, a synthetic double strand DNA of approximately 700 bp should be synthesized with the following design (Figure 8, panel B):

- At the beginning the gRNA sequence;
- Followed by the part of dCAS9 lost after digestion of LentiCRISPR second-generation RedPuro P300 with BsmBI;
- A strategically placed BsmBI at both ends of the sequence, which recreates the sticky ends compatible with the digested vector.

The gBlocks fragment should be intermediate-cloned with a Blunt-end vector, such as PCRv2Blunt, then excised with BsmBI and finally cloned into the p300-dCas9 module, as well as digested with BsmBI (Figure 8).

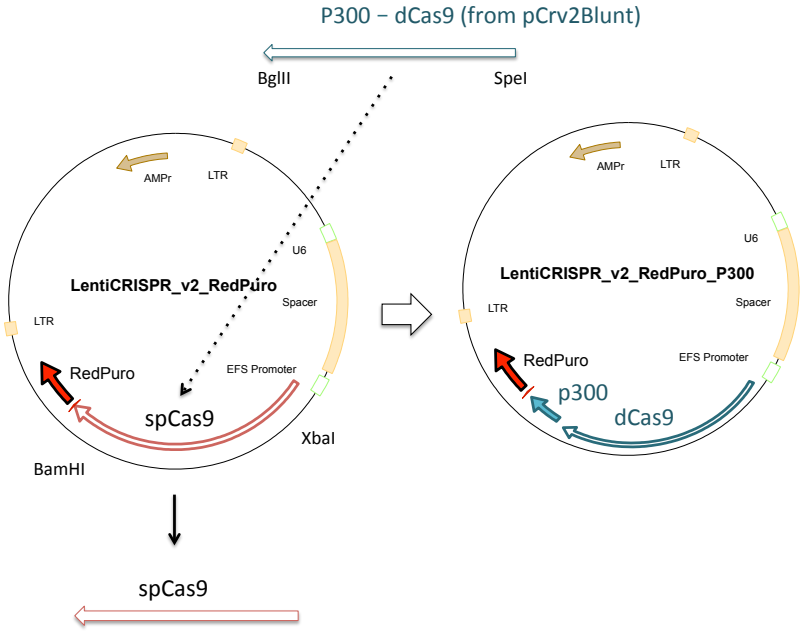


Figure 7 - Design for LentiCRISPR second-generation Red Puro P300
To clone the P300-dCas9 module we used as the backbone the LentiCRISPR second-generation RedPuro plasmid and as the insert the amplicon p300-dCas9, cloned from an intermediate vector (pCrv2Blunt). We took advantage of the complementarity of restriction enzymes (XbaI-SpeI, BamHI-BglII). The dotted arrow indicates the cloning site. The continuous arrow shows the excision of the insert. The right panel represents the final clone.

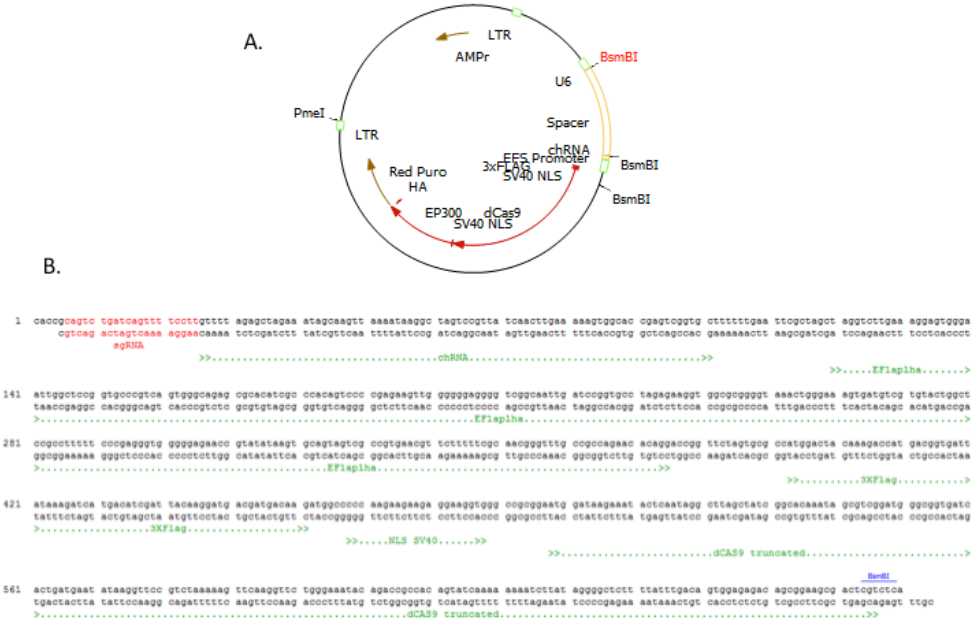


Figure 8 - gRNA cloning strategy variation for P300
Panel A: Map of LentiCRISPR second-generation RedPuro P300 showing an extra BmsBI site inside the coding sequence. Panel B: An alternative design to clone a gRNA, using a synthetic double strand DNA. In red, an example of gRNA sequence, which has to be modified according to the target.

Tet1 – when the going gets tough…

To clone the Tet1-dCas9 module, we used a similar strategy to the one outlined in the VP64 section, with the following differences:

- The template was taken from Addgene (ID 84475);
- There was no internal restriction site to be mutated;
- The flanking restriction sites of the insert were XbaI and BglII, harboring a compatible overhang with BamHI. This design was necessary because there were multiple BamHI internal sites.

The amplicon was cloned in an intermediate vector and processed exactly as described in the VP64 section (Figure 9).

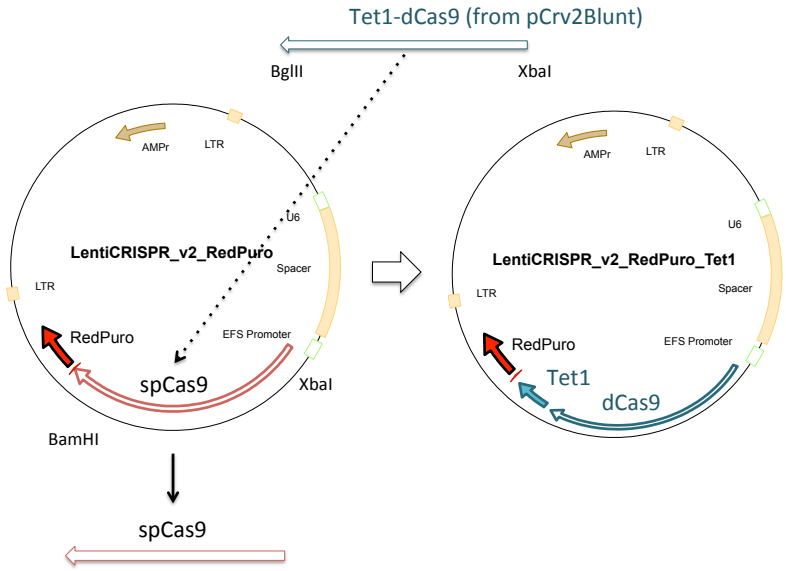
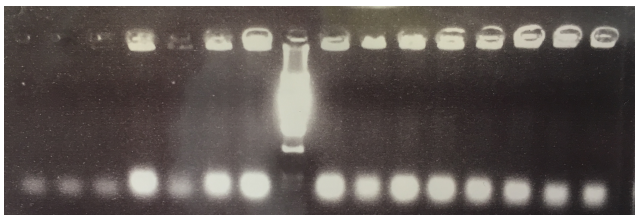


Figure 9 - Design for LentiCRISPR second-generation Red Puro Tet1
To clone the Tet1-dCas9 module we used as the backbone the LentiCRISPR second-generation RedPuro plasmid and as the insert the amplicon Tet1-dCas9, cloned from an intermediate vector (pCrv2Blunt). We took advantage of complementarity of restriction enzymes (BamHI-BglII). The dotted arrow indicates the cloning site. The continuous arrow shows the excision of the insert. The right panel represents the final clone.

During the screening step, Tet1 cloning presented an unexpected challenge. The electroporation consistently generated a large number of apparently false positive colonies. In fact, the standard PCR-based screening of several tens of randomly selected colonies did not detect any real positive (as shown in Figure 10, Panel A). By controlling the PCR screening, we discovered that the reaction with those particular primers and template was not compatible with a successful amplification. For this reason, instead of trying to optimize the reaction, we switched to the old "cracking colonies" method, which does not require any kind of previous manipulation, such as PCR amplification or DNA digestion. Briefly, the method consists of loading on the agarose gel directly a lysated bacterial colony (see Materials and methods). After the run, these samples reveal the presence of a supercoiled plasmid DNA as single or multiple bands. By comparing this profile with an empty backbone reference, it is possible to select any candidate showing a slower mobility on the gel. This method greatly simplifies the screening process and reduces the chance of showing false negatives signals (see Figure 10, Panel B).

A.



B.

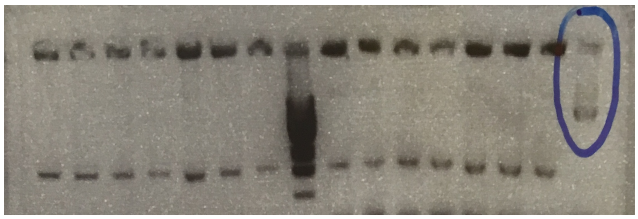


Figure 10 – Screening of candidate clones for Tet1 module
Panel A: an example of negative PCR-based screening. The large bands on the bottom of the gel represent the unincorporated primers. There are no amplicons compatible with the expected size. Panel B: Cracking method screening showing a positive candidate (blue circle) with a slower electrophoretic mobility compared to the negative clones, due to the presence of the correctly cloned insert.

DNMT3a, HDAC and KRAB – …the tough get going

The remaining cloning projects, for DNMT3a, HDAC and KRAB modules, were initially unsuccessful and required a careful troubleshooting process. We speculated that the main problem we were facing was due to a toxic effect of these dCas9 modules. In fact, the LTR promoter contains bacteria cryptic sequences, which allow the expression of the dCas9-fused protein inside the bacterial cells. This event leads to the creation of an environment that favors recombinant clones (23). We also considered that another potential problem was the large size of these three modules, therefore all together creating a molecular perfect storm. Based on both these conjectures, we devised a multi-step mitigation strategy:

- The origin of replication of LentiCRISPR second-generation RedPuro is based on PUC18, which at 37°C determines a high copy number of plasmids, while below 30°C drops to a lower one (24). Therefore, we proceeded with a clonal selection at room temperature instead of 37°C to mitigate the toxic effect;
- It has been reported in the literature that an inter-chain ligation between two backbones DNA strands generates a circular plasmid containing repetitive sequences, which are subsequently removed by the enzymatic apparatus of the bacterial cells (25). Therefore, we reasoned that dephosphorylation of the backbone would help to reduce recombinant clones, by preventing the ligation between them;
- It is well known that the size of the insert reduces the efficiency of the cloning. We took advantage of some unique internal restriction sites in dCas9, which allow us to split the insert in two smaller fragments. Therefore we proceeded with a two-step cloning procedure (see Figures 11, 12, 13).

Based on this strategy, we modified the cloning design used for VP64, P300 and Tet1 as detailed in the following sections.

DNMT3a

To clone the DNMT3a-dCas9 module, we used a similar strategy to the one outlined in the Tet1 section, with the following differences:

- The template was taken from Addgene (ID 84476);
- The insert was divided into two pieces
    - The first fragment was cloned with XbaI, EcoRV restriction enzymes in the corresponding sites of LentiCRISPR second-generation RedPuro plasmid (Figure 11); this construct was used as a backbone to clone the second fragment (see next point)
    - The second fragment was cloned using EcoRV, BglII for the insert, and EcoRV, BamHI for the backbone taken from the previous cloning;
- To further decrease the background, we digested the ligation product before electroporation with BglII and BamHI restriction enzymes. We reasoned that a correct ligation between these compatible overhangs would destroy both sites, therefore the digestion would linearize only recombined constructs, making them unable to transform bacterial cells.

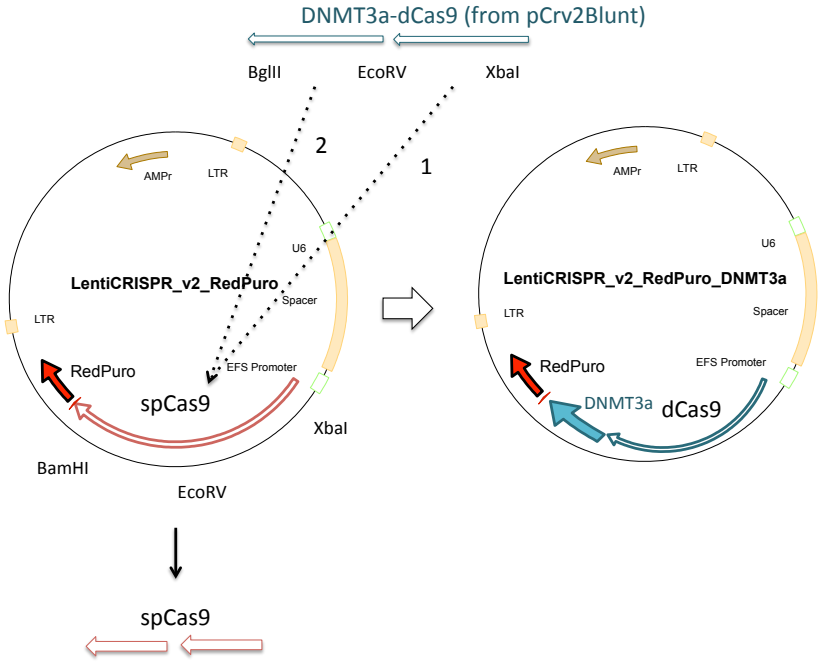


Figure 11 - Design for LentiCRISPR second-generation Red Puro DNMT3a
To clone the DNMT3a-dCas9 module the large insert was split and cloned in two parts. The restriction sites are indicated in the figure. The numbers indicate the order of cloning. The dotted arrow indicates the cloning site. The continuous arrow shows the excision of the insert. The right panel represents the final clone.

HDAC

HDAC was taken from Addgene (ID 98591). The plasmid, being a Lentiviral vector, was partially compatible with LentiCRISPR second-generation RedPuro. Therefore, by sacrificing the presence of the RedPuro selection marker, we could directly sub-clone HDAC into our backbone, allowing the project to be completed in the allocated timeframe. Briefly, the strategy was to clone first the fragment going from NheI to BamHI restriction sites and then the fragment going from BamHI to PmeI restriction sites, removing the RedPuro cassette up to the 3' LTR. The corresponding sites in the insert carry over the 3'end of dCas9-HDAC, an out-of-frame 2A peptide with Blasticidin coding sequence, followed by the 3' LTR, which was excised from the backbone. The ligation reconstitutes a functional Lentivirus but the final product does not have a selection marker (Figure 12). A future project will reintroduce the RedPuro part (see discussion).

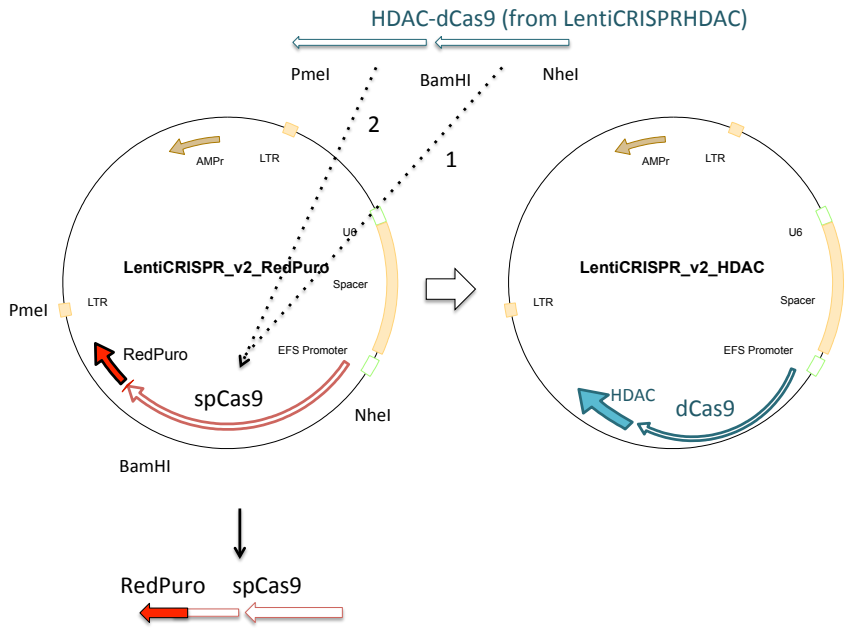


Figure 12 - Design for LentiCRISPR second-generation HDAC
To clone the HDAC-dCas9 module the large insert was split and cloned in two parts. The restriction sites are indicated in the figure. The numbers show the order of cloning. The dotted arrow indicates the cloning site. The continuous arrow points to the excision of the insert. The right panel represents the final clone, which shows that the RedPuro is lost.

KRAB

The KRAB module was taken from Addgene (ID 60954). As for HDAC, it was possible to avoid PCR amplification and directly subclone KRAB-dCas9 into the LentiCRISPR second-generation RedPuro plasmid. The strategy, similarly to DNMT3a and HDAC projects, was to clone the insert by dividing it into two fragments, to be inserted sequentially. The first piece was flanked by EcoRV, BamHI and the second one by XbaI, EcoRV. However, XbaI was not available in the source plasmid, therefore we added it by creating an intermediate clone into the BlueScriptSK+, using BamHI and EcoRV restriction sites. In this manner, an XbaI site was added next to BamHI. After this intermediate clone, we excised EcoRV, XbaI and completed the project by ligating it to the plasmid carrying the first fragment (Figure 13).

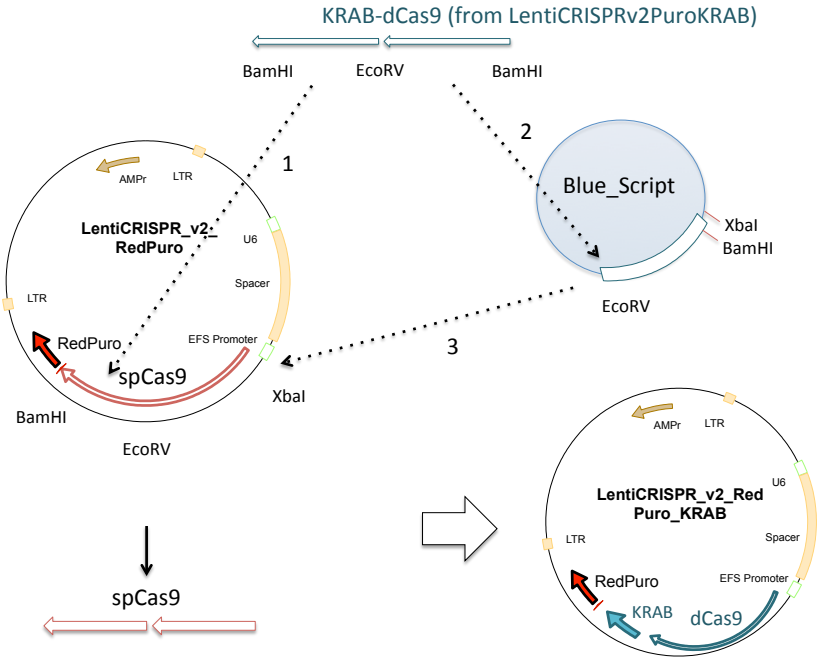


Figure 13 - Design for LentiCRISPR second-generation RedPuro KRAB
To clone the KRAB-dCas9 module the large insert was split in two parts. The first one was directly cloned into LentiCRISPR second-generation RedPuro (arrow 1). The second one was cloned into the intermediate vector BlueScriptSK+ (arrow 2). Subsequently, was excised using the external sites (XbaI, EcoRV) and cloned into the corresponding sites (arrow 3). The right panel represents the final clone.

Discussion

Pediatric sarcomas – there is still so much more to do

Soft tissue and bone malignancies are frequently initiated by a specific chromosomal translocation, which constitutes a "signature" of the disorder and leads to the formation of a fusion gene, which in most cases encodes an aberrant transcription factor or transcriptional regulator (26,27).

Recent studies have shown that some of these aberrant proteins affect the transcriptome by perturbing the chromatin state (2,4,5,11–16). The chemical modifications of the chromatin, including acetylation, methylation and phosphorylation of proteins and DNA, are commonly referred to as the epigenome, which constitute a control layer of genomic activity.

Chromatin modifications are emerging as one of the most important pathways that lead to the transformation to cancer cells, which is dependent on the initial state of the cell of origin and the activity of the aberrant protein (26,27). In the case of Ewing sarcoma, the second most common malignancy of bone in children and young adults, permissive cells for the activity of the fusion protein have been identified as mesenchymal stem cells (MSCs). In these cells the aberrant fusion protein (EWS-FLI-1) behaves as an oncogene, modifies the chromatin structure of a range of promoters and enhancers and behaves as a pioneer factor (28). Nevertheless, mechanisms that determine how these modifications act are currently unknown and their elucidation will be critical toward developing more effective therapeutic management of Ewing sarcoma.

For several other translocation-dependent sarcomas, including synovial sarcoma, rhabdomyosarcoma and dermatofibrosarcoma, there is the same necessity to clarify how the perturbation of chromatin works in the context of tumor development. Different genes, including NKX2-2, SOX-2, EGR-2, and NPY-1R have been demonstrated to play a central role in this cascade (29–32). However, more data are necessary to better understand the interplay between these genes and their regulators.

The discovery of nucleases that can target a specific region of a chromosome, including TALENs and CRISPRs, have opened new possibilities not only for genome manipulation but also for studying its regulatory control layer (7–9). In particular, CRISPR-Cas9 has emerged as a groundbreaking tool for its efficiency and simplicity and, by mutating the Cas9 nuclease and adding to it a chromatin modifier, can be used to modulate the epigenetic state of any region of interest.

The purpose of this project is to improve a set of tools based on the above mentioned technology and create a system that will offer the possibility to investigate the effect of the translocations on chromatin regulation. Zhang et al. have created a second-generation CRISPR-Cas9 lentiviral vector, with a single delivery system containing the entire molecular machinery, which has been extensively used by the research community to edit the genome. On the other hand, the tools available for modifying the epigenome were not so advanced; therefore we combined this vector with the most commonly used epigenetic modifiers (2,4,5,11–16). We also introduced an additional feature to the plasmid, by exchanging the puromycin selection marker with a puromycin-red fluorescent fusion peptide. This modification should have the advantage of allowing virus titration by FACS analysis while maintaining the selection marker.

Project's challenges – every problem has a solution

Perhaps unsurprisingly, the project has been extremely challenging. Even by using the most modern cloning technologies, which normally are very efficient, we faced several difficulties. The main problem we faced was the reduced number of clones as well as the extremely high background of false positives.

We reasoned that several factors led to this effect:

- The large size of both vector and insert makes cloning extremely challenging;
- The presence of the LTR regions induce recombination events even in RecA defective bacterial strains;
- Due to the presence of cryptic bacterial transcription activator inside the LTRs, the dCas9-epigenetic modifier is expressed also in the bacterial cells leading to a toxic effect and negative selection activity.

We had to develop a number of coping strategies to manage the above-mentioned difficulties:

- When necessary, we addressed the problem of large inserts by splitting the cloning in two sequential steps;
- We reduced the recombination risk by dephosphorylating the vector, therefore preventing inter-chain ligation events of the backbone, which is the primary cause of the deletion of repetitive regions (25);
- We decreased the toxic effect of dCas9-epigenetic module by switching to a low plasmid copy number condition by setting the growing environment at 25^{o}C;
- Furthermore, to improve the general efficiency of the system, we used a special protocol (see Materials and methods) to make highly competent bacteria cells.

Overall, it was a learning experience to optimize to such a high degree the management of molecular cloning projects, which is an essential skill to have in a research laboratory.

Future perspectives

This is an ongoing project that will benefit from several further improvements:

- Add more epigenetic modules, in particular EZH2
- Modify VP64 to the second-generation activators (33,34)
- Add a selection marker (RedPuro) to the HDAC vector

The toolbox is ready to be used in our lab in new research projects focused on Ewing and synovial sarcoma. We already have designed several gRNAs to test (see Materials and methods).

Materials and methods

LentiCRISPR second-generation plasmid RedPuro cloning

The RedPuro sequence was first amplified by plvmiRcontrol plasmid (BiOSETTIA), using the following primers:

- RedPuroF (5'GGATCCGGCGCAACAAACTTCTCTCTGCTGAAACAAGCCGGAGATGTCGAAGAGAATCCTGGACCGATGGTGAGCAAGGGCGAGGAG3')
- RedPuroR (5'AGTTCTTGCAGCTCGGTGAC3')

Subsequently the sequence was cloned in the pCRIIBlunt vector, according to the manufacturer's instructions. From this intermediate vector the RedPuro was digested with BamHI, BsiWI and cloned into the corresponding sites of the LentiCRISPR second-generation plasmid.

RedPuro transfection

The functionality of the RedPuro vector was tested by transfecting HEK 293T, using the FuGENE system (Promega) with 2 µg of plasmid and a 3:1 ratio of liposomes:DNA, according to the manufacture instructions. Puromycin resistance was assessed by selecting the cells at the concentration of 1 mg/mL G418. The fluorescence was tested on an epifluorescent microscope at 20x magnification objective, using a mercury lamp to acquire the signal at 564 nm.

PCR primers and conditions

VP64 PCR amplification:

The following primers were used to amplify a set of two overlapping amplicons:

1. VP64XbaI (5'AATTCTAGAGGCCACCATGAAAAGGCCGGCGGCCACG3')
2. VP64A8116CF (5'GTGGCGGCCGCTGGCTCCGGACGGGCTGACGCATTGGAC3')

3. VP64A8116CR (5'GCGTCAGCCCGTCCGGAGCCAGCGGCCGCCACCTTCCTC3')
4. VP64BamHI (5'GGCGGATCCTGTACAGTTAATCAGCATGTC3')

The first amplification was made using primers 1 and 3 with the following conditions: 30 cycles composed of 20'' at 98°C, 30'' at 65°C, 1'10'' at 72°C, using a Phusion HF buffer (New England Biolabs).

The second amplification was made using primers 2 and 4 with the following conditions: 30 cycles composed of 20'' at 98°C, 30'' at 60°C, 10'' at 72°C, using a Phusion HF buffer (New England Biolabs).

The final product was made using the external primers (1 and 4), harboring XbaI and BamHI, by gel purifying the set of the overlapping amplicons and using them as a template in a reaction with the following conditions: 30 cycles composed of 20'' at 98°C, 30'' at 65°C, 1'5'' at 72°C, using a Phusion GC buffer (New England Biolabs).

P300 PCR amplification:

The following primers were used to amplify a set of two overlapping amplicons:

1. P300SpeI (5'CCGACTAGTGCGCCATGGACTACAAAGAC3')
2. P300C5796TF (5'TGCGGAAGAAAGATGCATCAGATTTGTGTCCTTCACCATG3')
3. P300C5796R (5'ATCTCATGGTGAAGGACACAAATCTGATGCATCTTTCTTC3')
4. P300BglII (5'GCCAGATCTAGAAGCGTAGTCCGGAACGTCGTA3')

The first amplification was made using primers 1 and 3 with the following conditions: 30 cycles composed of 20'' at 98°C, 30'' at 60°C, 1'20'' at 72°C, using a Phusion HF buffer (New England Biolabs).

The second amplification was made using primers 2 and 4 with the following conditions: 30 cycles composed of 20'' at 98°C, 30'' at 60°C, 20'' at 72°C, using a Phusion HF buffer (New England Biolabs).

The final product was made using the external primers (1 and 4), harboring XbaI and BamHI, by gel purifying the set of the overlapping amplicons and using them as a template in a reaction with the following conditions: 30 cycles composed of 20'' at 98°C, 30'' at 56°C, 1'30'' at 72°C, using a Phusion HF buffer (New England Biolabs).

Tet1 PCR amplification:

The following primers were used to amplify:

1. Tet1XbaI (5'GGCTCTAGAGCCACCATGGACAAGAAGTATTCTATCG3')
2. Tet1BglII (5'GGCAGATCTGACCCAATGGTTATAGGGCCCCGCAACGTGTG3')

The reaction had the following conditions: 30 cycles composed of 20'' at 98°C, 30'' at 60°C, 1'45'' at 72°C, using a Phusion HF buffer (New England Biolabs).

DNMT3a PCR amplification:

The following primers were used to amplify:

1. DNMT3aXbaI (5'CCATCTAGAGCCACCATGGACAAGAAGTATTCTATCG3')
2. DNMT3aBglII (5'CCTAGATCTCACACACGCAAAATACTCCTTC3')

The reaction had the following conditions: 30 cycles composed of 20'' at 98°C, 30'' at 60°C, 1'45'' at 72°C, using a Phusion HF buffer (New England Biolabs).

Fast ligation protocol

The ligation was done for 10 minutes at 25°C using T4 ligase (New England BioLabs) with a homemade buffer containing PEG 8000 (Sigma) (35).

Protocol for super-competent cells

Stbl3 cells (Invitrogen) were made electrocompetent by using the established protocols (36,37). Briefly, a single colony cell was grown at 18°C for 36h in 500 mL of SOB medium. At the optical density of 0.25 at 600 nm the cells were harvested at 4°C, washed three times in a deionized water and resuspended in 7% DMSO in ultrapure water and stored at -80°C after snapshot freezing in liquid nitrogen. The competency achieved was 10^9 colonies/µg DNA.

Electroporation protocol

50 µl of competent cells were used in a 0.1 mm electroporation cuvette (BioRad) at 1.8 kV to transform the bacteria with our ligation product precipitated in butanol, washed with a 70% ethanol and resuspended in 4 µl of ultrapure water.

Pre-screening cracking colonies protocol

The cracking colonies method was based on the following protocol (38). Briefly, the colonies were first streaked on an agarose plate and growth overnight at 37°C. Subsequently, 30 µL of the following buffer (Composition 2x: 20% sucrose, 200 mM NaOH, 120 mM KCl, 10 mM EDTA, 0.5% SDS, 0.1% bromophenol blue) were diluted and pre-warmed at 37°C and aliquoted in a 384 PCR plate (30 µL). A small portion of bacteria was collected with a 10 µL pipette-tip and washed in the buffer. After a 5' incubation at 37°C and a 5' incubation in ice, 15 µL of the samples were loaded onto an agarose gel containing 1:10'000 of Red gel (Biotium).

Dephosphorylation protocol

The dephosphorylation was done for 10 minutes at 37°C, followed by 5 minutes at 65°C to inactivate the enzyme. We used FastAP enzyme (Thermo Fisher Scientific) with the correspondent buffer FastAP buffer (Thermo Fisher Scientific).

Some gRNA oligonucleotides to test the toolbox

Examples for DNMT3a and KRAB:

- SOX2 gene
  gRNA1: CACCTTATCCATCTAACAGGTGGG (forward cloning oligonucleotide)
  AAACCCCACCTGTTAGATGGATAA (reverse cloning oligonucleotide)
  gRNA2: CACCGAAAATTGATTGAACCCAGG (forward cloning oligonucleotide)
  AAACCCTGGGTTCAATCAATTTTC (reverse cloning oligonucleotide)

- EGR2 gene
  gRNA1: CACCCATCCAAATTAACATCAGTG (forward cloning oligonucleotide)
  AAACCACTGATGTTAATTTGGATG (reverse cloning oligonucleotide)
  gRNA2: CACCGACAGACAGGATAATATGGG (forward cloning oligonucleotide)
  AAACCCCATATTATCCTGTCTGTC (reverse cloning oligonucleotide)

References

1. Anderson JL, Denny CT, Tap WD, Federman N. Pediatric sarcomas: translating molecular pathogenesis of disease to novel therapeutic possibilities. Pediatr Res. 2012 Aug;72(2):112–21.

2. Laporte AN, Poulin NM, Barrott JJ, Wang XQ, Lorzadeh A, Vander Werff R, et al. Death by HDAC Inhibition in Synovial Sarcoma Cells. Mol Cancer Ther. 2017 Dec;16(12):2656–67.

3. Perani M, Ingram CJE, Cooper CS, Garrett MD, Goodwin GH. Conserved SNH domain of the proto-oncoprotein SYT interacts with components of the human chromatin remodelling complexes, while the QPGY repeat domain forms homo-oligomers. Oncogene. 2003 Nov 6;22(50):8156–67.

4. Souza BK, da Costa Lopez PL, Menegotto PR, Vieira IA, Kersting N, Abujamra AL, et al. Targeting Histone Deacetylase Activity to Arrest Cell Growth and Promote Neural Differentiation in Ewing Sarcoma. Mol Neurobiol. 2018 Feb 3;

5. Li Y, Li X, Fan G, Fukushi J-I, Matsumoto Y, Iwamoto Y, et al. Impairment of p53 acetylation by EWS-Fli1 chimeric protein in Ewing family tumors. Cancer Lett. 2012 Jul 1;320(1):14–22.

6. About NHGRI: A Brief History and Timeline [Internet]. National Human Genome Research Institute (NHGRI). [cited 2018 May 29]. Available from: https://www.genome.gov/10001763/about-nhgri-a-brief-history-and-timeline/

7. Ledford H. CRISPR: gene editing is just the beginning. Nat News. 2016 Mar 10;531(7593):156.

8. Addgene: CRISPR Guide [Internet]. [cited 2018 May 2]. Available from: https://www.addgene.org/crispr/guide/

9. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the

CRISPR-Cas9 system. Nat Protoc. 2013 Nov;8(11):2281–308.

10. Dominguez AA, Lim WA, Qi LS. Beyond editing: repurposing CRISPR–Cas9 for precision genome regulation and interrogation. Nat Rev Mol Cell Biol. 2016 Jan;17(1):5–15.

11. Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, et al. Editing DNA Methylation in the Mammalian Genome. Cell. 2016 Sep 22;167(1):233-247.e17.

12. Xu X, Tao Y, Gao X, Zhang L, Li X, Zou W, et al. A CRISPR-based approach for targeted DNA demethylation. Cell Discov. 2016 May 3;2:16009.

13. Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat Biotechnol. 2015 May;33(5):510–7.

14. Kwon DY, Zhao Y-T, Lamonica JM, Zhou Z. Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC. Nat Commun [Internet]. 2017 May 12 [cited 2018 May 1];8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5437308/

15. Thakore PI, D'Ippolito AM, Song L, Safi A, Shivakumar NK, Kabadi AM, et al. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat Methods. 2015 Dec;12(12):1143–9.

16. Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK. CRISPR RNA-guided activation of endogenous human genes. Nat Methods. 2013 Oct;10(10):977–9.

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

18. Addgene: Fluorescence Titering Assay [Internet]. [cited 2018 May 29]. Available from: https://www.addgene.org/protocols/fluorescence-titering-assay/

19. Abbate J, Lacayo JC, Prichard M, Pari G, McVoy MA. Bifunctional protein conferring enhanced green fluorescence and puromycin resistance. BioTechniques. 2001 Aug;31(2):336–40.

20. Liu Z, Chen O, Wall JBJ, Zheng M, Zhou Y, Wang L, et al. Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. Sci Rep. 2017 May 19;7(1):2193.

21. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, et al. Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Science. 2014 Jan 3;343(6166):84–7.

22. Jung V, Pestka SB, Pestka S. Efficient cloning of PCR generated DNA containing terminal restriction endonuclease recognition sites. Nucleic Acids Res. 1990 Oct 25;18(20):6156.

23. Lewin A, Mayer M, Chusainow J, Jacob D, Appel B. Viral promoters can initiate expression of toxin genes introduced into Escherichia coli. BMC Biotechnol. 2005 Jun 20;5:19.

24. Lin-Chao S, Chen WT, Wong TT. High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II. Mol Microbiol. 1992 Nov;6(22):3385–93.

25. Bzymek M, Lovett ST. Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. Proc Natl Acad Sci U S A. 2001 Jul 17;98(15):8319–25.

26. 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.

27. Suvà M-L, Cironi L, Riggi N, Stamenkovic I. Sarcomas: genetics, signalling, and cellular origins. Part 2: TET-independent fusion proteins and receptor tyrosine kinase mutations. J Pathol. 2007 Oct;213(2):117–30.

28. Riggi N, Knoechel B, Gillespie SM, Rheinbay E, Boulay G, Suvà ML, et al. EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma. Cancer Cell. 2014 Nov 10;26(5):668–81.

29. Riggi N, Suvà M-L, De Vito C, Provero P, Stehle J-C, Baumer K, et al. EWS-FLI-1 modulates miRNA145 and SOX2 expression to initiate mesenchymal stem cell reprogramming toward Ewing sarcoma cancer stem cells. Genes Dev. 2010 May;24(9):916–32.

30. Fadul J, Bell R, Hoffman LM, Beckerle MC, Engel ME, Lessnick SL. EWS/FLI utilizes NKX2-2 to repress mesenchymal features of Ewing sarcoma. Genes Cancer. 2015 Mar;6(3–4):129–43.

31. Grünewald TGP, Bernard V, Gilardi-Hebenstreit P, Raynal V, Surdez D, Aynaud M-M, et al. Chimeric EWSR1-FLI1 regulates the Ewing sarcoma susceptibility gene EGR2 via a GGAA microsatellite. Nat Genet. 2015 Sep;47(9):1073–8.

32. Hong S-H, Tilan JU, Galli S, Izycka-Swieszewska E, Polk T, Horton M, et al. High neuropeptide Y release associates with Ewing sarcoma bone dissemination - in vivo model of site-specific metastases. Oncotarget. 2015 Mar 30;6(9):7151–65.

33. Chavez A, Tuttle M, Pruitt BW, Ewen-Campen B, Chari R, Ter-Ovanesyan D, et al. Comparison of Cas9 activators in multiple species. Nat Methods. 2016;13(7):563–7.

34. Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, P R Iyer E, et al. Highly efficient Cas9-mediated transcriptional programming. Nat Methods. 2015 Apr;12(4):326–8.

35. Labtimes: Tips and Tricks of the Trade: Fast and Affordable Ligation [Internet]. [cited 2018 May 24]. Available from: http://www.labtimes.org/labtimes/trick/tricks/2013_07.lasso

36. Wu N, Matand K, Kebede B, Acquaah G, Williams S. Enhancing DNA electrotransformation efficiency in Escherichia coli DH10B electrocompetent cells. Electron J Biotechnol. 2010 Sep 15;13(5):0–0.

37. Growth of E. coli at low temperature dramatically increases the transformation frequency by electroporation. [Internet]. [cited 2018 May 29]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC306910/

38. Barnes W. Plasmid detection and sizing in single colony lysates. Science. 1977 Jan 28;195(4276):393–4.

Acknowledgements

First of all, thank you to Professor Ivan Stamenkovic, for letting me know the charming world of fundamental research. Thank you for the enthusiasm, the inspiration and the trust that you gave me from the beginning.

Thank you so much to Carlo Fusco. You had the patience and the kindness to teach me molecular biology starting from scratch. Thank you for your excellent coaching and, above all, for turning difficult moments into funny ones.

Thank you to the entire lab, for welcoming me like home. I spent a great time working with you all and I hope to have the opportunity to do it again.

Thank you to all my family, especially to my mother, my father, and Maria Sole, for constantly supporting me in so many different ways.

Last but not least, thank you to Alexis Renaud, my extraordinary class- and soul- mate. Thank you for being my driving-force since our first year of medical school.