UNIVERSITAT POLITÈCNICA DE VALÈNCIA

ibmcp
Instituto de Biología Molecular y Celular de Plantas

MÁSTER EN BIOTECNOLOGÍA MOLECULAR Y CELULAR DE PLANTAS

Synthetic transcriptional repressors design based on the CRISPR-Cas technology in *N. benthamiana*

TRABAJO DE FIN DE MÁSTER

Curso: 2018/2020

Realizado por:
Blanca Salazar Sarasua

Directores:

Sara Selma García

Diego Orzáez Calatayud

Valencia, enero 2020

A Fran

Gracias al laboratorio 2.10 por acogerme y enseñarme y gracias a todos los que habéis contribuido para que este trabajo fuera posible

ABSTRACT

Deactivated versions of Cas proteins like dCas9 and dCas12a open new posibilities for plant synthetic biology in the realm of negative transcriptional regulation. Here we describe two repression strategies tested on a luciferase reporter gene through transient expression in *Nicotiana benthamiana*. The first one consists of a single active repression domain (BRD, SRDX and KRAB repression domains were used) fused to dCas9 or dCas12a and guided to different positions inside the promoter or the target gene. Positions -35, +51 and +62 from the TSS were selected for dCas9 and positions -165, -66 and -9 were selected for dCas12a as the best working guides. These were then tested on pairs, increasing repression efficiency. KRAB domain was discarded from further assays for lower performance from the other domains. The second strategy tested consisted on the use of a repetitive peptide array called SunTag with the ability to recruit numerous antibody fusions. The SunTag was fused to dCas9 and dCas12a, and its antibody (ScFv) was fused to SRDX and BRD repression domains. Two SunTags were tested, one with 5 amino acids and another one with 22 amino acids as spacers between epitopes. Assays were carried out using previously selected guides alone and in pairs. Results show that dCas12a is a better endonuclease for transcriptional repression than dCas9. Both SRDX and BRD domains work, although SRDX is better for most strategies. Using more than one guide increases repression. SunTag 5aa does not seem to be able to increase repression efficiency, but recruiting more repression domains through the use of SunTag 22aa efficiently enhances repression. All in all, the best strategy out of all of the ones tested seems to be the use of dCas12a fused to the SunTag 22aa with either BRD or SRDX domains.

Key words: dCas9, dCas12a, CRISPR, synthetic biology, transcriptional regulation, repression, SRDX, BRD, KRAB, SunTag.

RESUMEN

Las versiones inactivas de CRISPR/Cas, como dCas9 y dCas12a, abren muchas posibilidades para la biología sintética, entre otras cosas, en el campo de la regulación transcripcional negativa. En este estudio describimos dos estrategias de represión empleando dichas formas inactivas probadas con luciferasa como gen reportero mediante expresión transitoria en *Nicotiana benthamiana*. La primera estrategia consistió en la unión de un único dominio de represión (SRDX, BRD o KRAB) a dCas9 y dCas12a. Se diseñaron y probaron guías en distintas posiciones dentro del promotor y dentro del gen para dCas9 y dCas12a. A partir de los resultados, se seleccionaron las posiciones -35, +51 y +62 (distancia en bases del punto de inicio de la transcripción) para dCas9 y las posiciones -165, -66 y -9 para dCas12a. Estos guías se probaron por parejas, aumentando la represión. El dominio KRAB se desestimó para futuros ensayos debido a no arrojar resultados satisfactorios. La segunda estrategia consistió en el empleo de una sucesión repetitiva de péptidos con la capacidad de reclutar numerosos anticuerpos llamado SunTag. Se probaron dos SunTags, una con 5 amino ácidos como espaciadores entre epítopos y una segunda con 22 amino ácidos. Las SunTags se fusionaron a dCas9 y dCas12a, y sus anticuerpos (ScFv) se fusionaron a los dominios de represión BRD y SRDX. Los ensayos con esta estrategia se llevaron a cabo empleando únicamente los guías seleccionados previamente, solos y en parejas. Los resultados de todos los ensayos mostraron una mayor eficiencia de dCas12a frente a dCas9 en la represión. Dicha eficiencia aumenta al emplear más de un guía. Tanto SRDX como BRD dan buenos resultados, aunque SRDX muestra mejores resultados que BRD en la mayoría de los ensayos. La SunTag 5aa no mostraba mejoras en la represión, pero el reclutamiento de más dominios de represión empleando la SunTag 22aa aumentó considerablemente la eficiencia. Finalmente, la mejor estrategia parece ser el empleo de dCas12a fusionado a la SunTag 22aa, empleando SRDX o BRD como dominios de represión.

Palabras clave: dCas9, dCas12a, CRISPR, synthetic biology, represión, regulación transcripcional, SRDX, BRD, KRAB, SunTag.

ABBREVIATIONS

35S: cauliflower mosaic virus 35S promoter.

aa: amino acids.

BRD: B3 repression domain.

Cas: CRISPR-associated proteins.

CDS: coding sequence.

CRISPR: Clustered regularly interspaced short palindromic repeats.

crRNA: CRISPR RNA.

dCas: deactivated CRISPR-associated proteins.

DNA: deoxyribonucleic acid.

EAR: ethylene-responsive element binding factor-associated amphiphilic repression.

Fluc/RLuc: firefly luciferase and renilla ratio.

GB: GoldenBraid.

HDV: Hammerhead Delta Virus rybozime.

KRAB: Krueppel-associated box.

Lb: *Lachnospiraceae.*

Luc: luciferase.

LB: Luria-Bertani.

MES: 2-(*N*-morpholino)ethanesulfonic acid.

***N. benthamiana*:** *Nicotiana benthamiana.*

PAM: protospacer adjacent sequence.

PCR: polymerase chain reaction.

pDGB: GoldenBraid destination plasmid.

Pnos: nopaline synthase promoter.

Ren: renilla.

RNA: ribonucleic acid.

rpm: revolutions per minute.

sgRNA: small guide RNA.

Sp: *Streptococcus pyrogenes.*

SRDX: EAR repression domain.

SunTag 5aa: SUperNova Tag with 5 amino acids as spacers.

SunTag 22aa: SUperNova Tag with 22 amino acids as spacers.

T35S: cauliflower mosaic virus 35S terminator.

Tnos: nopaline synthase terminator.

tracrRNA: trans-activating crRNA.

TSS: transcription starting site.

INDEX

FIGURES AND TABLES

Introduction

INTRODUCTION

1. Synthetic biology

Synthetic biology is a field of science that involves redesigning organisms and its metabolic pathways by engineering them to have new abilities. It involves putting together long strands of DNA either from other organisms or entirely synthetic and inserting them into the desired organism. It applies some principles of engineering like the standardization of the parts used (Benner & Sismour, 2005).

Transcriptional regulators based on CRISPR-Cas are an important tool for multiple purposes in synthetic biology. Previous results by our laboratory developed various strategies for positive transcriptional regulation (Selma et al., 2019). However, negative transcriptional regulation is still a challenge. The goal of this study is the development of synthetic negative transcriptional regulators based on the CRISPR-Cas system.

2. CRISPR-Cas

The CRISPR-Cas system was chosen for the design of negative transcriptional regulators due to its capabilities and ample experience by our laboratory. CRISPR (Clustered regularly interspaced short palindromic repeats) and CRISPR-associated genes (Cas) encode a defensive mechanism for prokaryotes that provides adaptation and immunity to virus pathogens. CRISPR loci consists of short (30-40 bp), palindromic, repetitive sequences with short spacers in between. These spacers are from viral origin and confer resistance to the virus they belong to. These short sequences, known as small guide RNAs (sgRNAs, crRNAs) are expressed and can be recognized by Cas proteins, which have endonuclease activity. The sgRNAs guide Cas proteins to the targeted site determined by the guide sequence and it is cut (Figure 1). For this process to occur, the targeted site must be preceded by a sequence known as the Protospacer Adjacent Motif (PAM) which is recognized by the Cas protein (Marraffini, 2015).

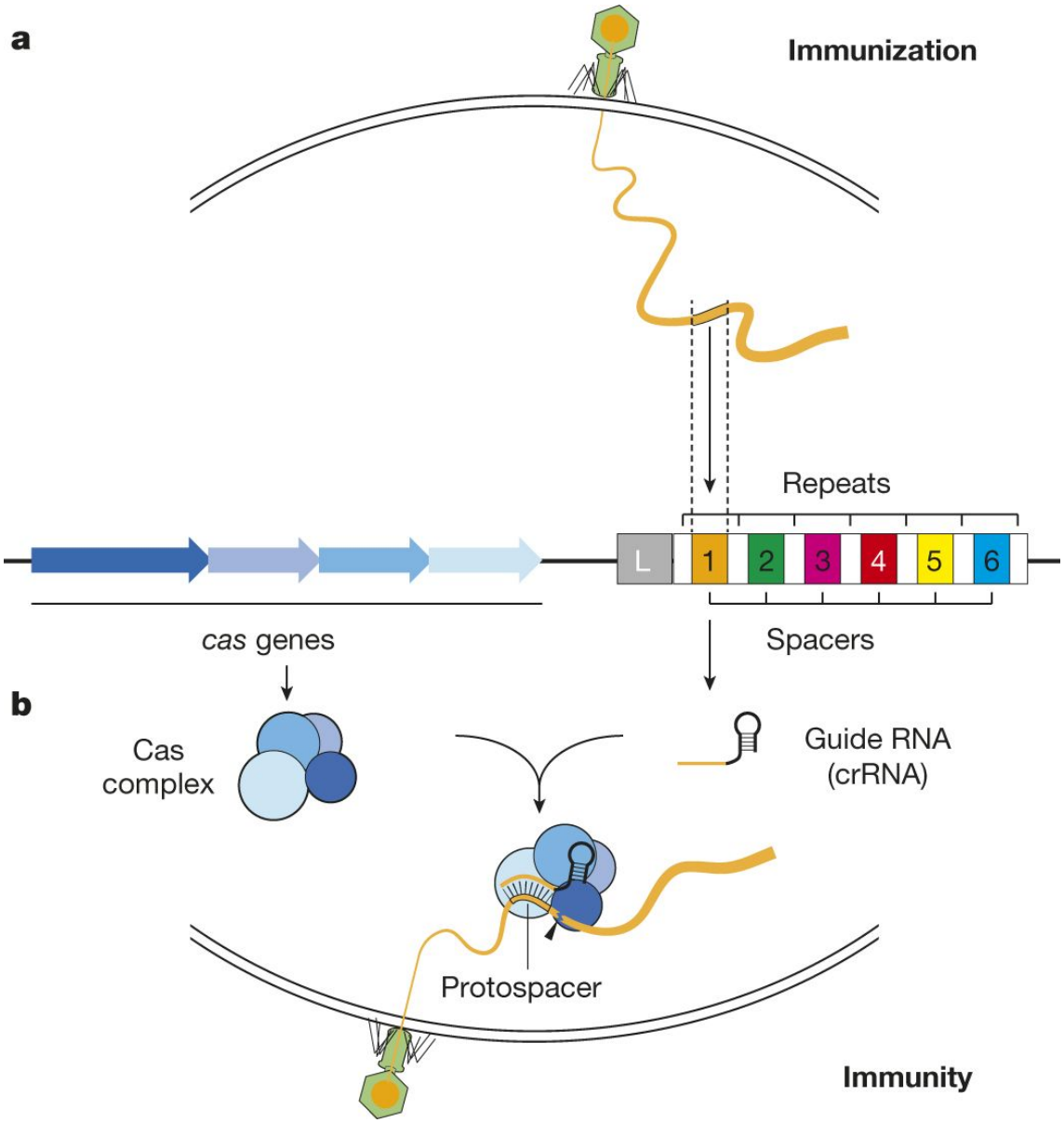

Figure 1. CRISPR immune system. A) Immunization stage, in which a short sequence of the invading virus is captured and integrated as a spacer in the CRISPR loci. B) Immunity stage, in which viruses with sequences already captured are recognized and cut by the Cas proteins. Image from (Marraffini, 2015).

Although this system was originally described as a defensive prokaryotic system (Mojica, Díez-Villaseñor, García-Martínez, & Soria, 2005), its potential as a genetic engineering tool was soon discovered, and its capability to introduce pointed, precise mutations was described by engineering an artificial guide RNA (Jinek et al., 2012).

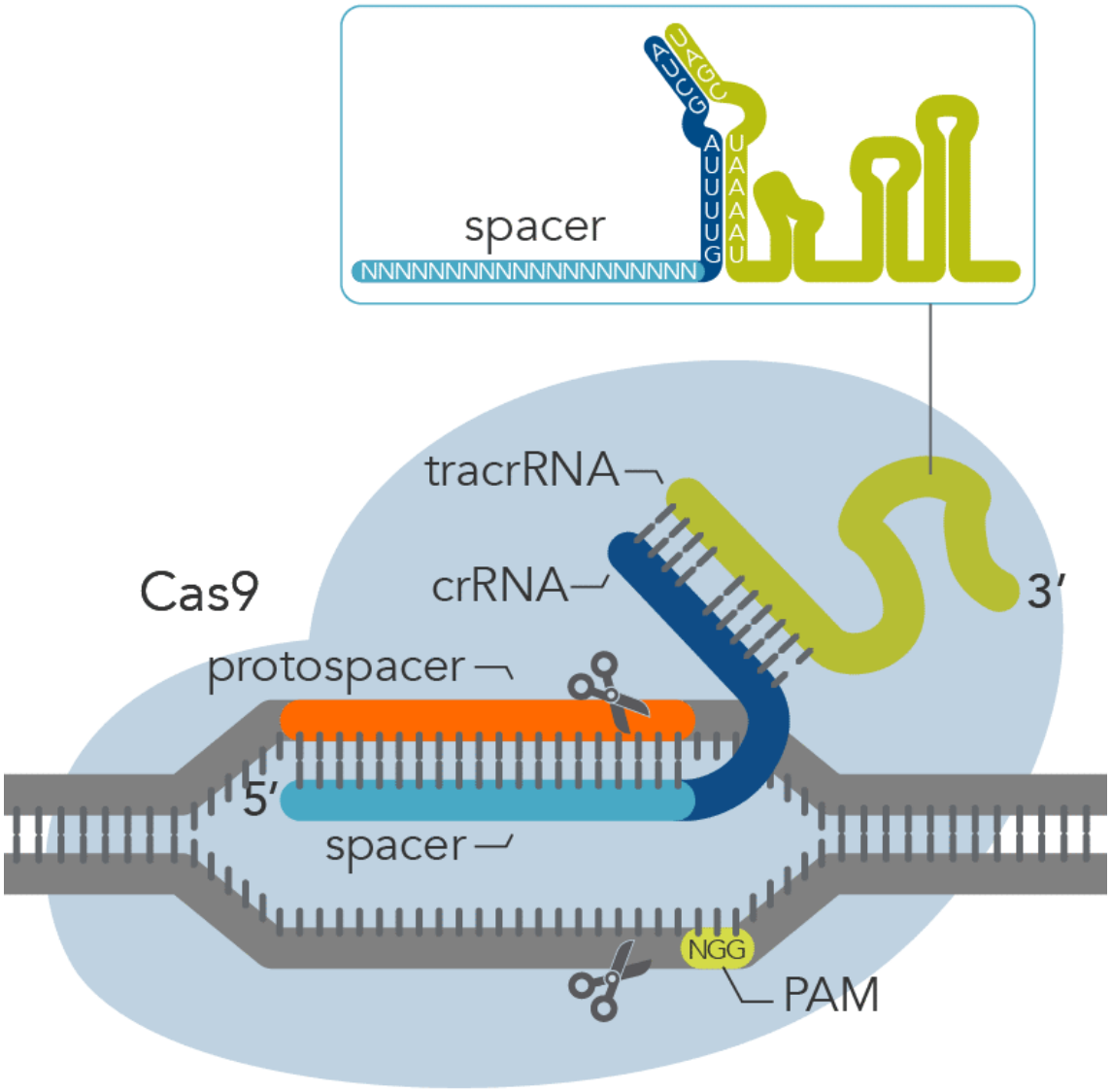

Figure 2. crRNA/tracrRNA hybrid complex. In type II endonucleases, the crRNA is coded by the spacers in the CRISPR loci. In between each spacer, a repeat is inserted, which codes for the end of the crRNA and the tracrRNA. It is partially complementary to the crRNA and is cleaved and processed to make the final guide RNA. Light blue corresponds to the spacer; dark blue and green corresponds to the repeat; blue corresponds to the final crRNA.

The CRISPR-Cas system has since been deeply studied and used as a genetic engineering tool. Several Cas proteins have been identified and characterized. The ones used for this study are known as Cas9 and Cas12a (formerly Cpf1). Cas9 is a type II Cas protein, which means the guide RNA needs a trans-activating crRNA to be processed. This tracrRNA is encoded by the repeats in the CRISPR loci (Figure 2) and is partially complementary to the pre-crRNA forming an RNA duplex. It is then cleaved by an RNA polymerase III to form a crRNA/tracrRNA hybrid that acts as a guide for Type II Cas proteins (Campa, Weisbach, Santinha, Incarnato, & Platt, 2019; Jinek et al., 2012). This can be artificially engineered as part of the guide design by fusing the 3’ end of the crRNA to the 5’ end of

the tracrRNA (Figure 3) (Bortesi & Fischer, 2015). Cas12a, on the other hand, is a type V Cas protein and only requires a single guide RNA to function (Cebrian-Serrano & Davies, 2017).

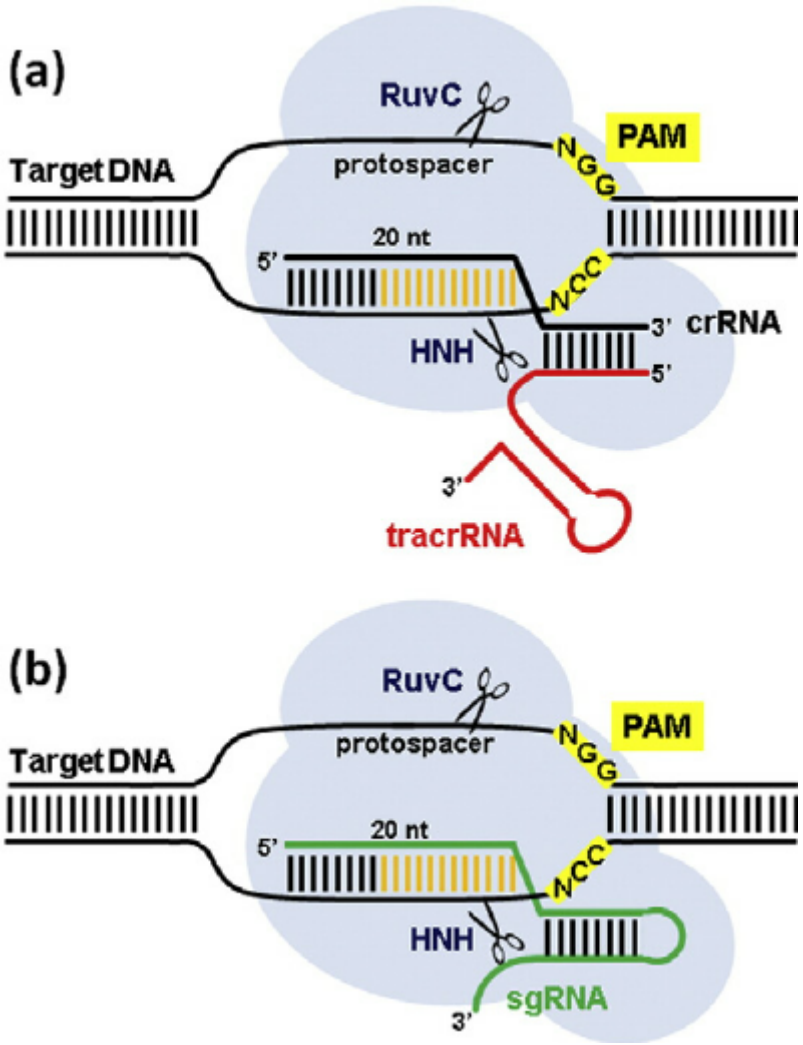

Figure 3. RNA-guided DNA cleavage by Cas9. A) In the native system, type II Cas proteins like Cas9 are guided by a structure formed by a crRNA and a trans-activating RNA (tracrRNA), which stabilizes the structure and activates these Cas proteins to cleave the target. The presence of a protospacer-adjacent motif (PAM) downstream from the target is necessary for Cas9-mediated cleavage. This sequence is different for each Cas protein. B) An artificial structure similar to the crRNA/tracrRNA can be engineered by fusing the 3' end of the crRNA to the 5' end of the tracrRNA, without the need to process the tracrRNA. Image from (Bortesi & Fischer, 2015).

2.1. Cas9

Cas9 is the most used out of all the known Cas endonucleases, belonging to type II Cas proteins. *Streptococcus pyrogenes* Cas9 (SpCas9) is the most commonly used and shows high levels of activity, but has some shortcomings. It has been shown to cause mutagenesis at genomic sequences resembling the target sequence (off-target), it needs an NGG PAM and its size can prove to be a problem for some purposes. However, it is

the most researched Cas protein and there is a lot of knowledge on its working conditions as well as improvements on its sequence to use for different purposes (Bortesi & Fischer, 2015; Cebrian-Serrano & Davies, 2017; Sanson et al., 2018; Vazquez-Vilar et al., 2016).

2.2. Cas12a (Cpf1)

Cas12a, also known as Cpf1, a type V Cas endonuclease that has been used for diverse projects. Although not as much information is known about this endonuclease, its main interest is the differences in structure and traits from Cas9. Since it is not a type II endonuclease, it operates with a single guide RNA, without the need for an additional tracrRNA, unlike Cas9. Its PAM is also different, and it has been defined as TTTV, where V being either A, G or C. This could make it an interesting option for T-rich genomes. Cleavage is also different, and while Cas9 leaves blunt ends, Cas12a leaves overhangs. It is also smaller than Cas9 (Figure 4). This differences make Cas12a an interesting alternative to use for certain projects (Cebrian-Serrano & Davies, 2017).

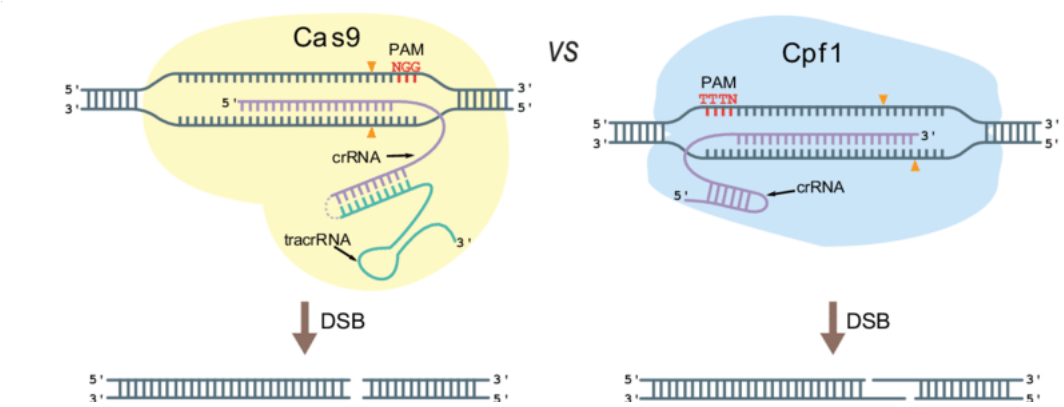

Figure 4. A comparison between Cas9 (left) and Cas12a (right). Cas9 requires an additional tracrRNA to activate, while Cas12a only needs a single guide RNA. PAM sequence and location is different, with Cas9 having NGG as a PAM in the 3' end and Cas12a having TTTV as a PAM on the 5' end. DNA cleavage is also different, leaving blunt ends with Cas9 and overhangs with Cas12a. Image from (Vanegas, Jarczynska, Strucko, & Mortensen, 2019).

These two endonucleases were chosen for this study due to the mentioned differences between them and previous results by our laboratory (Bernabé-Orts et al., 2019; Sarrion-Perdigones et al., 2013; Selma et al., 2019).

2.3. Transcriptional regulation

Aside from gene pointed mutations and editing, CRISPR-Cas systems have also been repurposed to allow for different engineering applications like transcriptional regulation. This is achieved by mutating the nuclease domain (RuvC, NHN) of the Cas endonuclease to obtain nuclease-deactivated Cas proteins (dCas) that cannot cleave but can still bind to DNA. This allows for the use of different strategies involving activation

or repression domains to efficiently regulate transcription (Dominguez, Lim, & Qi, 2016; Lo & Qi, 2017; Miao, Zhao, Qian, & Lou, 2019).

The objective of this study is to use deactivated forms of Cas9 and Cas12a (dCas9, dCas12a) with different strategies to efficiently repress transcription. *Streptococcus pyrogenes* Cas9 (SpCas9) and *Lachnospiraceae* Cas12a (LbCas12a) were used (Table 1).

Table 1. Endonucleases used in this study.

Endonuclease	Class and type	Size (aa)	PAM sequence	Reference
***Lachnospiraceae* Cas12a (LbCas12a)**	Class II, type V	1228	TTTV	(Cebrian-Serrano & Davies, 2017)
***Streptococcus pyrogenes* Cas9 (SpCas9)**	Class II, type II	1368-1424	NGG	(Cebrian-Serrano & Davies, 2017)

3. Repression domains

A lot of repression domains have been described for plants and other organisms, with different modes of action. Genetic and epigenetic repression can be a powerful tool for the remodeling of metabolic pathways, which is one of the aims of synthetic biology, as stated previously (Benner & Sismour, 2005).

For this study, three repression domains were tested and compared. These domains are the B3 repression domain (BRD), the EAR repression domain (SRDX) and the Krueppel-associated box (KRAB) domain.

The EAR motif (Ethylene-responsive element binding factor-associated amphiphilic repression) was the first active repression motif reported in plants. It is one of the principal mechanisms of plant gene regulation and the most predominant form of transcriptional repression motif identified in plants. It is defined by the consensus pattern of LxLxL or DLN xxP. Discoveries of co-repressors interacting with SRDX support a model of repression via recruitment of chromatin remodeling factors to facilitate an epigenetic regulation of gene expression (Kagale & Rozwadowski, 2011).

BRD is a plant domain that consists of 15 amino acids, needing only 8 of these 15 for its repression activity. It was described in *Arabidopsis thaliana*, and doesn't have any resemblance to EAR motifs or other known repression domains. As an active transcriptional repressor, it interacts directly with the transcription machinery to hinder translation (Ikeda & Ohme-Takagi, 2009).

The KRAB domain is a tetrapod vertebrate exclusive repression domain, which suggests its early evolution. It is a potent transcriptional repression domain that can be found in

the amino-terminal sequence of some zinc finger proteins, with KRAB-Zinc finger proteins being the largest family of transcription factors in the human genome (Mark, Åbrink, & Hellman, 1999). It is known to have protein-protein interaction, binding corepressor proteins or transcription factors to repress transcription (Urrutia, 2003). A key mechanism on the repression via KRAB seems to be the recruitment of heterochromatin protein 1 (HP1) and other chromatin modulating proteins to form heterochromatin, silencing the corresponding gene (Urrutia, 2003).

4. SUperNova Tag (SunTag)

In order to increase repression efficiency, it was reasoned that the recruitment of more than one repression domain could have an impact, since signals in many biological processes can be amplified by recruiting multiple copies of regulatory proteins to the target site. A protein scaffold, named SunTag, had been described and created for this purpose. It consists of a repeating peptide array that can recruit up to 24 copies of an antibody-fusion protein that bind to the repeating epitopes, with a 5 amino acid spacer between each epitope. At the time it was presented, it was also used to regulate gene activation via a dCas9 strategy by fusing the SunTag to dCas9 and making an antibody-activation domain (VP64) fusion protein (Tanenbaum, Gilbert, Qi, Weissman, & Vale, 2014) (Figure 5), so it was decided to adapt this strategy for negative transcriptional regulation as part of this study.

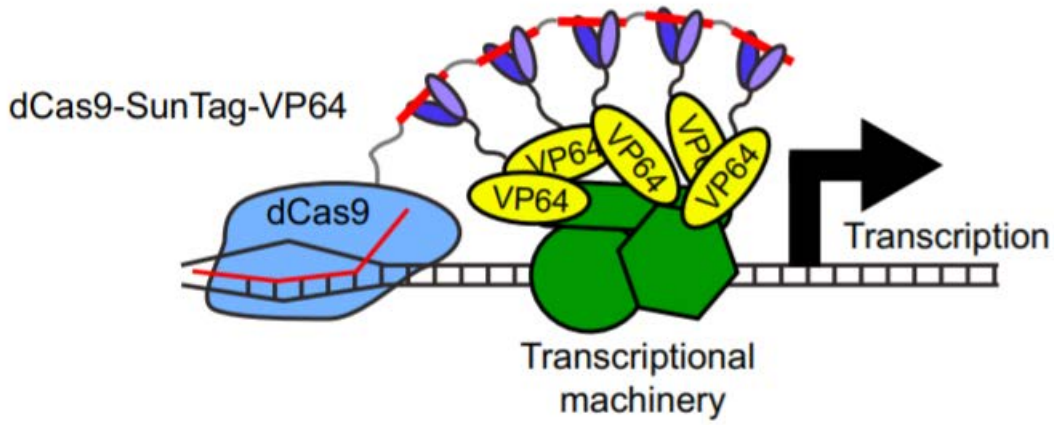

Figure 5. Transcriptional activation using the SunTag scaffold. The SunTag consists of a repeating peptide array that can recruit up to 24 antibodies. In this case, an antibody-VP64 (an activation domain) fusion protein was created and the scaffold was fused to dCas9 to bring as many activation domains as possible to the target gene. Image from (Tanenbaum et al., 2014).

This SunTag, however, proved to have some limitations due to the number of amino acids in the spacers (5 aa). Steric hindrances between antibodies when binding to the epitopes appeared due to the spacers being too short, which hindered the SunTag efficiency to recruit more copies of the antibody fusions. Another SunTag was developed to overcome this problem, using 22 amino acids as spacers instead of five (Papikian, Liu, Gallego-Bartolomé, & Jacobsen, 2019). Both SunTags were tested with an antibody-

repression domain fusion and compared to use for negative transcriptional regulation for this study.

5. *Nicotiana benthamiana*

The next step in this study was choosing an appropriate plant to conduct the assays. *Nicotiana benthamiana* was chosen for a number of reasons.

Nicotiana benthamiana is an angiosperm species native to Australia that is part of the *Solanaceae* family like tomato, peppers or tobacco, which makes it an interesting species for the study of agronomical traits. It is allotetraploid with a genome composed of 19 chromosomes. Five wild type accessions were classified and named according to the territory were they grow: Northern Territory (NT), North Western Australia (NWA), Western Australia (WA), Queensland (QLD) and South Australia (SA) (Bally et al., 2018). Originally, this plant was collected and used by virologists for the study of plant-pathogen interaction (Goodin, Zaitlin, Naidu, & Lommel, 2008). Evidence suggests that the original plant was plucked from the Northern Territory (NT). This original plant was used and grown systematically by self-propagation in every institution for research, which derived in a uniformity of *Nicotiana benthamiana* plants across the research community in a strain known as "*Nicotiana benthamiana* LAB strain" (Bally et al., 2018). This uniform strain allows for the comparison of results of different research groups and is the one used in this study.

Aside from its uniformity, *N. benthamiana* also has other qualities that make it a good model plant for a lot of different fields in plant research. It shares cell compartmentalization, cofactors and coenzymes with other plants, specially agronomically important plants from the *Solanaceae* family, which makes it relatively easy to transfer pathways from other plants without the need to extensively optimize the system (Reed & Osbourn, 2018). Although synthetic biology initially started in microorganisms, it soon spread to plants. Several plant species were suggested as the platform for engineering, one of them being *N. benthamiana* (Stewart, Patron, Hanson, & Jez, 2018).

Thus, *Nicotiana benthamiana* makes for a great species for this study due to all the characteristics listed above as well as its fast growth and the ease to grow it in a greenhouse.

6. Transient expression

Finally, choosing the correct assays to do was necessary. One of the main reasons to use *N. benthamiana* as a plant for synthetic biology studies is its susceptibility to infiltration by *Agrobacterium tumefaciens* and, specially, the use of agroinfiltration for transient expression. This involves infiltrating the leaves with a suspension of *A. tumefaciens* cells carrying the desired genes and allows for the rapid detection of the targeted protein or product. Agroinfiltration is a highly flexible process, allowing the simultaneous

expression of multiple genes by coinfiltrating different strains with different expression constructs (Reed & Osbourn, 2018). *N. benthamiana* also shows higher levels of expression for transient genes than most other model plants since it lacks some defensive capabilities related to salicylic acid (Canto, 2016).

Transient expression main constraints are that the results obtained through it are not always fully transferable to endogenous genes since levels of expression are not steady and depend on a lot of factors, including the quantity of *A. tumefaciens* infiltrated (Canto, 2016). While most of the T-DNA inserted via *A. tumefaciens* gets integrated in the genome (Figure 6), some of it may form complex structures on its own (Johansen & Carrington, 2001). The difference in structure for infiltrated genes may also alter the results observed. It is still, however, one of the most useful tools for initial tests since its quick and most of the results are transferable to a degree, and it was chosen for this reason.

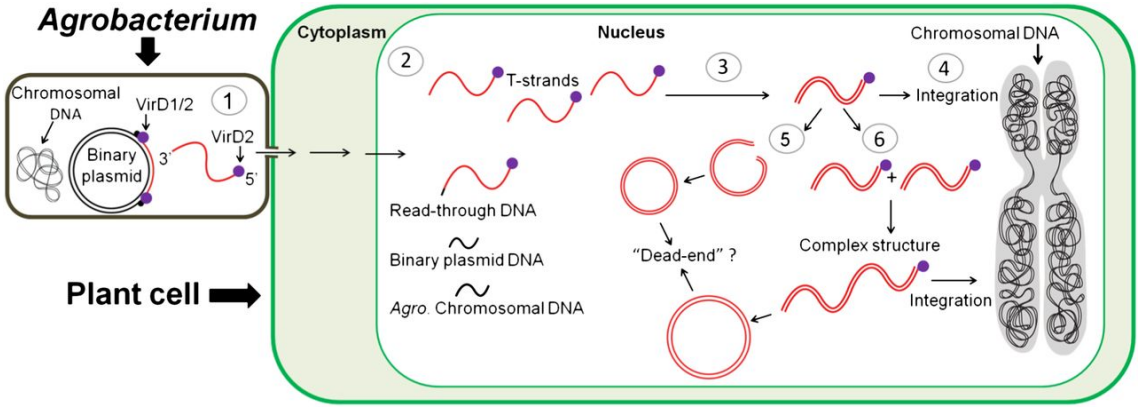

Figure 6. Proposed model for T-DNA integration. T-DNA may be integrated in the genome (4), form T-circles (5) or complex structures (6). Image from (Johansen & Carrington, 2001).

7. Background

A lot of studies have been made in both mutating and editing genes from different organisms using the CRISPR/Cas system. In plants, our laboratory designed a modular toolbox for gRNA-Cas9 genome engineering based on GoldenBraid (Vazquez-Vilar et al., 2016), making it easy to put together different tools using the CRISPR/Cas9 system for both editing and transcriptional regulation. Cas12a was also modularly adapted for use with the GoldenBraid system (Bernabé-Orts et al., 2019). All of the constructs used in this study were assembled via GoldenBraid.

dCas9 has also already been used in our laboratory to positively regulate translation both with a reporter and with endogenous genes, and the 5 aa SunTag was tested and reported decent results (Selma et al., 2019). Cas12a has also been assessed for gene editing (Bernabé-Orts et al., 2019). Gene activation using the 22 aa SunTag system was tested outside of our laboratory and made for an efficient system (Papikian et al., 2019).

Regarding gene repression with dCas9, results published by our laboratory suggest that guides upstream to the TATA-box but relatively close to the translation starting site (TSS) give the best results and that repression is stronger when using more than one guide (Vazquez-Vilar et al., 2016), although the repression efficiency obtained wasn't optimal. Unpublished results testing dCas12a for repression seem to indicate a stronger repression efficiency when using guides upstream the TATA-box but farther from the TSS than for dCas9. During the course of this study, the repression window for dCas9 was stablished between positions +25 and +75 from the TSS and was taken into account when testing for guide positioning with this endonuclease (Sanson et al., 2018).

Based on this background and the previous results mentioned on our laboratory, guide position, number of guides, different repression domains and two different systems were selected for testing and an experimental approach was designed.

Objectives

OBJECTIVES

1. **Determining optimal guide positioning (expressed as +/- bases from the TSS) for repression with dCas9 and dCas12a.** 5 guides will be designed and tested for each endonuclease for this purpose, both upstream and downstream from the TSS.
2. **Comparing repression efficiency between different repression domains (BRD, SRDX, KRAB domain).** All of the repression systems will be tested with these three domains.
3. **Comparing repression efficiency between both endonucleases (SpdCas9, LbdCas12a).** All of the repression systems will be tested for both endonucleases in similar conditions.
4. **Assessing if the number of guides used has an effect on repression efficiency.** After initial testing for guide positioning, working guides will be selected and tested in two-guide combinations.
5. **Testing and comparing both SunTag proteins (Named SunTag 5aa and SunTag 22aa for this study) and assess if repression efficiency increases when recruiting more repression domains through this system.** Both SunTag proteins will be tested for all of the conditions mentioned above and compared with each other and the single-domain system.

Materials and Methods

MATERIALS AND METHODS

1. GB phytobricks construction and assembly

1.1. Level 0 phytobricks

Level 0 GB phytobricks used in this work were created following the domestication strategy described in (Sarrion-Perdigones et al., 2013) (Vazquez-Vilar et al., 2017) following the GoldenBraid domestication tool ("Domestication," n.d.). The Domesticated phytobricks and other basic GoldenBraid phytobricks used are shown in Table 3.

Domestication of elements used was carried out by PCR following Phusion High-Fidelity DNA Polymerase (M0530) (New England Biolabs) manufacturer's protocol. For elements that were especially difficult to clone, a Touchdown (TD) protocol was used (Korbie & Mattick, 2008). Oligos used can be seen in Table 2.

Table 2. Primers used.

Primers used	
Name	**Sequence**
SunTag22aa Fw	GCGCCGTCTCGCTCGTTCGTATCCCTATGACGTGCCCGA
SunTag22aa Rv	GCGCCGTCTCGCTCAAAGCTTACCCTGAGCCTGATCCCC
sgRNA-35 Fw	GTGCAAGTGAATATGAGACTCTAAT
sgRNA-35 Rv	AAACATTAGAGTCTCATATTCACT
sgRNA+16 Fw	GTGCACGGGCCTTTCTTTATGTTTT
sgRNA+16 Rv	AAACAAAACATAAAGAAAGGCCCG
sgRNA+23 Fw	GTGCAGACGCCAAAAACATAAAGAA
sgRNA+23 Rv	AAACTTCTTTATGTTTTTGGCGTC
sgRNA+51 Fw	GTGCACCAGCGGTTCCATCTTCCAG
sgRNA+51 Rv	AAACCTGGAAGATGGAACCGCTGG
sgRNA+62 Fw	GTGCACCGCTGGAAGATGGAACCGC
sgRNA+62 Rv	AAACGCGGTTCCATCTTCCAGCGG
sgRNA+143 Fw	AGATACAGATGCACATATCGAGGT
sgRNA+143 Rv	GGCCACCTCGATATGTGCATCTGT
sgRNA+6 Fw	AGATTTTATGTTTTTGGCGTCTTC
sgRNA+6 Rv	GGCCGAAGACGCCAAAAACATAAA

PCR reactions were run in gel electrophoresis and purified directly or cut from gel depending on results following NucleoSpin Gel and PCR Clean-up manufacturer's protocol, then quantified using a NanoDrop ND-1000 Spectophotometer.

Table 3. Domesticated phytobricks and basic phytobricks used. Category syntax follows the standard syntax for plant synthetic biology (Patron et al., 2015).

GoldenBraid ID	Content	Category
Domesticated phytobricks		
GB2969	SunTag 22aa	B5
Other level 0 phytobricks used		
GB0030	CaMV 35S promoter	A1-A2-A3-B1-B2
GB1001	Arabidopsis thaliana U626 promoter	A1-A2-A3-B1-B2
GB0037	Nopaline synthase terminator (Tnos)	B6-C1
GB1443	U626:LbCas12aDR	None
GB1444	Hammerhead Delta Virus ribozyme (HDV)	None
GB1079	SpdCas9	B3-B4
GB1662	LbdCas12a	B3-B4
GB1463	ScFv single chain antibody	B3-B4
GB1175	BRD repression domain	B5
GB1185	SRDX repression domain	B5
GB1452	KRAB repression domain	B5
GB2514	SunTag 5aa	B5

1.2. Cas9 guide design and assembly

Guide RNA protospacers for Cas9 were manually designed on Benchling ("Cloud-Based Informatics Platform for Life Sciences R&D | Benchling," n.d.) using NGG as the PAM sequence. Editing on-target and off-target scores were not taken into account.

For the assembly of gRNAs to be used in the multiplexing strategy, GB level -1 plasmids (GB2245) containing the tRNA were designed following the plasmid structure described in(Vazquez-Vilar et al., 2016) (Vazquez-Vilar et al., 2016). Individual gRNAs were assembled in pUPD2 with a BsmBI restriction-ligation reaction that was performed with 75 ng of pUPD2 , 75 ng of the level -1 (GB2245) and a mix of complementary primers with the protospacer sequence.

Finally, they were assembled in an pDGB3_alpha2 vector with an AtU626 promoter (GB1001) and a nopaline synthase terminator (Tnos) (GB0037) with a BsaI restriction-ligation. The assembly reactions were as follows (Figure 7).

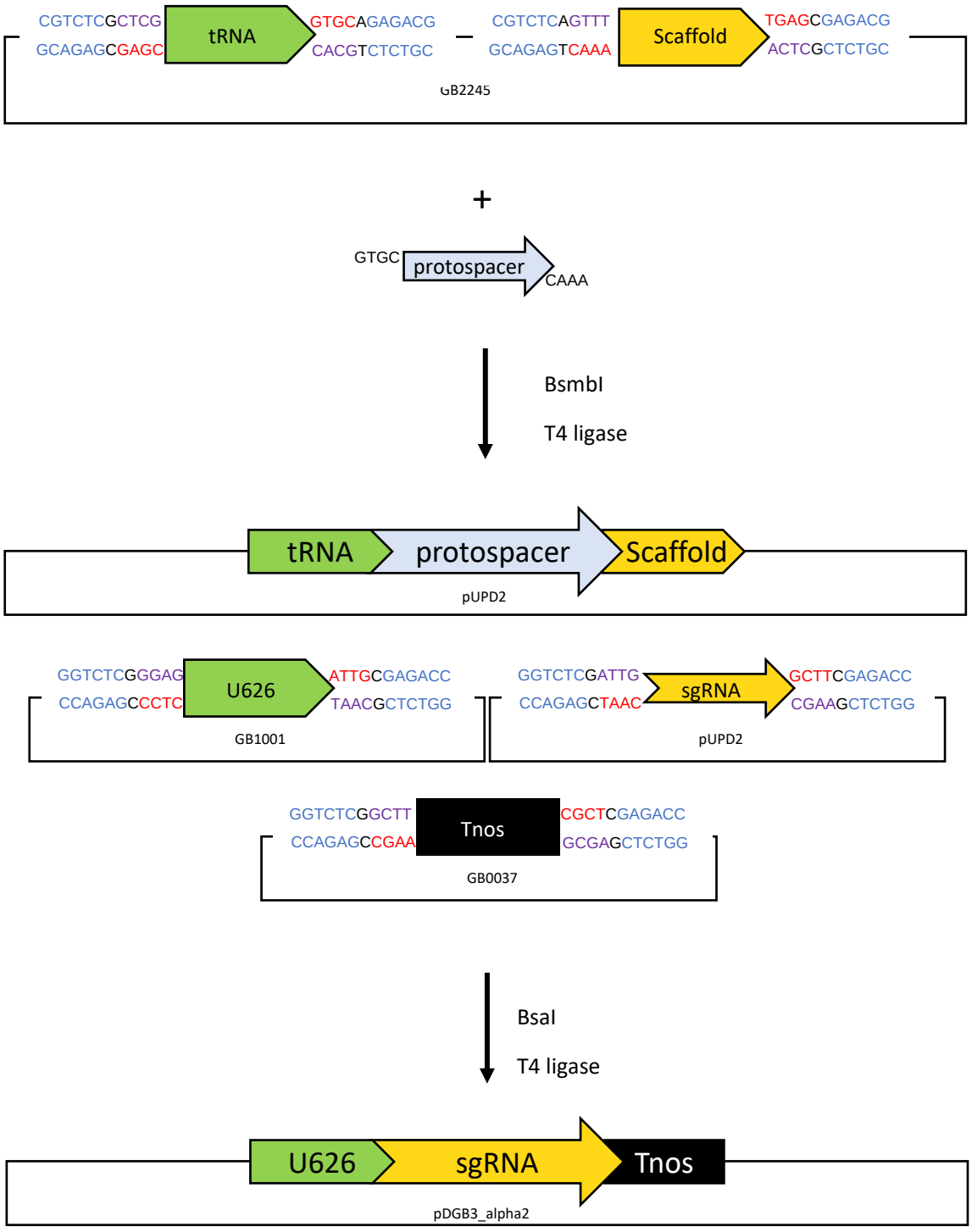

Figure 7. Cas9 guide assembly.

1.3. Cas12a guide design and assembly

Guide RNA protospacers for Cpf1 were manually designed on Benchling ("Cloud-Based Informatics Platform for Life Sciences R&D | Benchling," n.d.) using TTTV as the PAM sequence, where V can be any nucleotide aside from T. Editing on-target and off-target

scores were not taken into account. Primers were manually designed. LbCas12a sgRNAs primers were resuspended in water to final concentrations of 10 µM. Equal volumes of forward and reverse primers for each gRNA were mixed. The mixture was incubated at room temperature for 5 min for the hybridization of the primer pair. gRNA assembly in level 1 was carried out with a BsaI restriction–ligation reaction. The reactions were set up in 10 µl with 1 µl of primers mix ,75 ng of AtU626 adapted for Cas12a (GB1443), 75 ng of the Hammerhead Virus Ribozime (HDV) (GB1444) for the auto processing of the guide and 75 ng of pDGB3α destination vector (Figure 8).

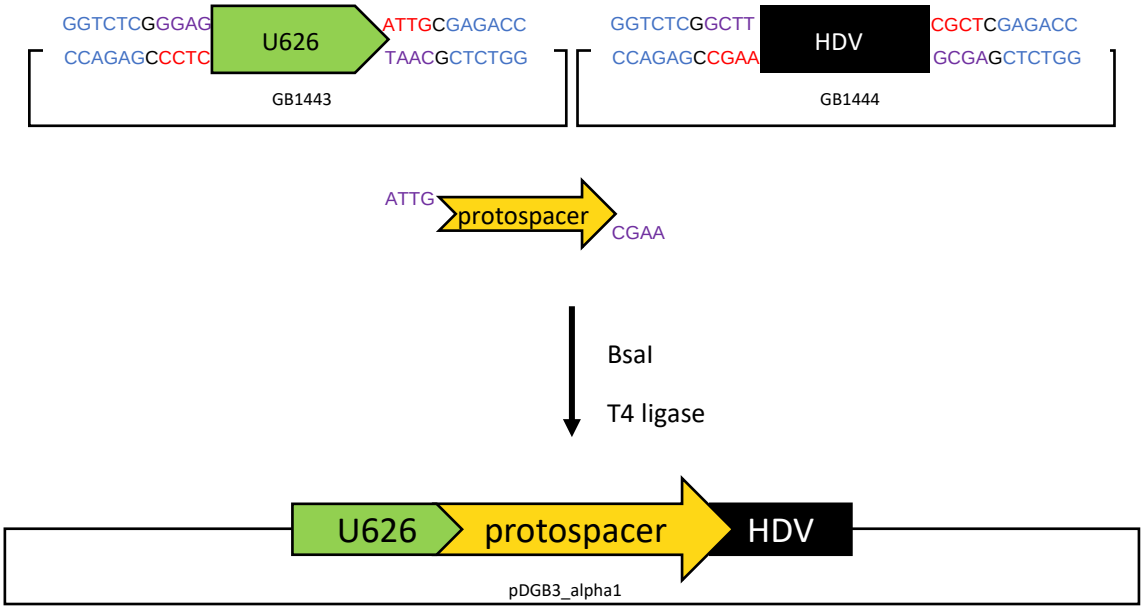

Figure 8. Cas12a guide assembly.

1.4. Transcriptional Units assembly (level ≥1)

Multipartite BsaI restriction–ligation reactions from level 0 parts and binary BsaI or BsmBI restriction–ligation reactions were performed as described in (Sarrion-Perdigones et al., 2013) to obtain all the level ≥1 assemblies.

Level 1 transcriptional Units were assembled on pDGB3_alpha1 or pDGB3_alpha2 vectors under regulation by CaMV 35S promoter (GB0030) with a tNos terminator (GB0037) following a multipartite reaction via GoldenBraid. Reactions were as follows (Figure 9).

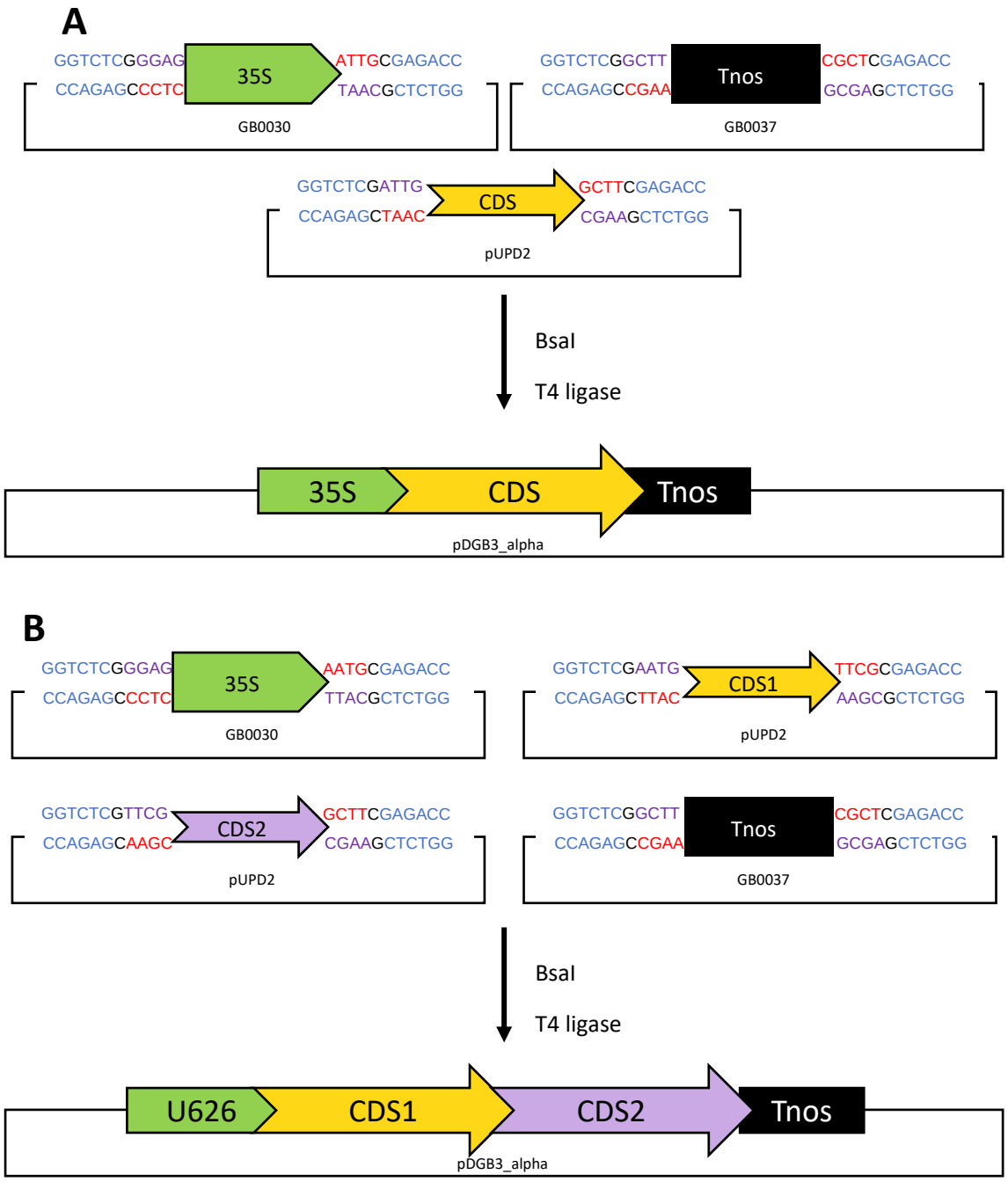

Figure 9. Multipartite assembly for transcriptional units with A) one CDS or B) more than one CDS. Transcriptional units assembled on either pDGB3_alpha1 or pDGB3_alpha2 vectors regulated by CaMV 35S promoter and nopaline synthase terminator (Tnos). CDS used for these constructs had been previously domesticated with nomenclature B3-B4 (CDS1) or B5 (CDS2) following the standard syntax for plant synthetic biology (Patron et al., 2015).

Every final construct assembled and other constructs used can be seen in Table 4.

Table 4. Constructs assembled and other constructs used.

GoldenBraid ID	Construct
Assembled constructs	
GB2642	pUPD2 Cas9 LUC sgRNA+16
GB2643	pUPD2 Cas9 LUC sgRNA+23
GB2644	pUPD2 Cas9 Pnos sgRNA-35
GB2645	pUPD2 Cas9 LUC sgRNA+51
GB2646	pUPD2 Cas9 Luc sgRNA+62
GB2647	pDGB3_alpha2 U626:sgRNA+16:Tnos
GB2648	pDGB3_alpha2 U626:sgRNA+23:Tnos
GB2649	pDGB3_alpha2 U626:sgRNA-35:Tnos
GB2778	pDGB3_alpha2 U626:sgRNA+51:Tnos
GB2779	pDGB3_alpha2 U626:sgRNA+62:Tnos
GB2976	pDGB3_alpha1 U626:sgRNA+6:HDV
GB2977	pDGB3_alpha1 U626:sgRNA+143:HDV
GB2970	pDGB3_alpha1 35S:dCas9-Suntag22aa:Tnos
GB2971	pDGB3_alpha1 35S:dCas12a-Suntag5aa:Tnos
GB2972	pDGB3_alpha1 35S:dCas12a-Suntag22aa:Tnos
GB2973	pDGB3_alpha2 35S:ScFv-BRD:Tnos
GB2874	pDGB3_alpha2 35S:ScFv-SRDX:Tnos
GB2875	pDGB3_alpha2 35S:ScFv-KRAB:Tnos
Other constructs used	
GB1812	pDGB3_alpha2 U626:sgRNA-165:HDV
GB1808	pDGB3_alpha2 U626:sgRNA-66:HDV
GB1807	pDGB3_alpha2 U626:sgRNA-9:HDV
GB1172	pDGB3_alpha2 35S:dCas9-BRD:Tnos
GB1188	pDGB3_alpha2 35S:dCas9-SRDX:Tnos
GB1459	pDGB3_alpha2 35S:dCas9-KRAB:Tnos
GB1668	pDGB3_alpha1 35S:dLbCas12a-BRD:T35S
GB1669	pDGB3_alpha1 35S:dLbCas12a-SRDX:T35S
GB1873	pDGB3_alpha1 35S:dLbCas12a-KRAB:T35S
GB1603	pDGB3_alpha2 35S:dCas9-Suntag5aa:Tnos7

Every construct was checked by digestion with the corresponding enzyme, gel electrophoresis and sequencing.

2. Cloning

2.1. Strains and growth conditions

Escherichia coli TOP10 was used for gene cloning and *Agrobacterium tumefaciens* C58 was used for plant agroinfiltration. For liquid culture, both strains were grown in LB medium with the corresponding antibiotic under agitation 120rpm at 37ºC and 28ºC, respectively.

2.2. *E. coli* transformation

For chemocompetent *E. coli* TOP10 transformation, a 100 µl aliquot was thawed on ice and 2µl plasmid was added and incubated on ice for 5 minutes. 300 µl SOC medium (2% tryptone, 0,5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) was added and the cells were incubated at 37ºC for 1 hour under agitation 120rpm, then plated on LB medium with the corresponding antibiotics, X-Gal and IPTG.

2.3. *A. tumefaciens* transformation

For electrocompetent *A. tumefaciens* C58 transformation, a 100 µl aliquot was thawed on ice and 1 µl plasmid was added. The mix was transferred to an electroporation cuvette and a 1440V pulse was applied. 500 µl LB was added to the cells and the mix was transferred to a 1.5 ml tube, then incubated at 28ºC for 2 hours under agitation 120rpm. 50 µl were plated on LB medium with the corresponding antibiotics.

2.4. Plasmid extraction

E. coli plasmids were extracted following E.Z.N.A Plasmid Mini Kit I (Omega Bio-Tek) manufacturer's protocol.

Agrobacterium tumefaciens plasmids were extracted following QIAprep Spin Miniprep Kit (250) manufacturer's protocol.

Plasmid DNA concentration was quantified using a NanoDrop ND-1000 Spectophotometer.

2.5. Gel electrophoresis

Gel electrophoresis was performed on 1% agarose 50 mL or 200 mL gels with 0.5 µl BrEt per 50 mL gel, and run through with a 100V or 120V current, respectively.

2.6. Sequencing

Plasmid inserts were sequenced by the IBMCP sequencing unit.

3. Plant material

Wild type *Nicotiana benthamimana* lab strain were grown for 5 to 6 weeks in a growing chamber in a 16h light (24ºC) and 8h dark (21ºC) photoperiod. They were watered with Hoagland solution.

3.1. *Nicotiana benthamiana* transient expression

Transient expression assays were carried out through agroinfiltration of *N. benthamiana* leaves. 5mL Fresh overnight *Agrobacterium* cultures were pelleted by centrifugation at 4500rpm for 10 minutes, then resuspended in 10mL agroinfiltration solution (10mM MES, pH 5.6, 10 mM $MgCl_2$, 200 µM acetosyringone) and incubated for 2h at room temperature on a horizontal rolling mixer in the dark. Optical density was measured for each culture at 600nm (OD600) and they were mixed for experiments in which more than one construct was used for a final optical density of 1. Agroinfiltrations were carried out through the abaxial surface of three out of the four youngest leaves of each plant with a 1ml needle-free syringe. A silencing suppressor (P19) was used for all assays.

4. Luciferase/Renilla assays

The assay conditions follow the experimental standards found in ("Add experiment," n.d.) with minor modifications. Samples were collected at 5 days post infiltration (5dpi) instead of 4 dpi. For determination of the Luc/Ren activity, one disc per leaf (~6,43 mg) was excised 5 days post inoculation and kept in 2ml Eppendorf tubes frozen in liquid nitrogen. Leaf discs were homogenized and extracted with 180 µl "Passive Lysis Buffer", then centrifuged for 10 minutes (13000 rpm) at 4ºC. Supernatant was used as the working plant extract. Luc/Ren activities were determined following the Dual-Glo Luciferase Assay System (Promega) manufacturer's protocol with minor modifications: 10 µ working plant extract, 40 µl LARII and 40 µl Stop&Glo Reagent were used. Measurements were made using a GloMax 96 Microplate Luminometer (Promega) with a 2-s delay and 10-s measurement time. Data was normalized to Pnos:luc and a 35S control was used following standard measurement protocols described in (Vazquez-Vilar et al., 2017).

Results and Discussion

RESULTS AND DISCUSSION

1. *N. benthamiana* transcriptional regulation

Comparisons among different repression strategies were performed transiently in *N. benthamiana* leaves. Repression levels using the different nucleases were assessed using Nopaline synthase promoter (pNos) coupled to firefly luciferase (Fluc) reporter. A constitutive Renilla luciferase (Rluc) was used as internal reference.

1.1. dCas9 gRNAs design

gRNA position was designed based on previous results obtained by our laboratory. The transcriptional repression achieved using different gRNAs that target upstream of the TATA box of the promoter had already been tested and did not show good results. For that reason, the next step was to design gRNAs downstream from the promoter or inside the promoter but closer to the transcription starting site (TSS), testing different distances from the transcription starting site (TSS) and both DNA strands (+16/+23, +51/+62 distance from the TSS). One single gRNA was designed for a position inside the promoter (-35) that was closer to the end of the promoter than other gRNAs previously tested, downstream from the TATA box and overlapping slightly with it. Designed gRNAs were assembled with a multiplexing strategy (see Materials and Methods). The final constructs were as follows (Figure 10).

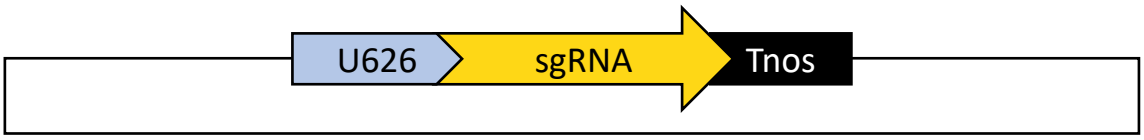

Figure 10. dCas9 sgRNA final construct. Each sgRNA consists of a pre-tRNA, the designed gRNA and a scaffold for the autoprocessing of the gRNA, and is under regulation of the *At*U626 promoter and the Tnos. dCas9 gRNAs were assembled an alpha2 vector using a GoldenBraid strategy.

1.2. dCas12a gRNA design

gRNA position was selected based on dCas9 results, previous studies by our laboratory and the theoretical repression window described by (Sanson et al., 2018). The closest PAMs to this theoretical window were selected to design gRNAs at position +6 and +143 from the TSS. Based on previous tests made in our laboratory, three gRNAs upstream of the TSS of the promoter were also tested (-165, -66, -9).

gRNAs were assembled under an *At*U626 promoter and with HDV (Hammerhead Delta Virus ribozyme) for gRNA autoprocessing, as is shown in Figure 11.

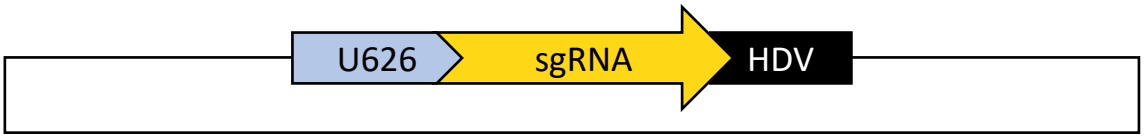

Figure 11 dCas12a sgRNA final construct. Each designed gRNA was assembled with *At*U626 promoter adapted for dCas12a and HDV (Hammerhead Delta Virus ribozyme) for autoprocessing of the gRNA. dCas12a gRNAs were assembled on an alpha2 vector using a GoldenBraid strategy.

2. Single domain repression strategy

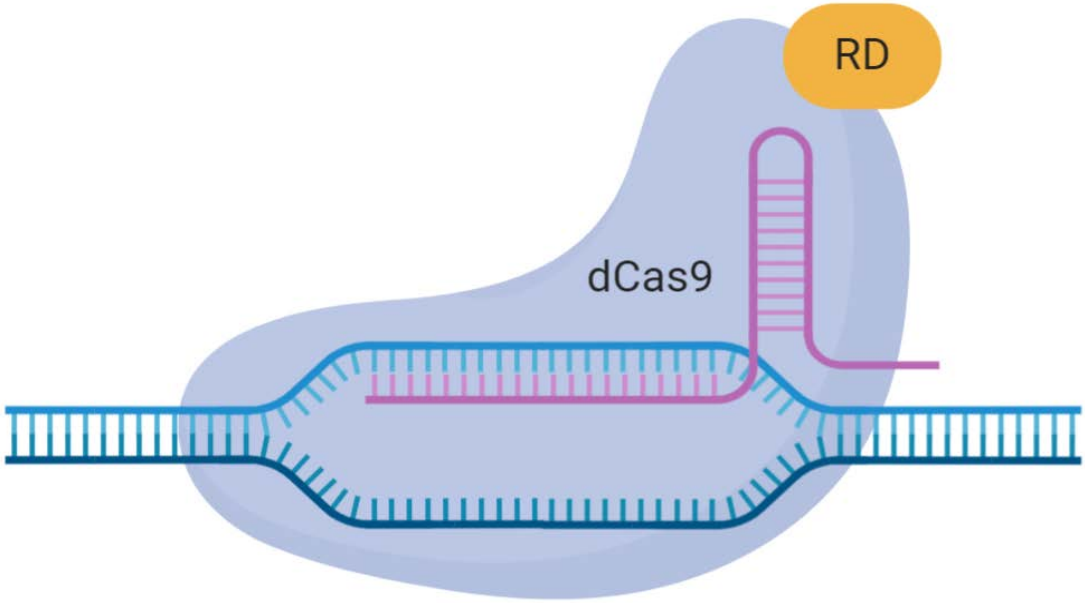

Figure 12. dCas9 repression strategy. dCas9 fused to a repression domain (RD). In blue, the targeted gene; in purple, the guide RNA. The same strategy was used for dCas12a. Image made with BioRender.

Initial tests were carried out by fusing a single repression domain to dCas9 and dCas12a. Different repression domains (BRD, SRDX, KRAB) were fused to dCas9 and dCas12a via GoldenBraid assembly under regulation by CaMV35S promoter and Tnos terminator in a pDGB3_alpha2 or pDGB_alpha1 vector using BsaI as a restriction enzyme. gRNAs were designed for different positions in the promoter or the luciferase gene to bring dCas9/dCas12a with the fused repression domain to the targeted site (Figure 12).

2.1. Single gRNA repression

2.1.1. dSpCas9 single gRNA repression

The direct fusion of RD (Repressor Domain) to dCas9 was initially tested. The domains selected for this assay were BRD (B3 repression domain) and SRDX (EAR repression domain) in order to compare the repression achieved with both domains and test the optimal gRNA positioning. Five gRNAs were designed and assayed separately to target

different positions of either pNOS or the luciferase gene (Figure 13). Significant repression was observed for position -35 inside the pNOS promoter for both repression domains. dCas9:BRD showed 30% repression (Figure 13A) while dCas9:SRDX showed 35% (Figure 13B). Only position +51 with the dCas9:SRDX domain (Figure 13B) showed significant repression inside the luciferase gene with 38% repression.

Both positions were selected for further assays, as well as position +62, which showed signs of repression although none of the previous assays were statistically significant.

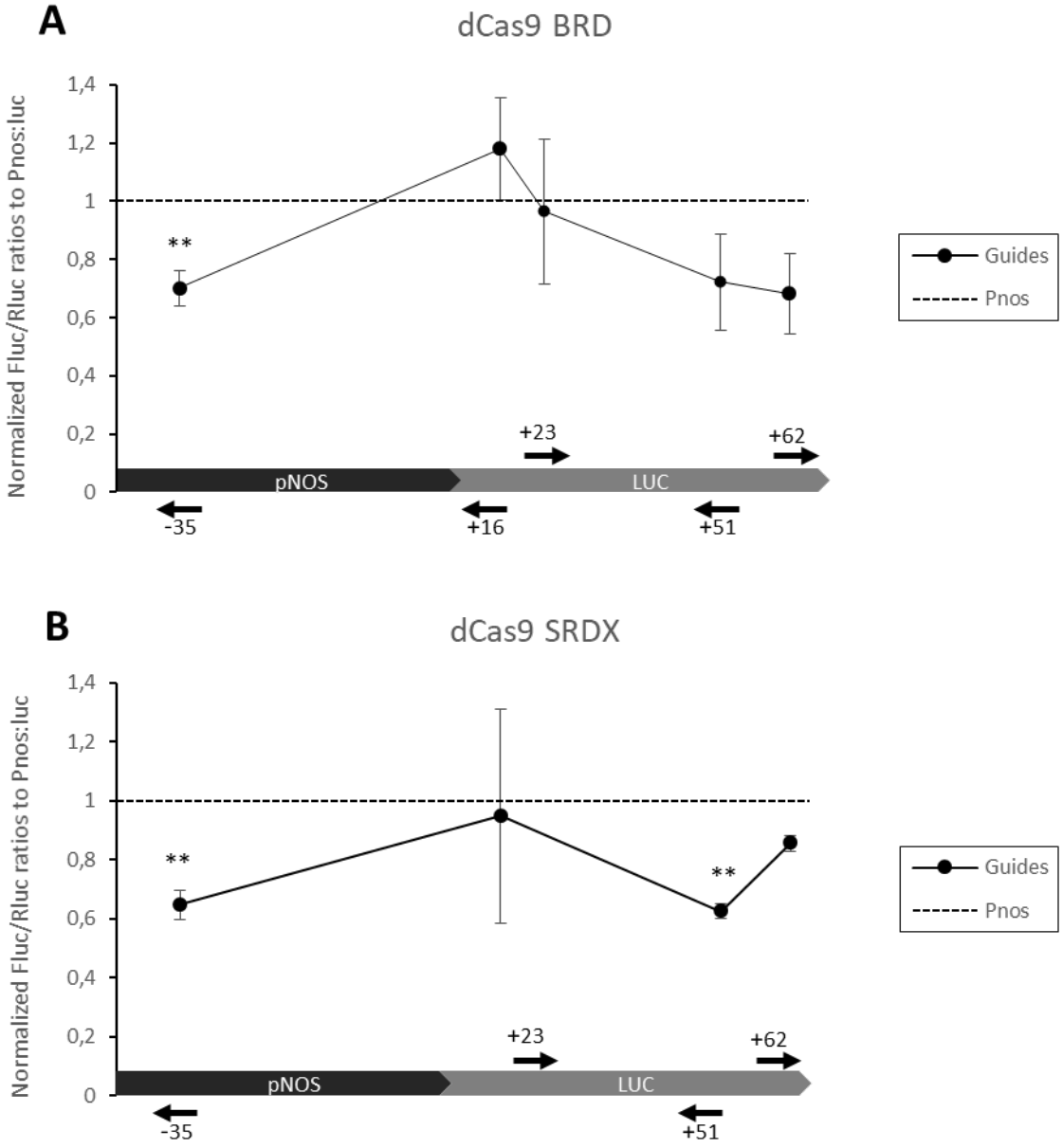

Figure 13. Transcriptional repression achieved for differente gRNA positions with a single domain strategy for BRD (A) and SRDX (B) domains. Normalized Luciferase/Renilla values to pNos:luc (represented as a dotted line) control are shown. An internal 35S:luc control was used. gRNA position and DNA strand is shown under the graph. Asterisks indicate Student's t-test significant values (one asterisk for $p < 0,05$; two asterisks for $p < 0,01$; three asterisks for $p < 0,001$).

Selected gRNAs were then tested with a different repression domain (KRAB) to test the repression efficiency of this domain (Figure 14).

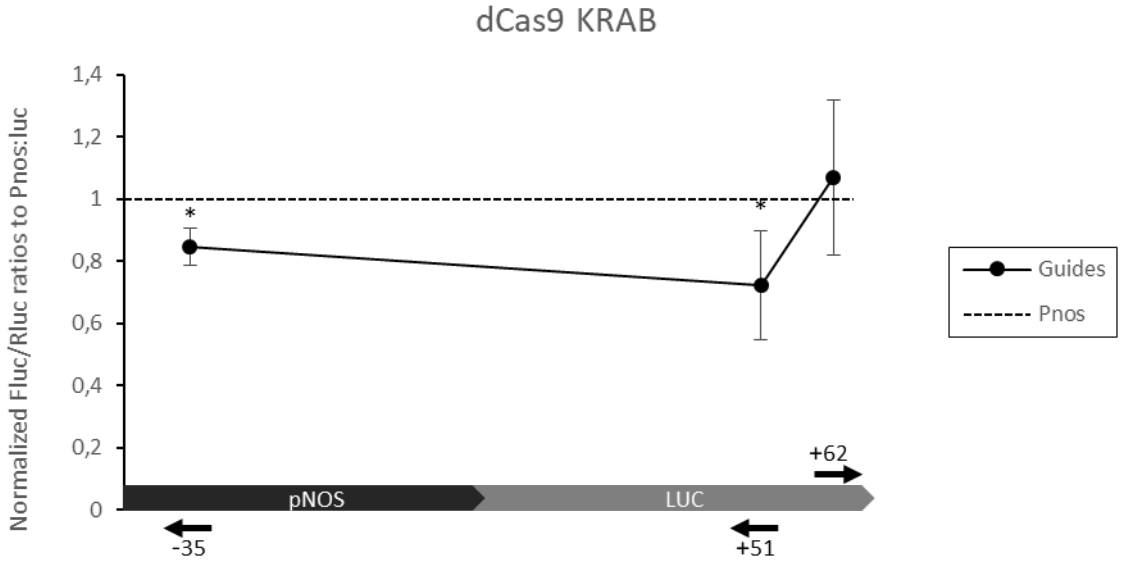

Figure 14. Transcriptional repression achieved for each dCas9 gRNA tested with a single repression domain for KRAB domain. Normalized Luciferase/Renilla values to pNos:luc (represented as a dotted line) control are shown. An internal 35S:luc control was used. gRNA position and DNA strand is shown under the graph. Asterisks indicate Student's t-test significant values (one asterisk for $p < 0,05$; two asterisks for $p < 0,01$; three asterisks for $p < 0,001$).

Both positions -35 and +51 showed significant repression for dCas9:KRAB domain. However, values were considerately lower in both cases compared to dCas9:BRD and dCas9:SRDX domains, with only 15% repression for position -35 and 28% repression for position +51. Due to these much lower values, KRAB domain was discarded from further assays with dCas9.

These results indicate higher repression efficiency for gRNAs inside de promoter and close to position +50 of the gene. During the course of this study, evidence was published indicating Cas9 repression window to be between bases +25 and +75 (Sanson et al., 2018). Both working gRNAs tested inside the luciferase gene are inside of that window.

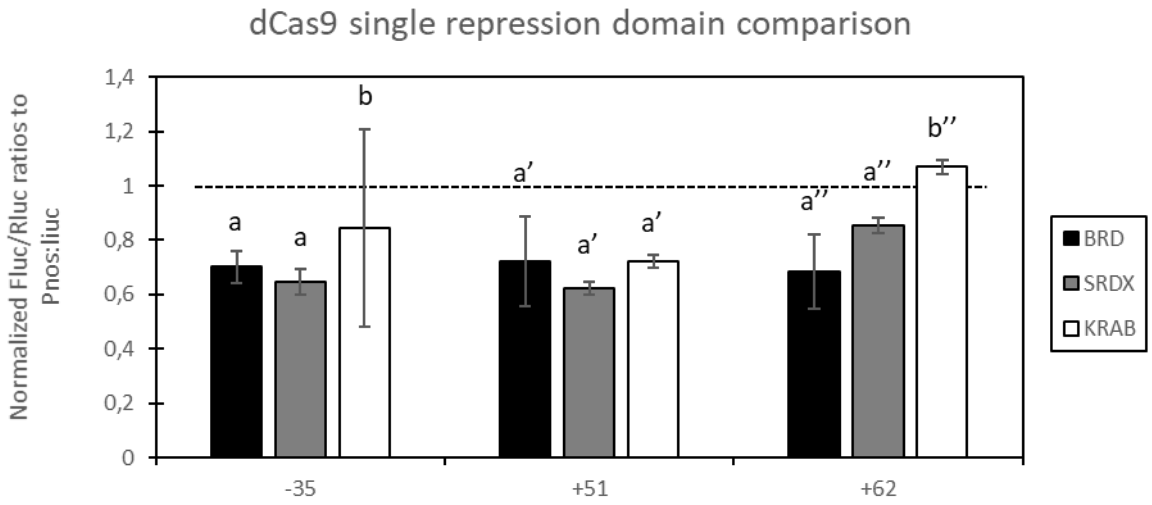

Figure 15. Comparison of the transcriptional repression achieved for each repression domain with dCas9. Only the best working gRNAs are shown. Normalized Luciferase/Renilla values to pNos:luc (represented as a dotted line) control are shown. An internal 35S:luc control was used. gRNA position and DNA strand is shown under the graph. Asterisks indicate Student's t-test significant values (one asterisk for $p < 0,05$; two asterisks for $p < 0,01$; three asterisks for $p < 0,001$).

Comparison between the three repression domains used (Figure 15) show similar repression levels for both dCas9:BRD and dCas9:SRDX domains and lower repression efficiency for dCas9:KRAB domain.

2.1.2. dLbCas12a single gRNA repression

Following the same experimental structure, gRNAs designed for dCas12a were tested in *N. benthamiana* leaves. gRNAs that target upstream from the TSS show significant repression with both BRD and SRDX domains (Figure 16), while gRNAs downstream from the TSS are not able to repress. Repression levels achieved with dCas12 were higher than with dCas9, getting as far as 80% with dCas12a:BRD domain and gRNA position -165. A pattern might be deduced from all three domains, with repression starting at the end of the promoter and peaking close to position -66. dCas12a:KRAB domain shows too much variability between samples, making it unreliable for this tool.

Due to the absence of repression for gRNAs +6 and +143 with dCas12:SRDX and lower levels with dCas12:BRD, only gRNAs upstream from the TSS were selected for further testing.

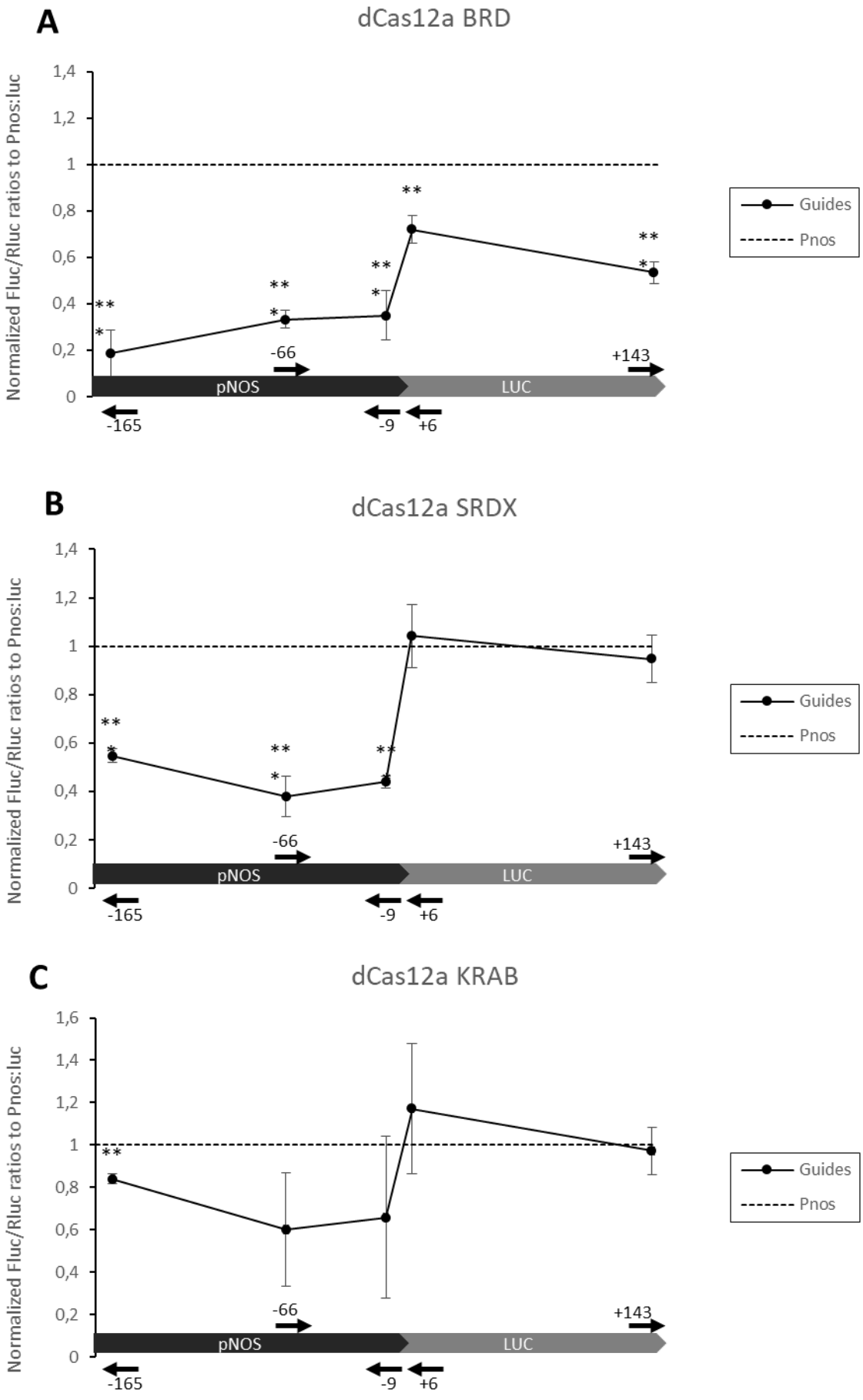
A
dCas12a BRD
Normalized Fluc/Rluc ratios to Pnos:luc
1,4
1,2
1
0,8
0,6
0,4
0,2
0
Guides
Pnos
-66
+143
pNOS
LUC
-165
-9
+6
B
dCas12a SRDX
Normalized Fluc/Rluc ratios to Pnos:luc
Guides
Pnos
-66
+143
pNOS
LUC
-165
-9
+6
C
dCas12a KRAB
Normalized Fluc/Rluc ratios to Pnos:luc
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0
Guides
Pnos
-66
+143
pNOS
LUC
-165
-9
+6

Figure 16. Transcriptional repression achieved for each dCas12a gRNA tested with a single repression domain for BRD (A), SRDX (B) and KRAB (C). Normalized Luciferase/Renilla values to pNos:luc (represented as a dotted line) control are shown. An internal 35S:luc control was used. gRNA position and DNA strand is shown under the graph. Asterisks indicate Student's t-test significant values (one asterisk for $p < 0,05$; two asterisks for $p < 0,01$; three asterisks for $p < 0,001$).

Comparison of the repression achieved for each gRNA with the three repression domains tested shows significant repression levels for dCas12a:BRD and dCas12a:SRDX and lower repression levels for dCas12a:KRAB (Figure 17). dCas12a:KRAB also shows greater variability between samples than dCas12a:BRD and dCas12a:SRDX. It was discarded from further assays for these reasons.

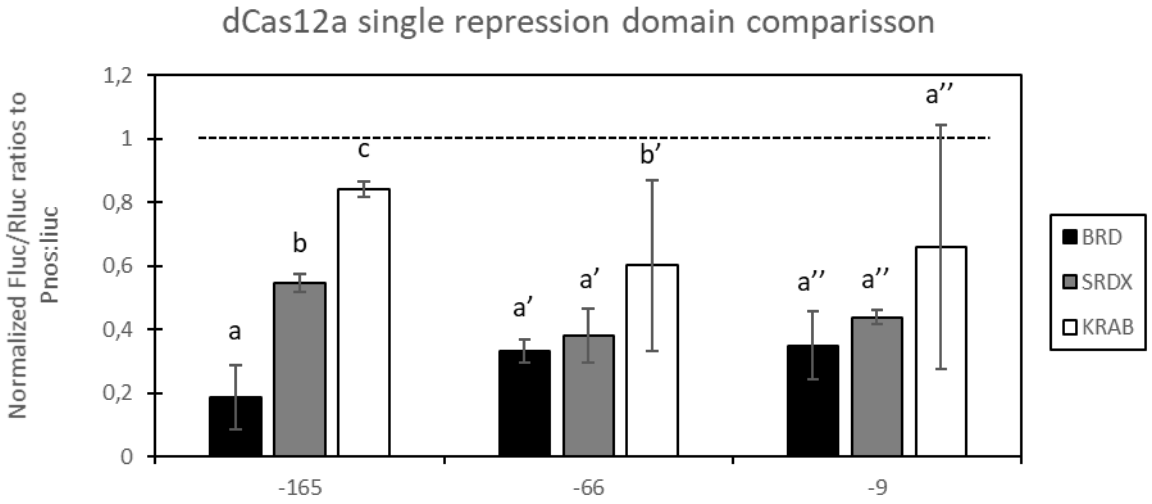

Figure 17. Comparison of the transcriptional repression achieved for each repression domain with dCas12a. Only the best working gRNAs are shown. Normalized Luciferase/Renilla values to pNos:luc (represented as a dotted line) control are shown. An internal 35S:luc control was used. gRNA position and DNA strand is shown under the graph. Asterisks indicate Student's t-test significant values (one asterisk for $p < 0,05$; two asterisks for $p < 0,01$; three asterisks for $p < 0,001$).

2.2. Double gRNA repression

In a further optimization step, the combination of two gRNAs was tested with both nucleases. The best gRNAs for each endonuclease were selected (positions -35, +51 and+62 for dCas9; positions -165, -66 and-9 for dCas12a) and tested in pairs to test if increasing the number of gRNAs could enhance repression efficiency. KRAB domain was discarded from these assays due to the low repression achieved previously.

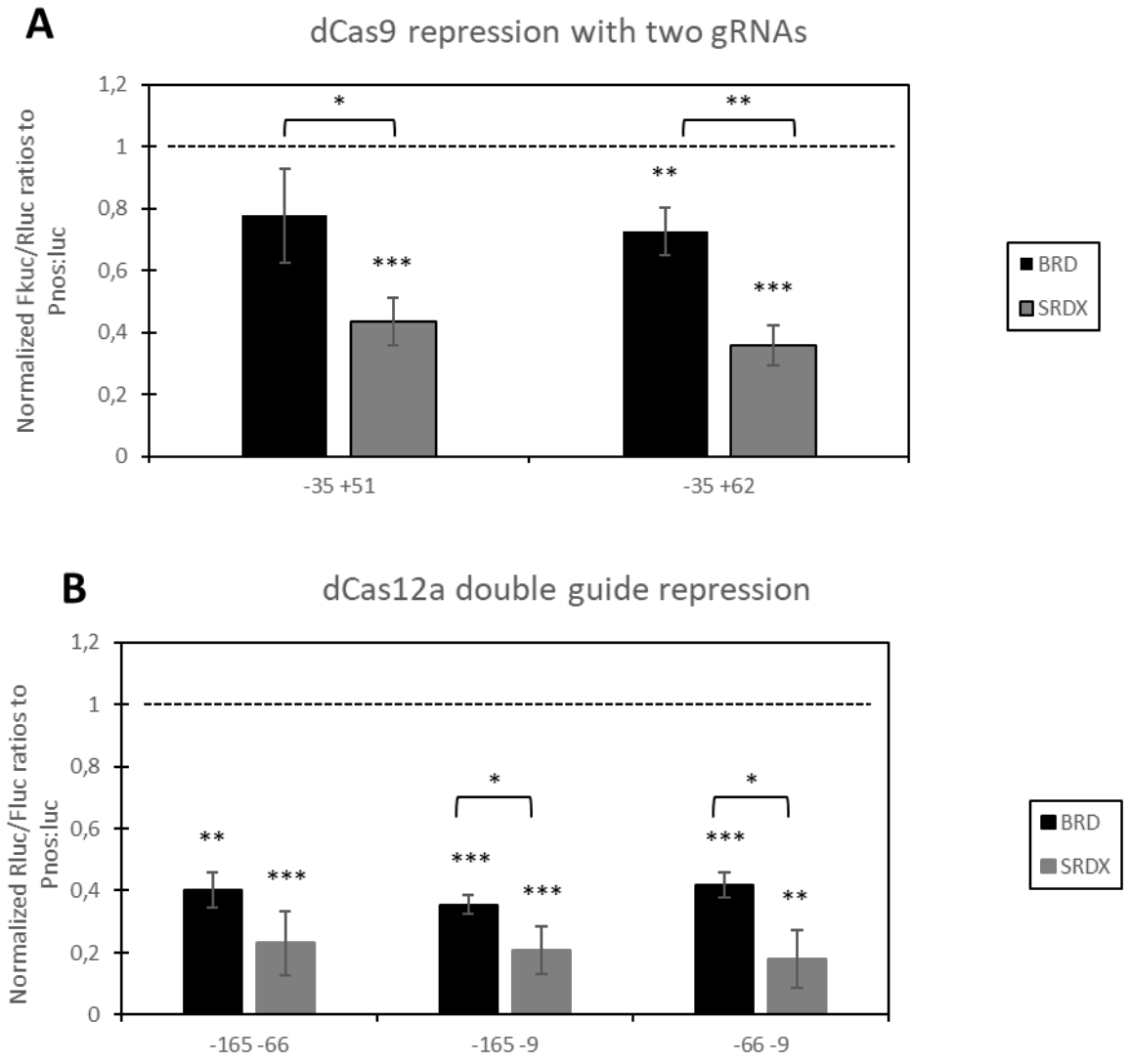

Figure 18. Transcriptional repression achieved using two gRNAs with a single domain strategy for dCas9 (A) and dCas12a (B). Normalized Luciferase/Renilla values to pNos:luc (represented as a dotted line) control are shown. An internal 35S:luc control was used. Asterisks indicate Student's t-test significant values (one asterisk for $p < 0,05$; two asterisks for $p < 0,01$; three asterisks for $p < 0,001$).

Significant repression was observed for gRNA combination -35/+51 for dCas9:SRDX but not for dCas9:BRD (Figure 18A). Both domains showed significant repression with gRNA combination -35/+62. Using this pair, dCas9:BRD repressed 28% of the luciferase activity, which is the same or less than the repression observed using each gRNA separately (Figure 18A). On the other hand, the repression levels achieved with dCas9:SRDX are 57% using the pair -35/+51 and 64% for -35/+62.

This data shows that dCas9:SRDX domain is more efficient than dCas9:BRD domain for repression with direct fusion to dCas9 using a pair of gRNAs. In parallel, every pair of gRNAs with dCas12a (Figure 18B), showed significant repression with both domains.

Repression levels achieved with dCas12a were higher than the repression levels obtained with dCas9, reaching levels of approximately 80% for all combinations with dCas12a:SRDX and 60% with dCas12a:BRD. dCas12a:SRDX repression levels were higher than dCas12a:BRD for combinations -165/-9 and -66/-9. This data agrees with the results obtained with dCas9, which show that the SRDX domain has more repression capability than BRD domain using more gRNAs.

3. SunTag strategy

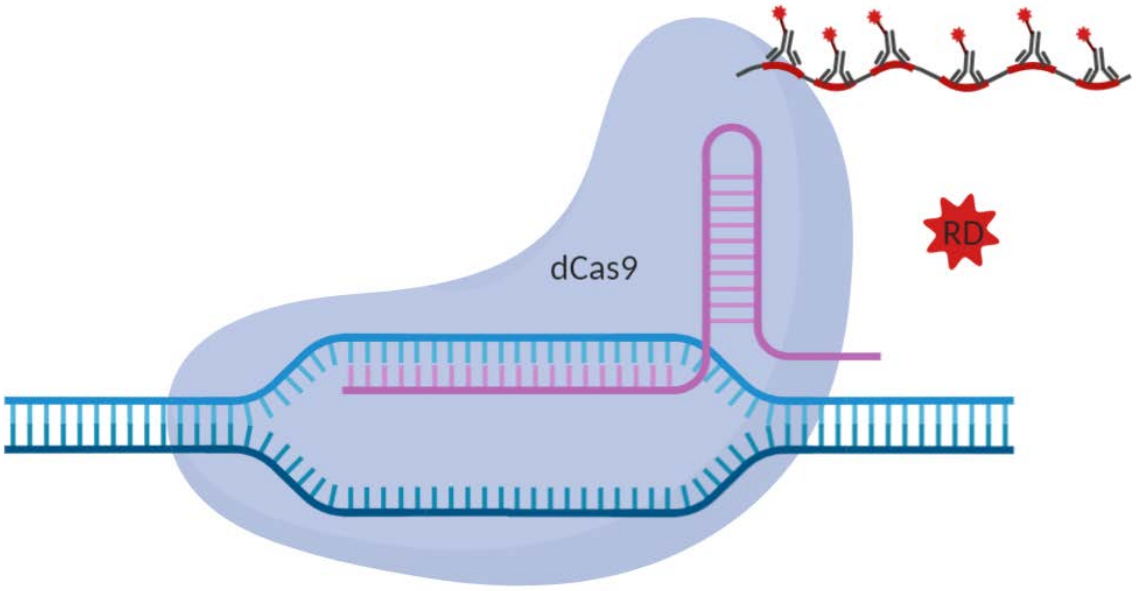

Figure 19. SunTag repression strategy. dCas9 fused to the SunTag protein, which recruits ScFv antibodies fused to different repression domains (RD). In blue, the targeted gene; in purple, the guide RNA. The same strategy was used for dCas12a. Image made with BioRender.

In order to optimize the transcriptional repression achieved, other strategies with dCas9 and dCas12a were explored. It was reasoned that using more than one repression domain could be a way to increase repression efficiency. To this end, a second strategy was devised which consisted on the use of a repetitive peptide array named SunTag, with the ability to recruit up to 24 antibodies that fuse to its epitopes (Tanenbaum et al., 2014). This SunTag is referred in this study as SunTag (5aa) due to having 5 amino acids as spacers in between epitopes.

To this end, dCas9 and dCas12a were bound to the SunTag (5aa) via GoldenBraid under regulation by a 35S promoter and a Tnos terminator (dCas9:SunTag5aa, dCas12a:SunTag5aa). The antibody ScFv was then fused to the corresponding repression domain (ScFv:BRD, ScFv:SRDX) via GoldenBraid under regulation by a 35S promoter and a Tnos.

This strategy allows the recruitment of numerous repression domains by a single dCas:SunTag5aa, (Figure 19) potentially increasing the repression efficiency.

3.1. SunTag 5aa repression

3.1.1. Single gRNA repression

Following the same conditions of previous assays, the gRNAs that showed the best repression levels with the direct fusion of RD were tested with the SunTag strategy (positions -35, +51 and +62 for dCas9; positions -165, -66 and -9 for dCas12a). KRAB domain was discarded due to inefficient results in previous assays. Results are shown on Figure 20.

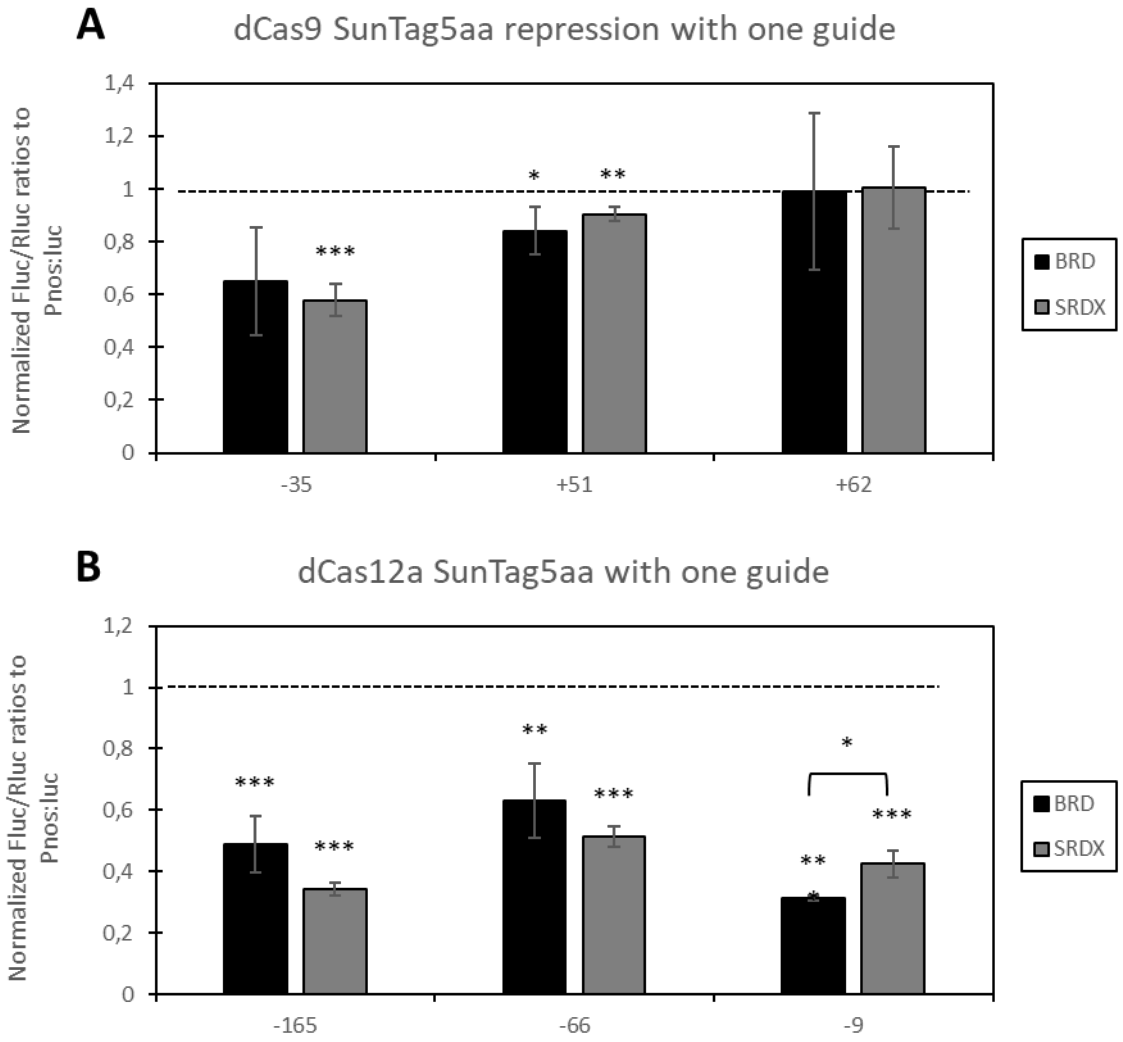

Figure 20. Transcriptional repression achieved using one gRNA with a SunTag 5aa strategy for dCas9 (A) and dCas12a (B). Normalized Luciferase/Renilla values to pNos:luc (represented as a dotted line) control are shown. An internal 35S:luc control was used. Asterisks indicate Student's t-test significant values (one asterisk for $p < 0,05$; two asterisks for $p < 0,01$; three asterisks for $p < 0,001$).

dCas9:SunTag5aa shows (Figure 20A) significant repression only for gRNA positions -35 and +51, and acceptable levels of repression were only observed for gRNAs that target position -35 to TSS with ScFv:SRDX, with approximately 40% repression. This is not a significant improvement from the direct fusion strategy (Figure 12), where gRNA that targets position -35 to TSS with dCas9:SRDX also achieved repression levels close to 40%.

dCas12a:SunTag5aa shows(Figure 20B) significant repression for all gRNAs selected with both repression domains. However, repression levels were similar to those obtained with the direct fusion (Figure 16).

3.1.2. Double gRNA repression

In order to increase the repression levels obtained with the SunTag strategy, selected, gRNAs were tested in pairs for both endonucleases with ScFv:BRD and ScFv:SRDX. Results are shown on Figure 21.

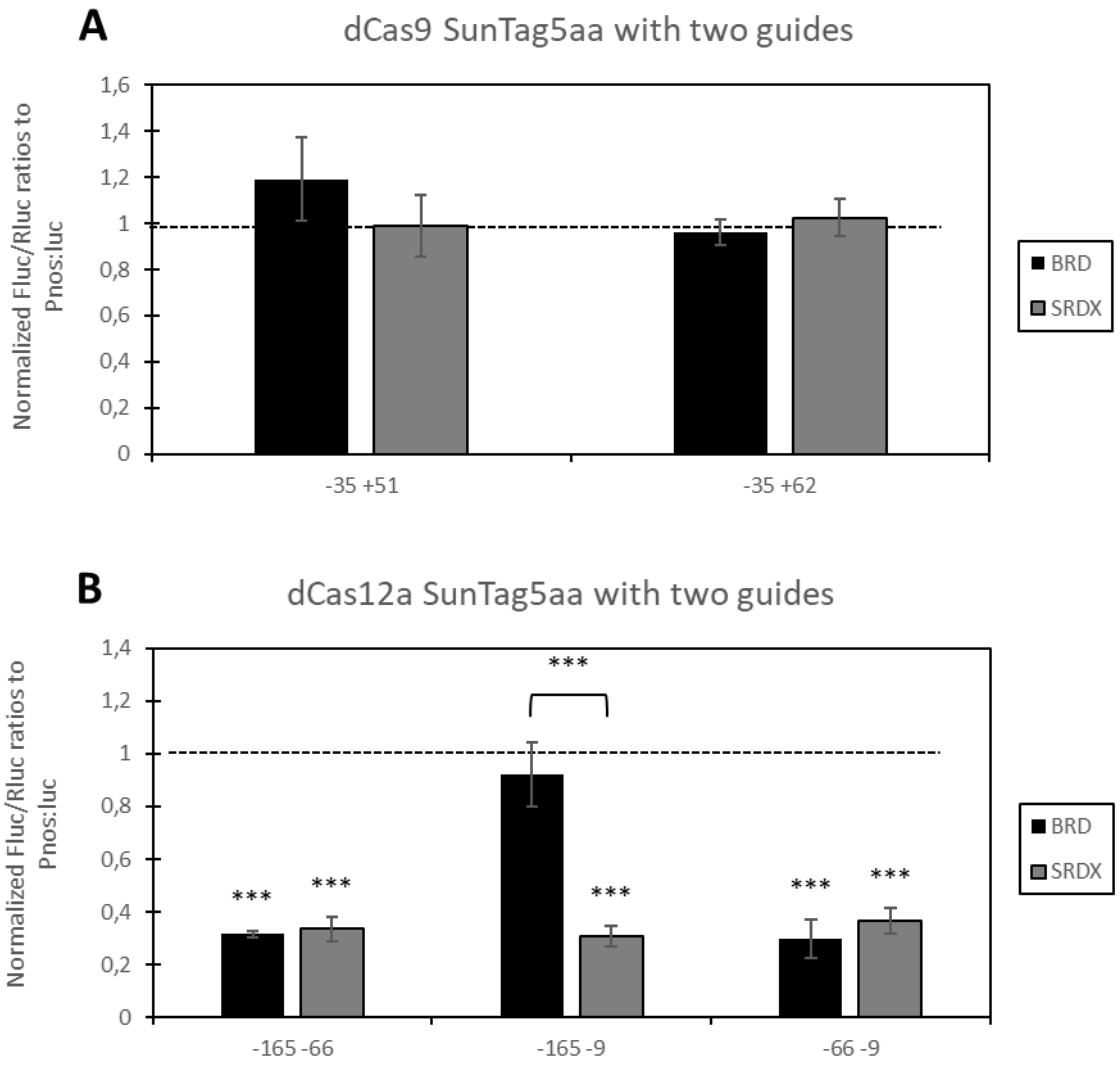

Figure 21. Transcriptional repression achieved using two gRNA with a SunTag 5aa strategy for dCas9 (A) and dCas12a (B). Normalized Luciferase/Renilla values to pNos:luc (represented as a

dotted line) control are shown. An internal 35S:luc control was used. Asterisks indicate Student's t-test significant values (one asterisk for $p < 0,05$; two asterisks for $p < 0,01$; three asterisks for $p < 0,001$).

Using two gRNAs, none of the combinations showed significant repression for dCas9:SunTag5aa (Figure 21A). Significant repression was observed for most combinations for dCas12a:SunTag5aa (Figure 21B) but repression levels were not different from those observed with the direct fusion of a repression domain (Figure 18).

The data observed suggests that this method is not an improvement from the direct fusion of a single repression domain. Evidence suggests that using only 5 amino acids for the spacers may cause steric impediments between antibodies, which would hinder the recruitment of more repression domains (Papikian et al., 2019).

3.2. SunTag 22aa repression

A recent study described an optimization of the SunTag strategy that includes 22 amino acids as spacers instead of 5, in order to avoid the steric hindrance between the ScFV antibodies (Papikian et al., 2019). This SunTag, named Suntag22aa, was assayed with the same system described previously (Figure 19) for selected gRNAs with ScFv:BRD and ScFv:SRDX in order to test if this new SunTag design could improve repression by efficiently recruiting more repression domains.

3.2.1. Single grRNA repression

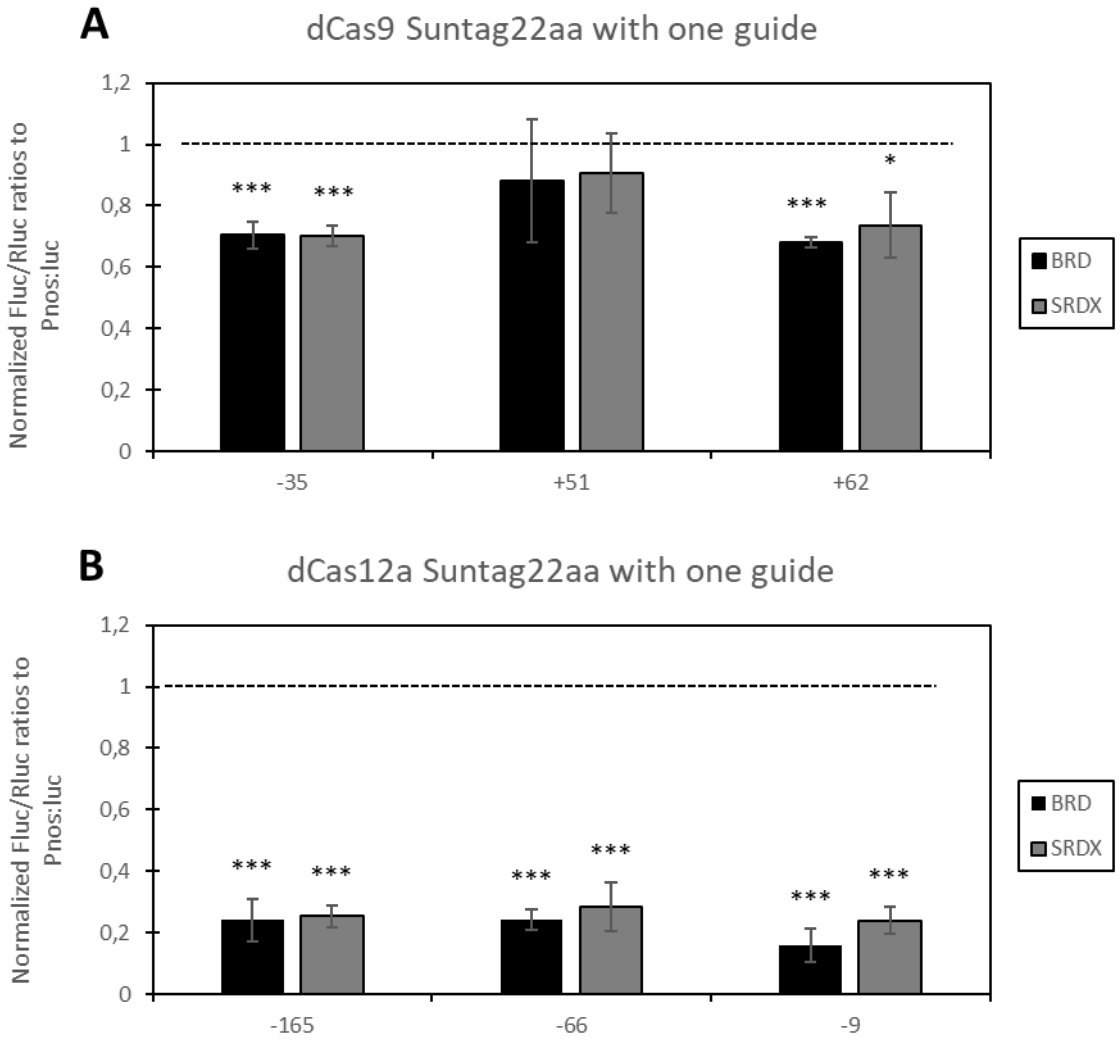

Figure 22. Transcriptional repression achieved using one gRNA with a SunTag 22aa strategy for dCas9 (A) and dCas12a (B). Normalized Luciferase/Renilla values to pNos:luc (represented as a dotted line) control are shown. An internal 35S:luc control was used. Asterisks indicate Student's t-test significant values (one asterisk for $p < 0,05$; two asterisks for $p < 0,01$; three asterisks for $p < 0,001$).

dCas9:SunTag22aa showed significant repression with ScFv:BRD and ScFv:SRDX positions -35 and +62 (Figure 22A). However, these repression levels are not higher than those observed with the single domain strategy (Figure 13).

dCas12a:Suntag22aa, however, showed significant repression for all gRNAs with ScFv:BRD and ScFv:SRDX (Figure 22B) and these levels are higher than those observed with the single domain strategy (Figure 16). All of the gRNAs showed close to 80% repression with both domains, and no difference between domains was observed. These repression levels are close to those observed with two gRNAs and a single repression domain (Figure 18) for dCas12a:BRD and dCas12a:SRDX, which suggests that recruiting more domains also increases repression efficiency.

3.2.2. Double gRNA repression

Following the same approach as with previous strategies, selected gRNAs were tested in pairs for both endonucleases with ScFv:BRD and ScFv:SRDX to try to increase repression efficiency. Results are shown on Figure 23.

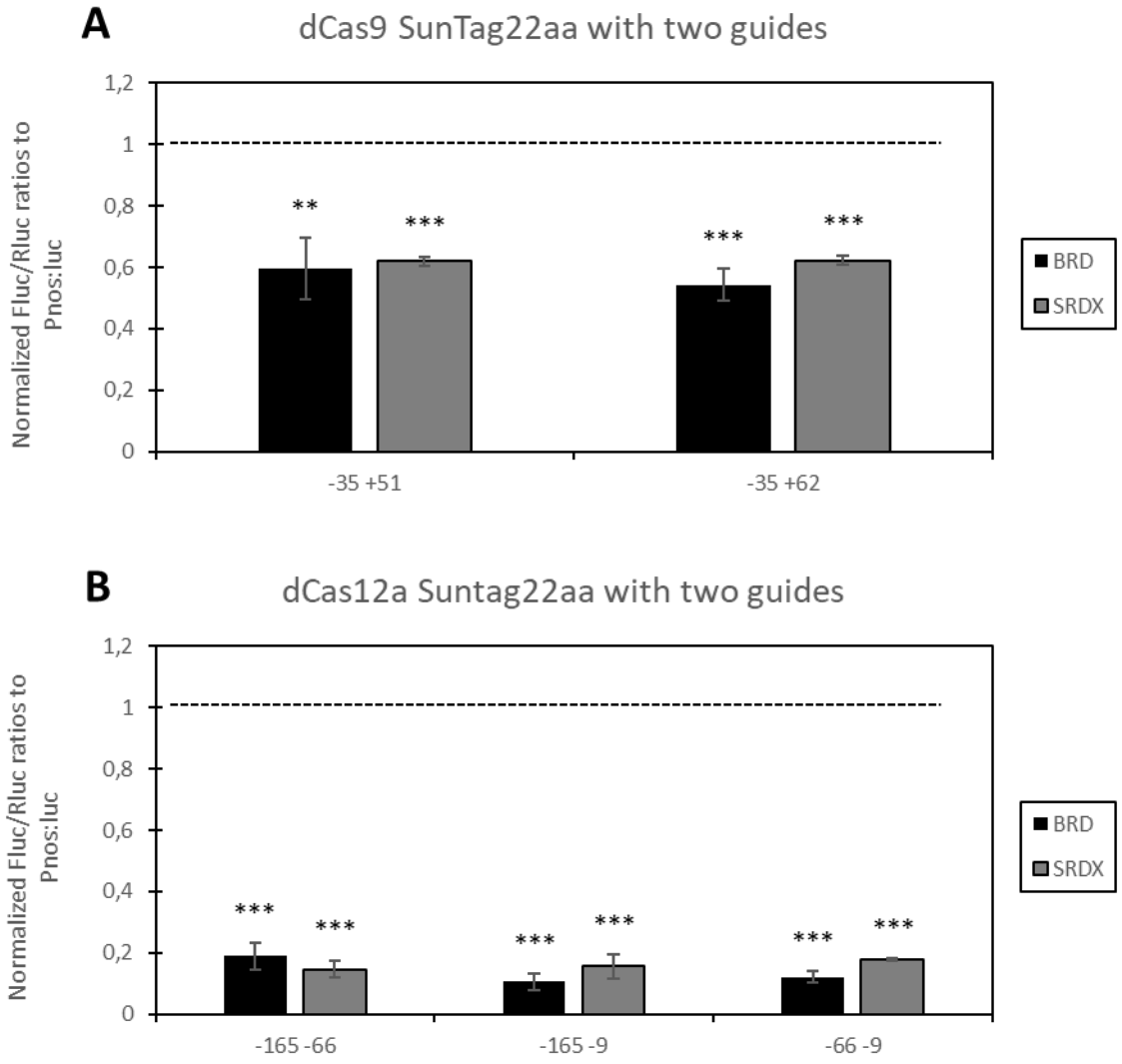

Figure 23. Transcriptional repression achieved using two gRNA with a SunTag 22aa strategy for dCas9 (A) and dCas12a (B). Normalized Luciferase/Renilla values to pNos:luc (represented as a dotted line) control are shown. An internal 35S:luc control was used. Asterisks indicate Student's t-test significant values (one asterisk for $p < 0,05$; two asterisks for $p < 0,01$; three asterisks for $p < 0,001$).

gRNAs combinations showed significant repression for both domains with dCas9:SunTag22aa (Figure 23A). As the previous assays showed, repression levels with dCas9 are not higher than those observed when recruiting a single repression domain (Figure 18). ScFv:BRD repression levels are similar to those observed with a single repression domain while ScFv:SRDX repression levels are lower.

Transcriptional repression obtained with dCas12a:SunTag22aa (Figure 23B), shows significant repression for every gRNA combination with both repression domains. For this endonuclease, repression levels were higher with this strategy than those observed when a single repression domain was used (Figure 16). Repression levels are as high as 90% with this strategy and no difference is observed between repression domains.

The data obtained with this study suggests that dCas12a is a better endonuclease for repression than dCas9. Previous evidence suggests dCas9, dCas12a and other similar endonucleases need to open the chromatin to bind to their target site (Barkal, Srinivasan, Hashimoto, Gifford, & Sherwood, 2016). Chromatin state is an essential regulator of gene expression (Roudier et al., 2011), where open chromatin allows for gene expression while closed, folded chromatin makes it more difficult for the translation machinery to access genes. This means genes in an open chromatin state will always be more easily expressed.

This poses a problem for gene repression using endonucleases such as dCas9. In order to bring the fused repression domains to the target gene, dCas9 and dCas12a need to access the gene by opening the chromatin to bind to their target site (Barkal et al., 2016), making target genes more accessible and enhancing gene expression. Gene activation using dCas9 has already been described (Selma et al., 2019) and could be enhanced due to dCas9 endogenous activity in opening the chromatin.

In the case of repression, endonuclease size could prove to be a key factor. Smaller endonucleases may have an easier time accessing the target gene and less chromatin will need to be opened to bind to its target site. In this case, SpdCas9, used in this study, has a size of approximately 1400 amino acids (Cebrian-Serrano & Davies, 2017) while LbdCas12a has a size of approximately 1250 amino acids (Koo et al., 2018). The difference in size could explain the higher levels of repression efficiency with dCas12a for all strategies used.

Repression efficiency also seems to be dependent on number of gRNAs used, which can be observed on the tests conducted with more than one gRNA at the same time (Figures 18, 21 and 23). This has also been described for gene activation (Selma et al., 2019) and gene editing (Bernabé-Orts et al., 2019), and for gene repression in other organisms (Miao et al., 2019).

Different repression domains also affect repression efficiency. Both BRD and SRDX are repression domains present in plants and performed similarly, with SRDX showing slightly better repression efficiency for the single repression domain strategy using more than one gRNA (Figures 16, 20). KRAB domain didn't seem to perform properly in most of the tests conducted (Figures 14, 16), with considerately lower repression levels than other domains tested (SRDX, BRD). KRAB domain works by recruiting other transcriptional repressors like the heterochromatin protein 1 (HP1) family, which in turn fold the chromatin to epigenetically repress translation (Janssen, Chen, Liu, & Gonçalves,

2019). Since this domain is exclusive to tetrapod vertebrates (Lupo et al., 2013), its low repression efficiency might be explained by a low capability of recruiting plant translational repressors. BRD and SRDX domains also work by recruiting other transcriptional repressors for epigenetic remodelling of the chromatin and may also be able to interact directly with the transcription machinery (Ikeda & Ohme-Takagi, 2009; Kagale & Rozwadowski, 2011) but they are endogenous plant domains.

Increasing the number of repression domains recruited also increases repression efficiency. The initial SunTag, with 5 amino acids as spacers (SunTag5aa), does not increase repression for either endonuclease (Figures 20, 21), probably due to the steric impediments already mentioned (Papikian et al., 2019). The improved SunTag22aa, however, improves repression considerately for dCas12a with both domains (Figures 22, 23). Lower impact of the SunTag strategy with dCas9 could be explained by the increase on its size, which would exacerbate the effect of the already big size of dCas9 on chromatin structure.

Other factors such as DNA strand do not seem to affect repression. However, during the course of this study new evidence appeared suggesting PAM sequence to be key in the regulation efficiency of dCas12a (Miao et al., 2019). Not enough gRNAs were tested in this study to extract any conclusion regarding PAM sequence, but it might be interesting to test in further assays.

Direct epigenetic repression without the need to recruit additional transcription factors could also be interesting to alleviate the effects of dCas9 and dCas12a binding on chromatin state. To this end, trying plant epigenetic repression domains like the *Nicotiana benthamiana* DRM methyltransferase (Papikian et al., 2019) in future experiments could be the next step to increasing negative transcriptional regulation efficiency in plants.

Finally, all of our data suggests that the best repression system out of those tested is dCas12a:Suntag22aa, increasing the number of repression domains recruited by utilizing the SunTag and using more than one gRNA. With this system, there seems to be no difference in repression efficiency between ScFv:BRD and ScDv:SRDX, although the SRDX domain does have better repression efficiency in most of the strategies tested.

Conclusions

CONCLUSIONS

I. Optimal gRNA positioning was determined as upstream and downstream from the transcription starting site (TSS) but close to it for dCas9, and as upstream from the TSS for dCas12a.

II. BRD and SRDX domains both show good repression efficiency for both endonucleases, with SRDX showing better results for some of the assays. KRAB domain has too much variability and shows lower repression values than the other two domains.

III. Using two gRNAs increases repression efficiency.

IV. The SunTag 5aa isn't an improvement from the single domain fusion strategy due to steric hindrances. However, SunTag 22aa greatly improves repression with dCas12a by recruiting more repression domains.

V. The best strategy for repression out of all of the ones tested is dCas12a:Suntag22aa using two gRNAs and ScFv:BRD or ScFv:SRDX as repression domains.

References

REFERENCES

Add experiment. (n.d.). Retrieved January 5, 2020, from https://gbcloning.upv.es/add/experiment/SE_002

Bally, J., Jung, H., Mortimer, C., Naim, F., Philips, J. G., Hellens, R., … Waterhouse, P. M. (2018). The Rise and Rise of Nicotiana benthamiana : A Plant for All Reasons. *Annual Review of Phytopathology*, *56*(1), 405–426. https://doi.org/10.1146/annurev-phyto-080417-050141

Barkal, A. A., Srinivasan, S., Hashimoto, T., Gifford, D. K., & Sherwood, R. I. (2016). Cas9 Functionally Opens Chromatin. *PLOS ONE*, *11*(3), e0152683. https://doi.org/10.1371/journal.pone.0152683

Benner, S. A., & Sismour, A. M. (2005). Synthetic biology. *Nature Reviews Genetics*, *6*(7), 533–543. https://doi.org/10.1038/nrg1637

Bernabé-Orts, J. M., Casas-Rodrigo, I., Minguet, E. G., Landolfi, V., Garcia-Carpintero, V., Gianoglio, S., … Orzaez, D. (2019). Assessment of Cas12a-mediated gene editing efficiency in plants. *Plant Biotechnology Journal*, *17*(10), 1971–1984. https://doi.org/10.1111/pbi.13113

Bortesi, L., & Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances*, *33*(1), 41–52. https://doi.org/10.1016/j.biotechadv.2014.12.006

Campa, C. C., Weisbach, N. R., Santinha, A. J., Incarnato, D., & Platt, R. J. (2019). Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nature Methods*, *16*(9), 887–893. https://doi.org/10.1038/s41592-019-0508-6

Canto, T. (2016). Transient Expression Systems in Plants: Potentialities and Constraints. In *Advances in Experimental Medicine and Biology* (Vol. 896, pp. 287–301). https://doi.org/10.1007/978-3-319-27216-0_18

Cebrian-Serrano, A., & Davies, B. (2017). CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. *Mammalian Genome*, *28*(7–8), 247–261. https://doi.org/10.1007/s00335-017-9697-4

Cloud-Based Informatics Platform for Life Sciences R&D | Benchling. (n.d.). Retrieved December 6, 2019, from https://www.benchling.com/

Domestication. (n.d.). Retrieved January 4, 2020, from https://gbcloning.upv.es/tools/domestication/

Dominguez, A. A., Lim, W. A., & Qi, L. S. (2016). Beyond editing: repurposing CRISPR–Cas9 for precision genome regulation and interrogation. *Nature Reviews Molecular Cell Biology*, *17*(1), 5–15. https://doi.org/10.1038/nrm.2015.2

Goodin, M. M., Zaitlin, D., Naidu, R. A., & Lommel, S. A. (2008). Nicotiana benthamiana: Its History and Future as a Model for Plant-Pathogen Interactions. */ 1015 MPMI*, *21*(8), 1015–1026. https://doi.org/10.1094/MPMI

Ikeda, M., & Ohme-Takagi, M. (2009). A Novel Group of Transcriptional Repressors in

Arabidopsis. *Plant and Cell Physiology*, *50*(5), 970–975. https://doi.org/10.1093/pcp/pcp048

Janssen, J. M., Chen, X., Liu, J., & Gonçalves, M. A. F. V. (2019). The Chromatin Structure of CRISPR-Cas9 Target DNA Controls the Balance between Mutagenic and Homology-Directed Gene-Editing Events. *Molecular Therapy - Nucleic Acids*, *16*, 141–154. https://doi.org/10.1016/j.omtn.2019.02.009

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, *337*(6096), 816–821. https://doi.org/10.1126/science.1225829

Johansen, L. K., & Carrington, J. C. (2001). Silencing on the Spot. Induction and Suppression of RNA Silencing in the Agrobacterium -Mediated Transient Expression System. *Plant Physiology*, *126*(3), 930–938. https://doi.org/10.1104/pp.126.3.930

Kagale, S., & Rozwadowski, K. (2011). EAR motif-mediated transcriptional repression in plants: An underlying mechanism for epigenetic regulation of gene expression. *Epigenetics*, *6*(2), 141–146. https://doi.org/10.4161/epi.6.2.13627

Koo, T., Park, S. W., Jo, D. H., Kim, D., Kim, J. H., Cho, H.-Y., … Kim, J.-S. (2018). CRISPR-LbCpf1 prevents choroidal neovascularization in a mouse model of age-related macular degeneration. *Nature Communications*, *9*(1), 1855. https://doi.org/10.1038/s41467-018-04175-y

Korbie, D. J., & Mattick, J. S. (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nature Protocols*, *3*(9), 1452–1456. https://doi.org/10.1038/nprot.2008.133

Lo, A., & Qi, L. (2017). Genetic and epigenetic control of gene expression by CRISPR–Cas systems. *F1000Research*, *6*, 747. https://doi.org/10.12688/f1000research.11113.1

Lupo, A., Cesaro, E., Montano, G., Zurlo, D., Izzo, P., & Costanzo, P. (2013). KRAB-Zinc Finger Proteins: A Repressor Family Displaying Multiple Biological Functions. *Current Genomics*, *14*(4), 268–278. https://doi.org/10.2174/13892029113149990002

Mark, C., Åbrink, M., & Hellman, L. (1999). Comparative analysis of KRAB zinc finger proteins in rodents and man: Evidence for several evolutionarily distinct subfamilies of KRAB zinc finger genes. *DNA and Cell Biology*, *18*(5), 381–396. https://doi.org/10.1089/104454999315277

Marraffini, L. A. (2015). CRISPR-Cas immunity in prokaryotes. *Nature*, *526*(7571), 55–61. https://doi.org/10.1038/nature15386

Miao, C., Zhao, H., Qian, L., & Lou, C. (2019). Systematically investigating the key features of the DNase deactivated Cpf1 for tunable transcription regulation in prokaryotic cells. *Synthetic and Systems Biotechnology*, *4*(1), 1–9. https://doi.org/10.1016/j.synbio.2018.11.002

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements.

Journal of Molecular Evolution, *60*(2), 174–182. https://doi.org/10.1007/s00239-004-0046-3

Papikian, A., Liu, W., Gallego-Bartolomé, J., & Jacobsen, S. E. (2019). Site-specific manipulation of Arabidopsis loci using CRISPR-Cas9 SunTag systems. *Nature Communications*, *10*(1), 729. https://doi.org/10.1038/s41467-019-08736-7

Patron, N. J., Orzaez, D., Marillonnet, S., Warzecha, H., Matthewman, C., Youles, M., … Haseloff, J. (2015). Standards for plant synthetic biology: A common syntax for exchange of DNA parts. *New Phytologist*, *208*(1), 13–19. https://doi.org/10.1111/nph.13532

Reed, J., & Osbourn, A. (2018). Engineering terpenoid production through transient expression in Nicotiana benthamiana. *Plant Cell Reports*, *37*(10), 1431–1441. https://doi.org/10.1007/s00299-018-2296-3

Roudier, F., Ahmed, I., Bérard, C., Sarazin, A., Mary-Huard, T., Cortijo, S., … Colot, V. (2011). Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *The EMBO Journal*, *30*(10), 1928–1938. https://doi.org/10.1038/emboj.2011.103

Sanson, K. R., Hanna, R. E., Hegde, M., Donovan, K. F., Strand, C., Sullender, M. E., … Doench, J. G. (2018). Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nature Communications*, *9*(1), 5416. https://doi.org/10.1038/s41467-018-07901-8

Sarrion-Perdigones, A., Vazquez-Vilar, M., Palaci, J., Castelijns, B., Forment, J., Ziarsolo, P., … Orzaez, D. (2013). GoldenBraid 2.0: A Comprehensive DNA Assembly Framework for Plant Synthetic Biology. *PLANT PHYSIOLOGY*, *162*(3), 1618–1631. https://doi.org/10.1104/pp.113.217661

Selma, S., Bernabé-Orts, J. M., Vazquez-Vilar, M., Diego-Martin, B., Ajenjo, M., Garcia-Carpintero, V., … Orzaez, D. (2019). Strong gene activation in plants with genome-wide specificity using a new orthogonal <scp>CRISPR</scp> /Cas9-based programmable transcriptional activator. *Plant Biotechnology Journal*, *17*(9), pbi.13138. https://doi.org/10.1111/pbi.13138

Stewart, C. N., Patron, N., Hanson, A. D., & Jez, J. M. (2018). Plant metabolic engineering in the synthetic biology era: plant chassis selection. *Plant Cell Reports*, *37*(10), 1357–1358. https://doi.org/10.1007/s00299-018-2342-1

Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S., & Vale, R. D. (2014). A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell*, *159*(3), 635–646. https://doi.org/10.1016/j.cell.2014.09.039

Urrutia, R. (2003). KRAB-containing zinc-finger repressor proteins. *Genome Biology*, Vol. 4. https://doi.org/10.1186/gb-2003-4-10-231

Vanegas, K. G., Jarczynska, Z. D., Strucko, T., & Mortensen, U. H. (2019). Cpf1 enables fast and efficient genome editing in Aspergilli. *Fungal Biology and Biotechnology*, *6*(1), 6. https://doi.org/10.1186/s40694-019-0069-6

Vazquez-Vilar, M., Bernabé-Orts, J. M., Fernandez-del-Carmen, A., Ziarsolo, P., Blanca, J., Granell, A., & Orzaez, D. (2016). A modular toolbox for gRNA–Cas9 genome engineering

in plants based on the GoldenBraid standard. *Plant Methods*, *12*(1), 10. https://doi.org/10.1186/s13007-016-0101-2

Vazquez-Vilar, M., Quijano-Rubio, A., Fernandez-del-Carmen, A., Sarrion-Perdigones, A., Ochoa-Fernandez, R., Ziarsolo, P., … Orzaez, D. (2017). GB3.0: a platform for plant bio-design that connects functional DNA elements with associated biological data. *Nucleic Acids Research*, *45*(4), gkw1326. https://doi.org/10.1093/nar/gkw1326