CHAPTER 4

Modulation of gene expression in plants via CRISPR/dCas9 technology

Carolina Vianna Morgante
Fabricio Barbosa Monteiro Arraes
Clidia Eduarda Moreira-Pinto
Bruno Paes de Melo
Maria Fatima Grossi-de-Sa

Introduction

The regulation of gene expression includes a diversity of cellular processes that occur in a coordinated manner and in multiple stages to trigger the increase or reduction of a specific gene product. Gene expression can be induced by endogenous and environmental stimuli and regulated at different cellular levels, such as in the initiation of the transcription, RNA processing, and post-translational modification of the protein.

The manipulation of target genes is of paramount importance for understanding gene function and reprogramming of cellular activities. This allows the deepening of basic knowledge about biochemical and molecular processes and the intensification of characteristics of agronomic interest. At this point, precision is essential to obtain the success in applications of genetic engineering and synthetic biology.

In the last decades, technologies using site-specific nucleases for the precise manipulation of DNA have undergone a profound advance, emerging as promising alternatives for site-directed mutagenesis and fine control of gene expression. Among these technologies stand out those of genome editing, such as zinc finger nuclease (ZFN), Transcription Activator-Like Effector Nucleases (TALENs), and more recently, the CRISPR/Cas (Clustered Regularly Interspaced Short Palindromic Repeats) associated with Cas nuclease. The latter has its revolutionary character, especially its specificity, universality, and relative simplicity (Pickar-Oliver; Gersbach, 2019). Also, CRISPR/Cas is a flexible tool that can be modified, which contributes to its continuous improvement and diverse applications in the study of cellular functions and biotechnology.

This chapter discusses CRISPR technology using the dead Cas9 variant (CRISPR/dCas9), emphasizing on its use for modulating gene expression in plants. The main strategies currently used are presented and discussed, for which a script was proposed, considering the main aspects for a good experimental design.

The dCas9 system

The CRISPR/Cas9 system was first identified in *Streptococcus pyogenes* and, subsequently, in other bacteria and most archaea, as a sophisticated adaptive immune system, guided by RNAs encoded by the CRISPR locus and Cas proteins (CRISPR associated protein), to provide acquired immunity mainly against bacteriophages (Hsu et al., 2014; Zhang et al., 2014). Its application in genetic engineering was made possible by the deep elucidation of the structures of Cas9 and guide RNA (gRNA).

The Cas9 from *S. pyogenes* is a large multifunctional protein (1.368 amino acid residues) composed of five domains and with nuclease activity (Jiang; Doudna, 2017). The structure of the apoenzyme Cas9 comprises the alpha-helical recognition lobe (REC) and the nucleic lobe (NUC). The latter contains the conserved nuclease domains HNH and tripartite RuvC, in addition to a more variable C-terminal domain (CTD) (Figure 1) (Jinek et al., 2014; Nishimasu et al., 2014).

In the native CRISPR/Cas9 system, mature gRNA is composed of two independent molecules of small non-coding RNAs, which interact with each other by base complementarity: (i) crRNA (CRISPR RNA), responsible for the recognition of the target DNA; and (ii) tracrRNA (transactivating crRNA or scaffold RNA), important for the anchoring of gRNA into Cas9 (Jinek et al., 2012). In vitro studies determined the minimum sequence of crRNA and tracrRNA required for its use in synthetic models. These two minimal sequences were joined, thus creating the sgRNA (single guide RNA) (Figure 1) (Anders et al., 2014; Dang et al., 2015; Hsu et al., 2013; Jinek et al., 2012; Nishimasu et al., 2014). After recognizing the target DNA sequence, through its complementarity with gRNA/sgRNA, Cas9 cleaves the double-stranded DNA (dsDNA) at three base pairs upstream from the PAM sequence (Protospacer Adjacent Motif), using its two distinct nuclease domains.

The HNH domain cleaves the DNA strand complementary to the sgRNA sequence (target strand), while the RuvC domain is responsible for cleaving the DNA strand opposite to the complementary strand (non-target strand) (Jiang; Doudna, 2017). Studies have shown that specific mutations in the domains HNH (for example, H840A) or RuvC (for example, D10A) can convert Cas9 into a nicking enzyme, which cleaves only one strand of dsDNA. Mutations in both domains abolish Cas9 endonuclease activity without interfering with its RNA-guided DNA-targeting ability, thus being designated as dead Cas9 or dCas9 (Table 1) (Jiang; Doudna, 2017; Jinek et al., 2014, 2012).

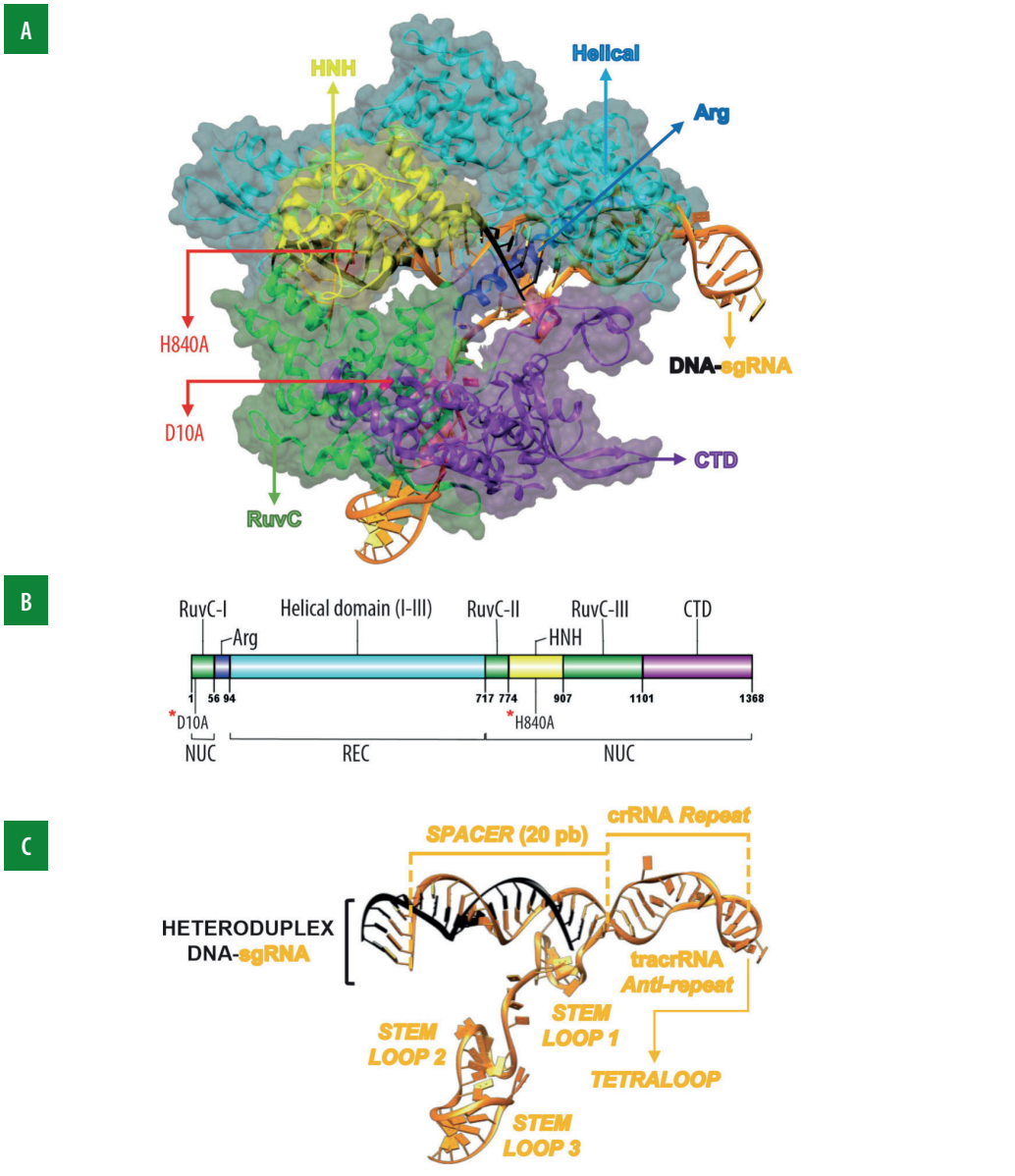


Figure 1. Structure of Cas9/dCas9 and sgRNA. Structural model (A) and diagrammatic representation (B) of the functional domains of Cas9/dCas9 associated with sgRNA, interacting with the target DNA strand (complementary strand represented in black). The NUC lobe comprises the C-terminal domain (CTD), RuvC-II and RuvC-III (tripartite) domains, and HNH nuclease domain. The dCas9 mutations in RuvC (D10A) and HNH (H840A) domains are represented in red (A) and with asterisks (B). Single mutation confers nicking activity to Cas9, while both together abolish its nuclease activity. The lobe REC is composed exclusively of the Helical domain. The structure of sgRNA (C) includes the crRNA (spacer + crRNA repeat) and scaffold RNA or tracrRNA (in this scheme, the anti-repeat tracrRNA is linked to crRNA repeat by a tetraloop, followed by three-stem loops). The structural representations are based on models deposited in a public database (PDB ID: 4008), available at https://www.rcsb.org/.

Table 1. Site directed mutagenesis on Cas9 nuclease of *Streptococcus pyogenes* (SpCas9) (UniProtKB - Q99ZW2) and its effects.

Position	Mutation[(1)]	Mutation effects	Reference
10	D → A	Non-complementary DNA strand to crRNA is not cleaved; nicking activity	Nishimasu et al. (2014)
15	S → A	Reduction of target DNA cleavage efficiency	Nishimasu et al. (2014)
66	R → A	Significant reduction of target DNA cleavage efficiency	Nishimasu et al. (2014)
70	R → A	Absence of target DNA cleavage	Nishimasu et al. (2014)
74	R → A	Significant reduction of target DNA cleavage efficiency	Nishimasu et al. (2014)
78	R → A	Moderate reduction of target DNA cleavage efficiency	Nishimasu et al. (2014)
97 – 150	Deletion	Absence of nuclease activity	Nishimasu et al. (2014)
165	R → A	Moderate reduction of target DNA cleavage efficiency	Nishimasu et al. (2014)
175 – 307	Deletion	50% loss of nuclease activity	Nishimasu et al. (2014)
302 – 409	Deletion	Absence of nuclease activity	Nishimasu et al. (2014)
475 – 477	PWN → AAA	Slight reduction in cleavage and interaction with target DNA	Jinek et al. (2014)
762	E → A	Cleavage of only one strand of the target DNA, probably the non-complementary to the crRNA	Nishimasu et al. (2014)
840	H → A	Non-complementary DNA strand to crRNA is not cleaved; nicking activity	Anders et al. (2014)
854	N → A	Reduction of target DNA cleavage efficiency	Nishimasu et al. (2014)
863	N → A	Cleavage of only one strand of the target DNA, probably the non-complementary to the crRNA	Nishimasu et al. (2014)
982 – 983	HH → AA	Correct sgRNA processing	Fonfara et al. (2014)
982	H → A	Reduction of target DNA cleavage efficiency	Nishimasu et al. (2014)
983	H → A	Cleavage of only one of the strands of the target DNA, probably not complementary to the crRNA	Nishimasu et al. (2014)
986	D → A	Cleavage of only one of the strands of the target DNA, probably complementary to the crRNA	Nishimasu et al. (2014)
1099 – 1368	Deletion	Nuclease activity is not detected	Nishimasu et al. (2014)
1125 – 1127	DWD → AAA	No change in cleavage efficiency, but a slight reduction in interaction with the target DNA	Jinek et al. (2014)
1132	G → C	Protein inactivation (probably)	Jiang et al. (2013)

To be continued...

Table 1. Continued.

Position	Mutation[(1)]	Mutation effects	Reference
1133 – 1135	RKR → AKA	Almost complete loss of nuclease activity	Anders et al. (2014)
1133	R → A	Dramatic reduction in the interaction with the target DNA, but a slight reduction in cleavage efficiency	Anders et al. (2014)
1135	R → A	Dramatic reduction in the interaction with the target DNA, but a slight reduction in cleavage efficiency	Anders et al. (2014)

[(1)] Abbreviation of amino acids: A - alanine; D - aspartic acid; E - glutamic acid; G - glycine; H - histidine; K - lysine; N - asparagine; P - proline; R - arginine; S - serine; W - tryptophan.

Applications of the dCas9 system in plants

The CRISPR system applications go beyond those related to site-directed mutagenesis by cleavage of dsDNA and activation of the repair system. The dCas9 protein offers a unique, multifunctional, and dynamic platform for the recruitment of proteins with different functions specific sites in the genome to promote, for example, the control of transcription regulation, epigenome editing, base editing, and genome imaging.

Base editing has become a powerful tool in plant biotechnology for site-directed mutagenesis. The development of herbicide-tolerant varieties for wheat, rice, corn, and watermelon crops, for example, was made possible by the introduction of point mutations into the genes encoding the enzymes acetolactate synthase and/or acetyl-coenzyme A, so that the final product is not transgenic (Kuang et al., 2020; Li et al., 2019; Tian et al., 2018; Zhang et al., 2019b). Base-editing systems require the fusion of Cas9, dCas9, or nCas9 variants, guided by a sgRNA, fused to a cytosine (C) deaminase, which induces the conversion of CG (Guanine) to TA (Thymine-Adenine) (CBE, cytosine base editor), or an A deaminase, which promotes the conversion of AT to GC (ABE, adenine base editor) (Gaudelli et al., 2017; Komor et al., 2016; Nishida et al., 2016). These systems have been successfully adapted for plants, including multiplex systems, in which more than one gene is simultaneously a target for mutations, with high efficiency, specificity, and without the occurrence of indels at the editing site. More satisfactory results were obtained with the use of nCas9 (Hua et al., 2018; Kang et al., 2018; Zong et al., 2017).

Another potent application of CRISPR/dCas9 is in genome imaging, used to visualize the spatial organization and temporal interactions of chromatin in real-time. For that, dCas9 is fused to fluorescent proteins, such as GFP (Green Fluorescent

Protein), resulting, for example, in a robust visualization of telomeric repetitions and their movements, multiple genomic loci in living cells, and protein-DNA interactions (Chen et al., 2013; Dreissig et al., 2017). Khosravi et al. (2020) optimized the method by inserting aptamers into the sgRNA scaffold able to recruit binding proteins fused to fluorescent proteins. Thus, a more significant number of reporter proteins is mobilized to the targeted sequence resulting in improved labeling.

The CRISPR/dCas9 system is a powerful tool to modulate the transcription of protein-coding and non-coding genes. It is a flexible and reversible tool since it can be used both for activation (CRISPRa, activator) and for repression (CRISPRi, interfering) of the transcription without permanently modifying the genome. Additionally, CRISPR/dCas9 can be used in multigene approaches to simultaneously modulate the expression of more than one gene, even in opposite directions. In this methodology, dCas9 and/or sgRNA are fused to transcriptional modulators (ModT), which act on activation or repression of the transcription. The ModT are proteins or protein domains that bind DNA to recruit key regulatory elements to control gene expression.

The fusion of ModT to dCas9 emerged as a strategy to enhance the effects of CRISPRi technology, initially developed based on the blockage of transcription, by physical interference on the binding of RNA polymerase and transcription factors to Transcription Start Site (TSS) or elongation process (Qi et al., 2013) (Figure 2A). The strategy of fusing ModT to dCas9 allowed the optimization of transcriptional repression and the development of a methodology for activation (Gilbert et al., 2013). Thus, CRISPR/dCas9 can be considered a generic and universal platform, since it is capable of promoting transcriptional activation and repression at different degrees in different species. In this way, the same transcriptional repressor or activator can be used to regulate transcription in cells of different species, with efficiencies varying according to the target genomic region, sgRNA design, delivery strategy (transformation method), and biological system in study.

The epigenetic regulation of chromatin through ModT, as acetyl and methyltransferases, is called epigenome editing (Hilton et al., 2015; Thakore et al., 2015). In plants, this method was used to induce drought tolerance in *Arabidopsis thaliana*, by the overexpression of *AtAREB1* gene (Abscisic Acid-Responsive Element Binding Protein 1), using dCas9 fused to an acetyltransferase 1, and also to alter the flowering time, through epigenetic regulation of the *AtFT* gene (Flowering Locus T) in *A. thaliana*, using the CRISPR/dCas9 system associated with an acetyltransferase or a methyltransferase (Lee et al., 2019; Roca-Paixão et al., 2019).

Next, the main ModTs, including transcriptional activators (AtvT) and transcriptional repressors (RepT), are presented, as well as the main strategies for modulating gene expression through the CRISPR/dCas9 system.

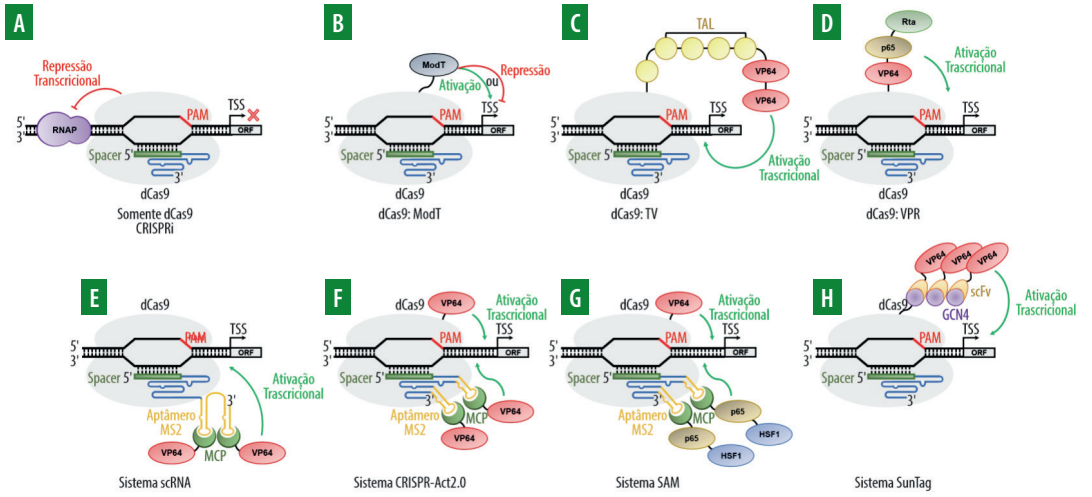


Figure 2. Strategies for the use of CRISPR/dCas9 technology in modulating gene expression. (A) Repression (CRISPRi, interfering) via RNA Polymerase (RNAP) blockage by the dCas9-sgRNA complex; (B) Activation or suppression by the fusion in tandem of transcriptional modulators (ModT) to dCas9; (C) Activation by the combination of ModTs (VP64 and TAD) fused to dCas9 (dCas9: VT); (D) Activation by the combination of ModTs (VP64: p65: Rta, VPR) fused to dCas9; (E) Scaffold RNA system (scRNA), in which the sgRNA containing an aptamer MS2 at its 3 'end recruits ModT (VP64) via fusion with MCP (MS2 Coat Protein); (F) CRISPR-Act2.0 system in which dCas9: VP64 is used in combination with a modified sgRNA with two aptamers MS2 that recruit ModT via MCP; (G) SAM system, optimization of the CRISPR-Act2.0 system, in which dCas9 is fused to a combination of ModTs (p65-HSF1); (H) SunTag system that consists of fusing dCas9 to tandem repeats of the GCN4 peptide that recruits ModT (VP64) via antibody (scFv) that binds to GCN4. TSS - transcription start site; ORF - open reading frame; PAM - protospacer adapter motif.

Strategies for modulating gene expression via CRISPR/dCas9

Increase of gene expression (CRISPRa)

Fusion of transcriptional activators to dCas9 protein (dCas9: AtvT)

This can be considered the pioneer strategy for activating transcription in plants by the CRISPR/dCas9 system. The sequence of a transcriptional activator (AtvT) is fused in tandem to the C-terminal region of the dCas9 protein. This complex is specifically targeted to the promoter region of a target gene by sgRNA (Figure 2B). Activators such as VP64, EDLL, TAD, and HAT (Table 2) have been used successfully to increase the expression of coding and non-coding genes (microRNAs) in plants, which can

be enhanced with the use of multiple sgRNAs complementary to sequences of the same promoter region (Lowder et al., 2015; Piatek et al., 2015; Roca Paixão et al., 2019; Vazquez-Vilar et al., 2016). However, the ideal number of sgRNAs varies depending on the dCas9:AtvT system used, and the excess may lead to an increase in non-specific targets (off-targets). Furthermore, the overabundance of sgRNAs and dCas9 can reach the saturation point and cause the attenuation of transcription due to the structural interference of physical binding of multiple dCas9-sgRNA complexes to the target DNA region (Piatek et al., 2015; Roca Paixão et al., 2019).

Table 2. Main transcriptional modulators used in the CRISPR / dCas system in plants.

Modulator	Type[(1)]	Source	Reference
VP64	AtvT	Artificial tetrameric repetition of the minimal activation domain of VP16 protein of Herpes Simplex Virus	Beerli et al. (1998)
EDLL	AtvT	Domain of the transcriptional activator AtERF98 (ethylene response factor) of *Arabidopsis thaliana*	Tiwari et al. (2012)
TAD	AtvT	*Xanthomonas campestris* Hax3 transcriptional activator DNA binding domain	Mahfouz et al. (2011)
TV	AtvT	6x TAD + 2x VP64	Li et al. (2017)
VPR	AtvT	Tripartite activator containing VP64, p65 (NF-kB trans-activator subunit) and Rta	Chavez et al. (2015)
HAT	AtvT	Catalytic nucleus of *A. thaliana* histone acetyltransferase 1	Roca Paixão et al. (2019)
p65-HSF	AtvT	NF-kB trans-activating subunit fused to the human heat-shock factor (HSP1) activation domain	Konermann et al. (2015)
SRDX	RepT	Domain of repression of the EAR motif (amphiphilic repression motif associated with ERF)	Ohta et al. (2001)
KYP	RepT	SET domain of the H3K9 KRYPTONITE (KYP) methyltransferase from *A. thaliana*	Jackson et al. (2002)
BRD	RepT	B3 repression domain of *A. thaliana* transcription factors	Ikeda e Ohme-Takagi (2009)
KRAB	RepT	Repressive domain of transcription factors that contain the Krüeppel zinc finger pattern of tetrapods	Mark et al. (1999)

[(1)] AtvT – Transcriptional Activator; RepT- Transcriptional Repressor

Another point to be considered is the location of sgRNAs. In general, sgRNAs targeting the region upstream of the TATA box and TSS correlate positively with

dCas9-mediated gene activation, probably due to the interference with the transcriptional initiation complex. When dCas9 is directed downstream or close to the TATA box, gene expression is negatively affected, probably through the physical blockage of transcription machinery by dCas9 (Deaner; Alper, 2017; Farzadfard et al., 2013). However, the optimal sgRNA distance upstream of the TSS for maximum gene activation can vary depending on the type of AtvT used (Piatek et al., 2015).

The choice of DNA strand (sense or antisense) also influences the efficiency of the method. Although there is no consensus, sgRNAs located on the sense strand and close to TSS are generally more efficient in activating transcription. Those located on the antisense strand can block RNA polymerase binding and initiation of transcription or lead to premature transcription termination (Howe et al., 2017; Piatek et al., 2015).

Thus, the increasing knowledge on the structure and functionality of the target promoter region (composition and role of cis-elements, recruitment of transcription factors, TSS position, among others) favors the success of dCas9 technology. Therefore, it is strongly recommended, as a first step, to test different sgRNAs, in different combinations, positions, and orientations in transient expression systems.

Another variation of this strategy is the fusion of multiple activators to the dCas9 sequence. In the dCas9: VP system, for example, the dCas9 protein was fused to six TAD modules followed by two VP64, which provided an expression increase of up to 190 times compared to the basal expression of the endogenous gene, both in monocotyledons and in dicotyledons (Figure 2C). The dCas9: VP system proved to be efficient in a multiplex system, promoting a simultaneous increase in the expression of three genes, and it was also effective in a DNA-free system, using complexes of ribonucleoproteins (RNP), but with a lower level of gene activation (Tables 2 and 3) (Li et al., 2017). However, the effects of fusing more than one activator to the dCas9 are not always additive. The dCas9: VP64: EDLL fusion, for example, resulted in a modest activation of endogenous genes, inferior to the classic dCas9: VP64 system, often considered weak and inefficient in plant cells (Table 3) (Li et al., 2017; Lowder et al., 2018). It must be taken into account that the fusion of an excessive number of activators can trigger the degradation of mRNAs due to the high number of repetitive sequences (Li et al., 2017).

An interesting observation is that there is a negative correlation between the endogenous basal expression of the gene and the magnitude of increased expression via dCas9: AtvT. Endogenous genes with lower basal expression are more likely to be strongly induced by this system, compared to those with higher basal transcription (Li et al., 2017; Lowder et al., 2018).

Table 3. Strategies for using the CRISPR/dCas9 system for the activation of gene expression in plants.

Plant species	Target-gene	CRISPR/dCas9 strategy	Transformation system	Activation level	Reference
A. thaliana	*AtPAP1*	dCas9:VP64 dCas9:VP64-EDLL dCas9:VP64-MS2-VP64 dCas9:VP64-MS2-EDLL dCas9:VP64- MS2-p65-HSF	Transgenic plants	7.0 x 4.0 x 45.0 x 30.0 x 3.0 x	Lowder et al. (2015) Lowder et al. (2018) Park et al. (2017)
	AtFIS2	dCas9:VP64 dCas9:VP64-EDLL dCas9:VP64-MS2-VP64 dCas9:VP64-MS2-EDLL	Transgenic plants	200.0 x 2.5 x 1500.0 x 30.0 x	Lowder et al. (2018)
	miR319	dCas9-VP64 dCas9:VP64-MS2-VP64 dCas9:VP64-MS2-VP64 + 3sgRNA	Transgenic plants	7.5 x 6.0 x 2.5 x	Lowder et al. (2015); Lowder et al. (2018)
	AtULC1	dCas9:VP64-MS2-VP64 + 3 sgRNAs	Transgenic plants	40.0 x	Lowder et al. (2018)
	AtWRKY	dCas9:TV dCas9:TV + RNP	Transgenic plants Protoplast/PEG	139.0 x 11.7 x	Li et al. (2017)
	pWRKY::luciferase	dCas9:TV dCas9:VP64-MS2-TV dCas9:VP64-	Protoplast/PEG	510.0 x 170.0 x 6.7 x	Li et al. (2017)
	AtRLP23	dCas9:TV dCas9:TV dCas9:TV + RNP	Transgenic plants Protoplast/PEG Protoplast/PEG	30.0 x 44.0 x 9.3 x	Li et al. (2017)

To be continued...

Table 3. Continued.

Plant species	Target-gene	CRISPR/dCas9 strategy	Transformation system	Activation level	Reference
A. thaliana	*AtWRKY* *AtRLP23* *AtCDG1*	dCas9:TV multiplex	Protoplast/PEG	80.0 x 37.0 x 192.0 x	Li et al. (2017)
	AtAREB1	dCas9:HAT	Transgenic plants	1.7 x	Roca Paixão et al. (2019)
	AtFWA *AtAPETALA3* *AtCLAVATA3* *Evadé transposon*	dCas9:SunTag-VP64	Transgenic plants	1.5 x 500.0 x 100.0 x 4000.0 x	Papikian et al. (2019)
	AtAVP1	dCas9:VP64-MS2-p65-HSF SAM	Transgenic plants	5.0 x	Park et al. (2017)
	AtFT	Cas9-MS2-VP64	Transgenic plants	30.0 x	Lee et al. (2019)
N. benthamiana	*NbPDS*	dCas9:EDLL + 3 sgRNAs dCas9:TAD + 3 sgRNAs	Agroinfiltration	3.4 x 4.0 x	Piatek et al. (2015)
	pNOS::luciferase	dCas9:VP64 dCas9:EDLL dCas9:EDLL-MS2-VP64	Agroinfiltration	3.0 x 3.0 x 4.0 x	Vazquez-Vilar et al. (2016)
		dCas9:SunTag-VP64 dCas9:SunTag -EDLL		3.5 x 3.0 x	Selma et al. (2019)

To be continued...

Table 3. Continued.

Plant species	Target-gene	CRISPR/dCas9 strategy	Transformation system	Activation level	Reference
O. sativa	*OsGW7*	dCas9:TV	Protoplast/PEG	79.0 x	Li et al. (2017)
	OsER1	dCas9:TV dCas9:TV + RNP	Protoplast/PEG	62.0 x 13.0 x	Li et al. (2017)
	Os03g01240	dCas9:VP64 dCas9:VP64-MS2-VP64 + 2 sgRNAs	Protoplast/PEG	2.0 x 3.0 x	Lowder et al. (2018)
	Os04g39780	dCas9:VP64 dCas9:VP64-MS2-VP64 + 2 sgRNAs	Protoplast/PEG	0.0 4.0 x	Lowder et al. (2018)
	Os03g01240 *Os04g39780* *Os11g35410*	dCas9:VP64 multiplex	Protoplast/PEG	2.0 x 2.0 x 2.0 x	Lowder et al. (2018)
	Os03g01240 *Os04g39780* *Os11g35410*	dCas9:VP64-MS2-VP64 multiplex	Protoplast/PEG	3.0 x 6.5 x 2.5 x	Lowder et al. (2018)

Modification of the sgRNA structure for activator recruitment

Even with all the improvements, the strategy of fusing ModT to dCas9 may not be able to modulate the expression levels of target genes satisfactorily. Also, the one-to-one ratio of dCas9-sgRNA in classical systems restricts the programming of multiple gene circuits to only one direction of regulation: repression or activation. This does not correspond to the complexity and sophistication underlying native transcription networks. However, similarly to ModTs and dCas9 fusions, the engineering of the sgRNA structure proved to be a modular and adjustable platform to diversify the positions to be recognized in the target genomes and the direction of transcriptional regulation mediated by the CRISPR system. This strategy is based on the modification of the sgRNA structure by the addition of RNA aptamers, making it capable of recruiting adapter proteins fused to ModTs to the target promoter region. Thus, a greater number of ModT in the target site can increase the magnitude of the activation levels (Konermann et al., 2015; Zalatan et al., 2015).

In the first systems using engineered sgRNAs, the sequence of RNA aptamers was inserted at its 3'end. These aptamers, such as MS2, fused to sgRNA scaffold, could interact with specific proteins such as MCP, fused with ModTs, allowing the specificity of the target in the genome, together with the modulating function. This system was called scaffold RNA (scRNA) (Figure 2E) (Jensen et al., 2017; Kiani et al., 2015; Zalatan et al., 2015).

Later, in the development of a second system, it was shown that the scaffold RNA region assumes a specific secondary structure, formed by a tetraloop and 3 stem-loops, important for the sgRNA anchoring in Cas9/dCas9 (Figure 1). It was evident that the tetraloop and stem-loop 2 protrude out of the dCas9-sgRNA complex, with the four distal base pairs of each stem completely free of interactions with the protein's amino acid side chains. Functionality studies have shown that substitutions and deletions in the tetraloop and stem-loop 2 regions do not affect the target DNA recognition function by dCas9 and that they can tolerate the addition of RNA aptamers (for example, MS2), which in turn, are recognized by specific proteins (for example, MCP). These specific proteins recruit a greater number of ModT to the target region in the genome, recognized by the dCas9-sgRNA complex. The technology of combining ModTs fused to dCas9, and the secondary structure of sgRNA was named as Synergistic Activation Mediator (SAM). A similar system adapted for plants using AtvT VP64 was called CRISPR-Act2.0 (Figure 2 F and G) (Konermann et al., 2015; Li et al., 2017; Lowder et al., 2018; Nishimasu et al., 2014).

In plants, these systems have already been proven effective with the use of various combinations of AtvT (Table 2) (Lee et al., 2019; Lowder et al., 2018; Selma et al., 2019). The CRISPR-Act2.0 system showed superior ability to activate three

genes simultaneously in a multiplex scheme when compared to the classic dCas9: VP64 system (Table 3) (Lowder et al., 2018). However, the efficiency of the system can vary depending on the target gene, type of activators used, and the number of sgRNAs. For example, at the same time that the dCas9: VP64-MS2: VP64 system was able to activate the expression of the *A. thaliana AtFIS2* gene up to 1,500 times, a similar strategy, using dCas9: VP64-MS2: EDLL for activation of this same gene, presented a lower result than that obtained by the classic dCas9 system: VP64 (Table 3) (Lowder et al., 2018). These same authors launched the hypothesis that, for many genes, there must be a maximum limit for increasing its expression. Higher levels can trigger post-transcriptional gene silencing through the production of small interfering RNAs. Thus, strategies that should optimize the gene activation system via CRISPR/dCas9 do not always result in the desired effect, as exemplified in the activation of miR319 in *A. thaliana* (Table 3) (Lowder et al., 2018, 2015).

SUperNova Tag (SunTag) system for gene activation

This strategy was first applied in mammalian cells in an attempt to optimize the increase in gene expression by enhancing the number of activators in the target promoter region. In this method, dCas9 is fused with tandem repeats of the GCN4 peptide, which binds to an antibody (scFv) fused to ModT to (Figure 2H) (Tanenbaum et al., 2014). In plants, this strategy was adapted using the AtvT VP64 to increase in the transcription of endogenous genes and transposable elements up to 4,000 times (Table 3) (Papikian et al., 2019). A variation of this technique is the replacement of VP64 by the catalytic domain of methyltransferases to promote methylation of promoter regions and control of the transcription. Thus, the dCas9-SunTag system can be a potent tool for site-specific epigenome modulation. It is a highly specific system, capable of promoting transcriptional activation of genes, both in euchromatin and heterochromatin, and with heritable effects over multiple generations, even in the absence of the transgene (Papikian et al., 2019). In addition, the use of the dCas9-SunTag system associated with a catalytic domain of methyltransferases significantly minimizes generalized non-specific methylation, common when these domains are directly fused to dCas9 (Pflueger et al., 2018).

Reduction of gene expression (CRISPRi)

The dCas9, when directed by the sgRNA to a promoter region, in the vicinity of TATA-Box or TSS or a coding region of a gene, can physically block the assembly of the transcription complex and the initiation and elongation of transcription (Figure 2A). The CRISPRi is an efficient strategy of directed negative regulation

of gene expression, with reversible effects, low incidence of off-targets, and the possibility of being applied to multigene analyses (Qi et al., 2013). It is a functional platform in eukaryotes and an alternative strategy for modulating gene expression in prokaryotes, which does not have RNA interference (RNAi) machinery.

In plants, the efficiency of CRISPRi was demonstrated in transient transformation systems by the repression of a reporter gene in up to 80% (Table 4) (Piatek et al., 2015; Vazquez-Vilar et al., 2016). These latter authors still demonstrated the effects of dCas9-sgRNA complex location at promoter region on gene repression. They showed that gene expression was lower for sgRNAs located nearby TSS. They also observed a modest and non-significant additive effect of the simultaneous use of two or three sgRNAs on reducing gene expression (Table 4) (Vazquez-Vilar et al., 2016).

Fusion of transcriptional repressors to dCas9 (dCas9:RepT)

In this strategy, the dCas9 sequence is fused in tandem with one or more transcriptional repressors (RepTs) (Table 2), similarly to previously described for AtvT. In plants, the RepT mainly used so far is the SRDX, which is compatible with a multiplex system and reached maximum efficiency when three domains were fused to dCas9, with a reduction by up to 80% in gene expression (Lowder et al., 2018, 2015; Sarasua, 2020; Vazquez-Vilar et al., 2016). Although KRAB RepT is widely used in animal models, it was unstable and inefficient for gene expression modulation in plants, even when tested in different contexts, varying the type of Cas nuclease and the location of sgRNA in relation to the TSS (Table 4). KRAB is an exclusive motif of tetrapods, a fact that would limit its ability to recruit chromatin-modulation proteins in plant cells (Sarasua, 2020).

Interestingly, the dCas9: SRDX system was able to mask the effects on gene activation of dCas9: EDLL and dCas9: TAD so that gene expression remained unchanged when these ModTs were used jointly (Piatek et al., 2015). Although the dCas9: RepT strategy is a useful tool for modulating gene expression, in some cases, its efficiency is similar to or less than that obtained with the classical CRISPRi system without RepT (Table 4) (Sarasua, 2020; Vazquez-Vilar et al., 2016).

Another point to highlight is that the use of a different type of dCas protein can improve the efficiency of dCas: RepT systems. Tang et al. (2017) tested the dCas12a nuclease (also called dCpf1) fused with three domains of SRDX RepT [dCas12a: 3x (SRDX)] and observed a reduction of up to 90% in the transcription of the *miR159B* of *A. thaliana*. In a similar system, using dCas9: 3x (SRDX), this value was approximately 70% (Table 4) (Lowder et al., 2018). Sarasua (2020) also reported the superiority of dCas12a nuclease compared to dCas9 when both were

fused to the SRDX and BRD RepTs to reduce the expression of a reporter gene in a transient transformation system. Moreover, the use of two sgRNAs in these systems leads to a variable significant additive effect. Superior results were detected only for dCas9: SRDX combination, with no difference when one or two sgRNAs were used in dCas9: BRD and dCas12a: SRDX. For dCas12a: BRD, the use of two sgRNAs was less efficient (Table 4). The superior efficiency of dCas12a in suppressing gene transcription, compared to dCas9, may be related to the size of these nucleases. While dCas12a has approximately 1.250 amino acids, dCas9 has 1.368, which would facilitate dCas12a access to the target genomic region (Sarasua, 2020).

Modification of the sgRNAs structure for the recruitment of repressors

Despite being a widely used strategy for activating gene transcription, CRISPR-Act2.0 methodology has been used only once so far to suppress gene transcription in plants (Lee et al., 2019). The authors reported the failure of the method using SRDX RepT, but it was possible to achieve median gene repression levels with another RepT (KYP) (Table 4).

SUperNova Tag (SunTag) system for gene repression

This strategy was extensively explored in a detailed study by Sarasua (2020), who varied several parameters of the method, such as sgRNA location in relation to TSS, the number of sgRNAs (1 or 2), type of Cas nuclease (dCas9 or dCas12a), type of RepT (SRDX or BRD), and size of spacers (5 or 22 amino acids) between the GCN4 epitopes, using a transient system for monitoring the expression of the luciferase reporter gene guided by the nopaline promoter (*pNOS*). The system using dCas12a nuclease gave higher efficiency (Table 4), probably due to dCas12a smaller size, compared to dCas9, as previously discussed. The use of larger spacers of 22 amino acids between epitopes promoted better results since smaller spacers can offer steric impediment to antibody binding. In general, the SunTag system induced gene repression in a similar way as dCas: RepT strategy, showing to be slightly superior when two sgRNAs were used with dCas12a (dCas12a: Suntag - SRDX / BRD + 2 sgRNA). There was no significant difference between the two RepTs tested (SRDX and BRD) (Table 4).

Table 4. Strategies for using CRISPR / dCas9 system to suppress gene expression in plants.

Plant species	Target gene	CRISPR/dCas9 strategy	Repression level	Transformation system	Reference
A. thaliana	*AtCSTF64*	dCas9:3x (SRDX) + 3 sgRNAs	60	Transgenic plants	Lowder et al. (2015)
	miR159A *miR159B*	dCas9: 3x (SRDX) multiplex	80 70	Transgenic plants	Lowder et al. (2015)
	miR159B	dCas12a:3x (SRDX)	90	Transgenic plants	Tang et al. (2017)
	AtFT	dCas9-MS2-SRDX dCas9-MS2-KYP	0 40	Transgenic plants	Lee et al. (2019)
N. benthamiana	*NbPDS*	dCas9 dCas9:SRDX	20 50	Agroinfiltration	Piatek et al. (2015)
	pNOS: luciferase	dCas9 dCas9 + 2 sgRNA dCas9 + 3 sgRNAs dCas9:SRDX dCas9:SRDX + 3 sgRNAs dCas9:BRD dCas9:BRD + 3 sgRNAs	80 85 85 50 50 40 70	Agroinfiltration	VazquezVilar et al. (2016)
		dCas9:BRD dCas9:BRD + 2 sgRNAs dCas9:SunTag – BRD dCas9:SunTag – BRB + 2 sgRNA dCas9:SRDX dCas9:SRDX + 2sgRNAs dCas9:SunTag - SRDX dCas9:KRAB dCas12a:BRD dCas12a:BRD + 2sg RNA dCas12a:SunTag – BRD dCas12a:SunTag – BRD + 2sgRNAs dCas12a:SRDX dCas12a:SRDX + 2sg RNA dCas12a:SunTag – SRDX dCas12a:SunTag – SRDX+ 2 sgRNA dCas12a:KRAB	30 30 30 40 40 65 40 30 80 60 80 90 80 80 80 90 30	Agroinfiltration	Sasuna (2020)

Experimental design: important aspects to have in mind

As mentioned previously, gene expression modulation mediated by CRISPR technology contributes significantly to the understanding and orthogonal control of transcriptional and post-transcriptional regulations. Compared to other methods, such as RNAi technology, CRISPR technology is easy and offers RNA-mediated targeting of an individual or multiple genes, making it a powerful tool for the study of multifactorial native transcriptional regulation in diverse organisms, including plants (Deaner et al., 2017; Jensen et al., 2017; Zalatan et al., 2015). Thus, unique sets of multiplex experiments can be easily applied, and the use of more than one sgRNA allows rapid progression through interactive cycles of genetic engineering. By varying the expression of a single gene, it is possible to rapidly evaluate the combinatorial effects on a network of gene expressions to identify previously unknown targets.

There are many aspects regarding the experimental design that must be taken into account when using the CRISPR system for base editing or modulation of gene transcription. Thus, two important points are addressed in the following tutorials: (i) sgRNA design; and (ii) transformation cassette design (plasmid vectors).

sgRNAs design

Theoretical basis for choosing the best software/platform

Two main challenges for the biotechnological application of the CRISPR/Cas system are associated with the correct design of sgRNA: (i) reliable prediction of off-targets; and (ii) efficiency in recognition of on-targets (Xu et al., 2015; Zhang et al., 2015).

The parameters for a successful sgRNA design in CRISPR/dCas9 technology are the same as for native Cas proteins, even considering that DNA cleavage is not an important factor in CRISPR/dCas9. Thus, three main steps must be followed in sgRNA design: (i) identification of PAM sequences in the target region; (ii) selection of a 20 nucleotides sequence upstream from PAM, excluding it, named as spacer; and (iii) attachment of the spacer to the 5′ end of scaffold RNA, for the correct anchoring of sgRNA to Cas9/dCas9. A successful experimental design, especially of the sgRNA, includes the prediction of on-targets and off-targets (Doench et al., 2016; Tang et al., 2018).

For sgRNAs used with Cas9, guanines are preferred at positions 1 and 2 before the PAM sequence (NGG), while the presence of thymines at positions +4 and -4, flanking PAM, interferes negatively with nuclease cleavage efficiency (Wang et al., 2014; Wu et al., 2014). Also, the regions downstream from PAM can affect cleavage efficiency and sgRNA specificity, while the upstream sequences generally have no significant effect on these parameters (Doench et al., 2014). Cytosine is preferred at Cas9 cleavage site (position -3 next to PAM) (Cong et al., 2013; Xu et al., 2015), and a high GC content in the region comprising bases 4-13, downstream from PAM sequence, favors cleavage efficiency and sgRNA specificity. The sgRNA size is also a crucial factor for the specificity of the Cas9 target, with the minimum sgRNA size being 17 nucleotides for knockout gene experiments (Fu et al., 2014). For dCas9-mediated transcriptional and epigenetic modulation, the best sgRNA size is 20 nucleotides since a reduction in its size decreases the effectiveness of dCas9-mediated transcriptional regulation (Kiani et al., 2015; Smith et al., 2016). There is a certain degree of flexibility in the design of distal positions of sgRNAs used to recognize gene promoters in regions rich in nucleosomes or upstream activation sequences (UASs) (Kiani et al., 2015; Smith et al., 2016).

Powerful computational tools are continually developed and updated to facilitate in silico sgRNA design for biotechnological purposes (Chuai et al., 2017; Graham; Root, 2015; Wilson et al., 2018). However, how to choose the best software/platform for sgRNA design, given the availability of numerous options? There are several models for predicting sgRNA efficiency in silico, based on experimental evidence. To facilitate this choice, it is necessary to know what each software offers, highlighting its positive points and limitations.

Rule Set 1 is a predictive model based on a scoring system associated with an in silico learning method, which classifies the data in a linear and generalized way, in order to estimate the cutting efficiency of Cas9, with high correlations with experimental results (Doench et al., 2014; Liu et al., 2020). Independent studies contributed to improve the accuracy of this model, including more data sets to build a new model, called Rule Set 2, in which nucleotide analyses, regardless of the position and location of the sgRNA within the target sequence, can improve the prediction of off-targets and on-targets (Doench et al., 2016; Liu et al., 2020).

The software and algorithms developed to predict the specificity of sgRNAs are based on the determination of off-targets and have different types of data input (Table 5). All are built on two main methods: (i) alignment method, mainly used to predict all off-*targets in silico*; and (ii) a score method, in which sgRNAs are scored and classified according to off-targets identified through an alignment matrix.

In theory, putative off-targets can be identified by aligning the sgRNA sequences to a reference genome through sequence similarity. The tools Bowtie (Langmead et al., 2009) and BWA (Li and Durbin, 2009) allow the detection of off-targets but fail to identify small PAMs, since they were developed to align reads from Next Generation Sequencing (NSG) (Haeussler et al., 2016; Wilson et al., 2018). Another point is that these tools allow a limited number of mismatches, restricting the identification of putative off-targets (Langmead and Salzberg, 2012). However, not all nucleotide mismatches have the same effect on off-target cleavage, and the alignment-based prediction always identifies redundant off-targets, of which many are false positives. Considering these factors, CasOFFinder and FlashFry softwares, among those based exclusively on alignments, would be the best options to identify possible off-targets for any Cas nuclease (Liu et al., 2020). FlashFry software stands out by its high processing speed and easy-to-understand output.

Table 5. Most used online tools for sgRNA design [1]

Tool	Website	Species	Input data
Benchiling	https://www.benchling.com	Diverse	DNA sequence; gene ID
CCTop	https://crispr.cos.uni-heidelberg.de	Diverse	DNA sequence
CHOPCHOP	https://chopchop.cbu.uib.no	Diverse	DNA sequence; TSS; location or gene ID
CRISPETa	http://crispeta.crg.eu	Model plants	DNA sequence; location or gene ID
CRISPOR	http://crispor.tefor.net	Diverse	DNA sequence; location or gene ID
CRISPR-ERA	http://CRISPR-ERA.stanford.edu	Diverse	DNA sequence; TSS; ID do gene
CRISPR-GE	http://skl.scau.edu.cn	Plants	DNA sequence; location or gene ID
CRISPR-P 2.0	http://crispr.hzau.edu.cn/CRISPR2	Plants	DNA sequence; location or gene ID
CRISPR-PLANT	https://www.genome.arizona.edu/crispr2	Plants	DNA sequence; location or gene ID
CRISPR RGEN Tools	http://www.rgenome.net	Diverse	DNA sequence; location or gene ID sgRNA

To be continued...

Table 5. Continued.

Tool	Website	Species	Input data
CRISPRscan	https://www.crisprscan.org	Model plants	DNA sequence; location or gene ID
E-CRISP	http://www.e-crisp.org/E-CRISP	Diverse	DNA sequence; gene ID
EuPaGDT	http://grna.ctegd.uga.edu	Eukaryotic Pathogens	DNA sequence
FlyCRISPR	https://flycrispr.org/	Insects	DNA sequence
Yeastriction	http://yeastriction.tnw.tudelft.nl/#!/	*Saccharomyces cerevisiae*	Gene ID

(1) The information presented here is subject to website updates.

To improve off-target detection, the score-based method was developed using two approaches: (i) hypothesis-oriented, in which off-targets are scored based on the contributions of genome context specific factors to sgRNA specificity; and (ii) learning approach, in which sgRNAs are scored and predicted based on a training model, which considers the different factors that affect specificity.

In the hypothesis-guided approach, the MIT score (Hsu-Zhang) evaluates mismatches'contributions in different positions along sgRNA sequence using a weight matrix to identify possible off-targets. The MIT score is included in sgRNA design tools, such as CHOPCHOP (Labun et al., 2016) and CRISPOR (Haeussler et al., 2016). The Cutting Frequency Determination (CFD) is another popular score system for off-target prediction, which is taken into account the presence of non-canonical PAMs (NAG, NCG, and NGA), indels, and mismatches in the target sequences. CFD score shows a better performance than MIT score and is used by the softwares GUIDE-Seq (Tsai et al., 2015), CRISPRscan (Moreno-Mateos et al., 2015), GuideScan (Perez et al., 2017), CRISPOR (Haeussler et al., 2016), and Benchling (Doench et al., 2016).

The prediction of sgRNA specificity based on the structural characteristics of the Cas9-sgRNA complex proved to be superior to that based only on sequence characteristics. The CRISPRoff (Alkan et al., 2018) and uCRISPR (Zhang et al., 2019a) tools use energetic properties to predict off-targets and were more accurate in predicting off-targets when compared with tools that use other methods, such as MIT and CFD. However, none of them has been evaluated in large-scale experiments, and care should be taken when using them.

Empirical algorithms may fail in predicting off-targets since they consider few data sources. On the other hand, learning approaches build complex models using combinations of numerous resources. The CRISPR Target Assessment

(CRISTA) algorithm uses BWA as an off-target search tool and implements several parameters (PAM, nucleotide composition, GC content, chromatin structure, DNA methylation, secondary sgRNA structure, etc.) to predict cleavage (Abadi et al., 2017). CRISTA performs better than MIT and CFD. DeepCRISPR is a new innovative computational platform that unifies sgRNA on-target and off-target site prediction into one framework with deep learning (Chuai et al., 2018). This tool automates the identification of sequence and epigenetic features that may affect sgRNA efficacy. It was able to overcome other available tools for predicting off-targets (Doench et al., 2016; Hsu et al., 2013; Liu et al., 2020; Singh et al., 2015; Stemmer et al., 2015).

Many online tools combine some of the score methods mentioned above and even develop a unique evaluation system to provide alternative options for users, such as CHOPCHOP (Labun et al., 2019, 2016; Montague et al., 2014) and CRISPRscan (Moreno-Mateos et al., 2015) (Table 5). E-CRISP uses an exclusive score system denominated SAE (Specificity, Annotation, Efficacy) to determine the quality of sgRNAs. CCTop empirically assigns scores for the prediction of off-targets (Stemmer et al., 2015). CRISPOR (Haeussler et al., 2016) is a versatile platform that classifies sgRNAs according to different score systems to prediction-target activity and putative off-targets in the genome of interest. It also shows a list of PAM sequences to be selected by the user. The CRISPR RGEN platform (Sullenger, 2020) provides several computational tools, as well as information for sgRNA design for several Cas types. The platform has nine useful tools, including Cas-Designer (Park et al., 2015) and the BE-Designer (Hwang et al., 2018). Compared to other tools, Cas-Designer allows for mismatches when detecting DNA/RNA hybrids off-targets. This detection method is faster than others due to the association with the CasOFFinder algorithm. Sequences, genomic coordinates, and FASTA files are allowed as inputs. More than 350 genomes and 20 types of PAM sequences are available to users. The output results include on-targets, out-of-frame score, calculated by micro-homology, incompatibility number (0–2), and off-target sequences with mismatches of up to 1 base pair. On/off-targets are redirected to the Ensembl[1] genome browser for further evaluation. Unlike Cas-Designer, BE-Designer is used mainly for base editing. This tool offers four methods of base editing and the editing region is adjustable. The CRISPR-ERA (Liu et al., 2015) and CHOPCHOP v3 (Labun et al., 2019) permit sgRNA design for dCas9 system, and CHOPCHOP includes 30 different plant species. CRISPETa (Pulido-Quetglas et al., 2017) is used mainly for genome deletion using two sgRNAs.

The appropriate tool's choice should be made with caution since there are tools developed specifically for some organisms or cell types. For example, Yeastriction

[1] Available at: https://chopchop.cbu.uib.no/

(Mans et al., 2015) is developed for yeasts, FlyCRISPR (Gratz et al., 2014) for *Drosophila* and other insects, and EuPaGDT (Peng; Tarleton, 2015) for eukaryotic pathogens. CRISPR-P/CRISPR-P 2.0 (Lei et al., 2014; Liu et al., 2017), CRISPR-PLANT (Minkenberg et al., 2019), and CRISPR-GE (Xie et al., 2017) are applied for plant genome. CRISPR-P 2.0 is one of the most used platforms for genome editing in plants nowadays and allows sgRNA design and prediction of on-targets and off-targets. It supports the design of sgRNAs for 49 plant genomes, including species for which there is a well-assembled genome, and also permits sgRNA identification using user-supplied sequences as reference. It applies a modified score system to predict off-targets and sgRNA targeting efficiency using Cas9 and Cpf1 (Zetsche et al., 2015). This platform informs the GC's content, restriction endonuclease sites, microhomology of the sequence flanking the region to be edited, and the secondary structure of the sgRNA.

We chose CHOPCHOP[2] v3 (Labun et al., 2019) to present a tutorial on sgRNA design using dCas9, considering that this software contains the specific function for its nuclease, provides genomic data of more than 250 species, and undergoes constant updates. If you have previous in silico/ in vivo information about the promoter region of your interest, you can use other software, such as CRISPR RGEN or CRISPR-P 2.0.

Tutorial 1: sgRNA design for transcriptional modulation via dCas9

Step 1: Access URL https://chopchop.cbu.uib.no

Step 2: Choice of species of interest (Section *In*). The eukaryotic and prokaryotic species available so far are divided into the following categories: Arthropoda (37), Bacteria (58), Chordata (46), Cnidaria (1), Ctenophora (1), Echinodermata (2), Fungi (30), Mollusca (2), Nematoda (2), Others (12), Parasites (11), Plants (30), and Virus (20). In this tutorial, the plant species *Glycine max* (GCA_000004515.3) was chosen. You can request to the software developers the insertion of the genome sequence of a non-provided species. For that, you have to supply a FASTA file with the genome assembly, as well as a valid GFF3 file referring to its functional annotation (validated by gff3ToGenePred program, available on CHOPCHOP platform) (Figure 3). The ID assigned in the GFF3 file must match the identifier you want to use to search for the genes. Chromosome names must be identical in GFF3 and FASTA files, and gene coordinates must correspond to the FASTA file. Whether the requirements have been attended, a request must be sent by email (available in the "Add new species" section, below the place for selecting the species of interest). It is worth mentioning that the platform's maintainers perform the insertion of a new genome. The user is not allowed to insert the target genome manually.

[2] Available at: https://chopchop.cbu.uib.no/

Step 3: Choice of the target gene (section "*Target*"). The gene ID chosen for this tutorial is Glyma_09G153900. The target gene ID is inserted in this field and must correspond precisely with the genome deposed in the CHOPCHOP database. If there is no match, the platform itself will send the error message "Error status: 501" and the correct way to write the ID for the selected is given. This field acts synchronously with the gray "Paste Target" icon. This tool allows to find the gene according to the ID (RefSeq or ENSEMBL format), gene name, or gene coordinates in the genome (for example, chrX: 15,560,138-15,602,945). When clicking on the gray "Paste Target" icon, the user changes the icon's function to the "Gene Target" option, where it is possible to place the sequence of nucleotides to be analyzed. For dCas9-based strategy, sequence self-placement is not widely used since, in most cases, the genomic context of the promoter region of the target gene is lost, which is essential for transcriptional modulation via dCas9 (Figure 3).

Step 4: Choice of nuclease (section "*Using*"). In this field, the platform has so far made available five different types of nucleases for genome editing: (i) CRISPR/Cas9; (ii) CRISP/Cas9 nickase; (iii) CRISPR/Cpf1 (CasX); (iv) CRISPR/Cas13 (c2c2); and (v) TALEN. For this tutorial, the selected nuclease is "CRISPR/Cas9" (Figure 3).

Step 5: Choice of strategy (Section "*For*"). In this field, the approach or context in which the sgRNAs will be applied is selected (for this tutorial, the option "Activation" was selected (Figure 3):

a) *Knockout* (for mutations intending changes in the Open Reading Frame (ORF) of the target gene; available for almost all nucleases on the platform, exceptionally for CRISPR/Cas13): for applications with Cas9, it is possible to predict the rate of reading frame changes provided by each sgRNA. Some recommendations are suggested: (i) use the traditional Cas9 (20 nucleotides with PAM sequence 5'-NGG-3'); (ii) make sure that sgRNA has the fewest number of putative off-targets, preferably none; (iii) ensure that the selected sgRNA can recognize all isoforms of the target gene (not available for all genomes); and (iv) select sgRNAs downstream of any ATG in a frame within the ORF-structure (green boxes in the blue coding-region) to avoid the expression of truncated proteins.

b) *Knock*-down (for mRNA editing with CRISPR/Cas13; available only for humans and mice): in this mode, it is possible to search for off-targets in transcriptomes. It is possible to select the Cas13 system in "*Options*" (after selecting "CRISPR/Cas13" in the "*For*" section) and specify the protospacer flanking site (PFS) and the size of the sgRNA.

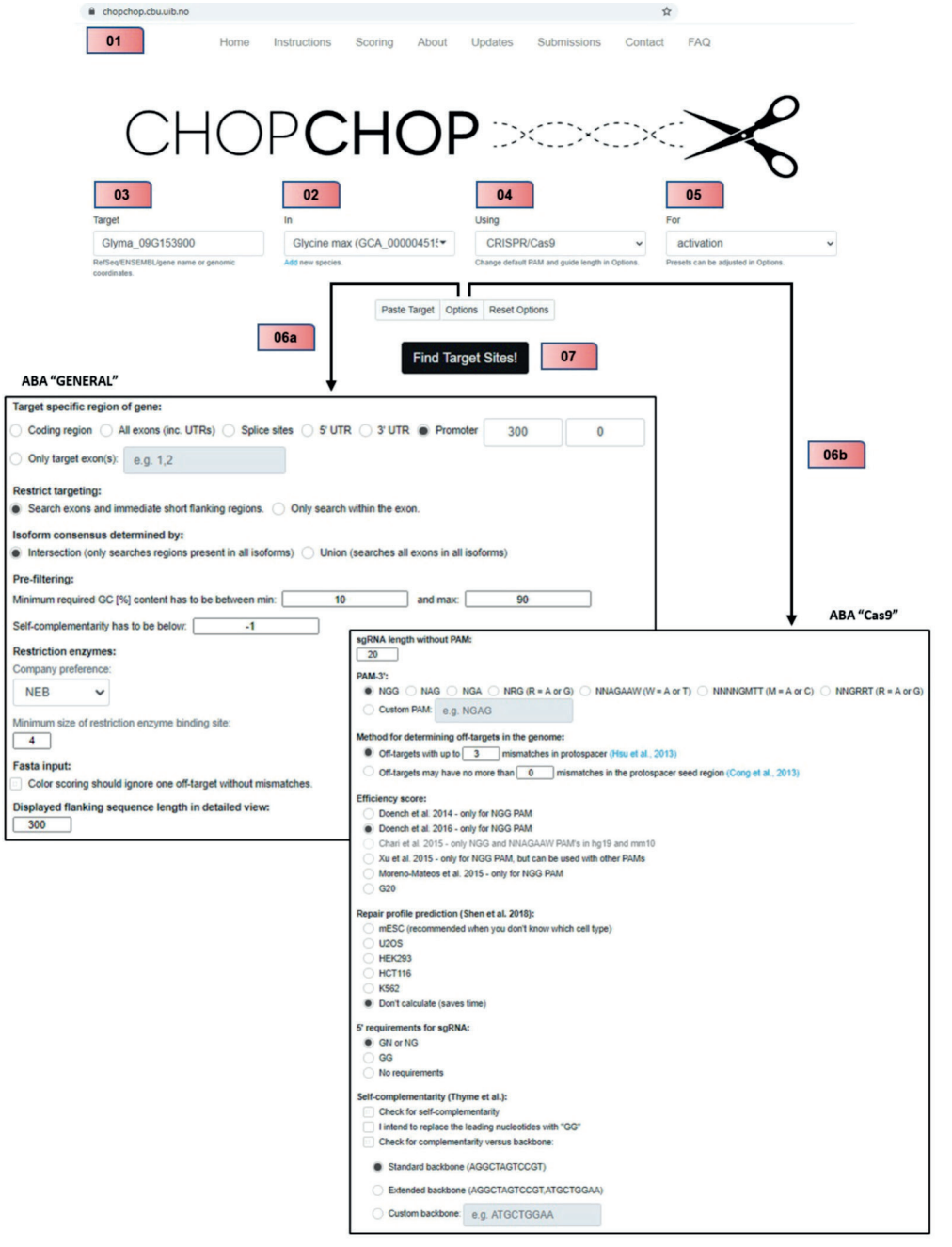


Figure 3. CHOPCHOP v3 – Parameters. Introduction of Tutorial 01 workflow for choosing the main parameters of sgRNAs design aiming at activation of a target gene from *G. max* (ID: Glyma_09G153900) via dCas9. The numbering shown in red rectangles corresponds to steps 1 to 7 of Tutorial 1.

c) *Knock-in* (for inserting DNA sequences in a region of interest; unavailable for CRISPR/Cpf1, CRISPR/Cas13, and TALEN): it is highly recommended to be familiarized with the different types of knock-ins available for choosing the most appropriate method (Nami et al., 2018). Knock-in by CRISPR technology allows site-specific genetic engineering through nuclease-induced double-stranded DNA (DSB) breakdown and, subsequently, DSB repair via cell repair pathways: non-homologous final junction (NHEJ) and homology-directed repair (HDR). An important aspect to consider when deciding the type of technique to use is whether the modification has a dividing- or non-dividing-cell as a target. HDR is restricted to the S and G2 phases of the cell cycle and is not suitable for non-dividing-cells. The sequences for the homology arms are provided on the sgRNA detailed results page. It is possible to adjust the position of the microhomology arms to the 5' end of the sgRNA (default: nucleotide -3 from the PAM sequence) and specify the size of the arm. It is recommended to check the complementarity between the inserted sequence and its microhomology arms. In the default settings for knock-in, homology arms up to 800 bp can be obtained. If a larger homology arm is desired (up to 2,000 base pairs), it is necessary to change the size of the flanking sequence shown on the home page ("*Options*" > "*General*" > "*Displayed flanking sequence length in detailed view*").

d) *Activation/Repression* (for sgRNAs targeting promoter regions; available for CRISPR/Cas9 and CRISPR/Cpf1 nucleases): in this option, the default for sgRNA design is: (i) activation mode: 300 bp upstream from the TSS; (ii) repression mode: 200 base pairs downstream and upstream from the TSS. It is recommended to use more than one sgRNA for these applications.

e) *Nanopore Enrichment*: used to enrich specific DNA fragments for sequencing through the Oxford Nanopore platform. It is only available for CRISPR/Cas9 and CRISPR/Cpf1 nucleases.

Step 6: Choice of advanced options (*Options section*). This section is found in the gray icon in the center of the home page and is extremely relevant to refine the results according to the characteristics of the target genome and sequence (Figure 3). When clicking on it, with the CRISPR/Cas9 nuclease previously selected, new options appear in three tabs: (i) "*General*", in which it is possible to adjust general parameters for the CRISPR technology; (ii) "*Cas9*" option, to adjust specific parameters for the Cas9 nuclease; and (iii) "*Primers*", to adjust parameters related to the design of primers that will be used in PCR experiments for mutations detection (basically used for knockout experiments). For dCas9 experiments, only parameters displayed in the "*General*" and "*Cas9*" tabs will be adjusted:

a) "*General*" tab:

- *Target a specific region of the gene:* it is possible to specify the region where the sgRNAs will be designed: (i) only the coding region (default); (ii) the entire exonic sequence (including 5'and 3'-UTR); (iii) splicing sites; (iv) only 5'-UTR; (v) only 3'-UTR; (vi) the promoter (it is possible to specify how many bp upstream and downstream from the TSS will be analyzed); or (vii) a specific exon or a specific subset of exons. If you want to target an intron, specify its genomic coordinates (maximum size, 40.000 base pairs).
- *Restrict targeting:* when searching for on-targets in a region of interest, in default mode, CHOPCHOP allows the sgRNA to link only outside that region, ensuring a cut on the target region. It is possible to disable this option in experiments with dCas9.
- *Isoform Consensus*: when analyzing genes with multiple isoforms, the "*Intersection*" or "*Union*" modes can be activated. In "*Intersection*" mode, CHOPCHOP searches only for sgRNAs present in all isoforms. In "*Union*" mode, CHOPCHOP looks for sgRNAs in all exons of all isoforms. Therefore, you can use this mode to target one or more specific isoforms. It is possible to disable this option in experiments with dCas9.
- *Pre-filtering:* it is possible to pre-filter sgRNAs based on GC content and self-complementarity score (set to -1 to disable, which is not recommended). sgRNAs that do not attend these requirements will not be reported in the final result table.
- *Restriction enzymes:* mutagenesis can be assessed by digestion with restriction enzymes. CHOPCHOP displays restriction sites at the target site and allows you to restrict your search to enzymes distributed by specific companies (NEB™, Fermentas™, Promega™, Roche™, Sigma™, and Stratagene™), as well as by the number of bases recognized by the enzyme. This parameter is not relevant in dCas9-mediated transcriptional modulation experiments.

b) "*Cas9*" tab:

- *sgRNA length without PAM:* since some studies suggest that truncated sgRNAs may show a greater cleavage efficiency, it is possible to design them shorter than 20 nucleotides (default), which is not recommended for experiments with dCas9.
- *PAM-3' sequence:* the default PAM sequence for Cas9 is NGG. Alternatively, it is possible to select a PAM from an orthologous system or insert a

personalized PAM. For *S. pyogenes* dCas9, it is recommended to use the default PAM.

- *Method for determining off-targets in the genome:* there are two options: (i) one-base mismatch from the first to the eleventh base from the 5' end of PAM sequence makes Cas9 cleavage unviable (Cong et al. 2013). However, mutations further upstream of PAM maintain cleavage activity. This method searches for mismatches only in the first nine nucleotides from the PAM sequence, considering that mismatches after PAM will not promote cleavage; and (ii) mismatches can be tolerated in any position, except in the PAM. This method looks for mismatches only in the 20 base pairs upstream from PAM (Hsu et al., 2013). This is the default method, and the most used for experiments with dCas9.
- *Efficiency score:* there are several methods to evaluate and calculate the cleavage efficiency by Cas9 nuclease. The simplest form of efficiency scoring is the "G20" (Doench et al., 2016) (default), in which guanine is prioritized at position 20, right after the PAM. This value is normalized on the CHOPCHOP platform ranging from 0 to 1. Since cleavage efficiency is not relevant in experiments with dCas9, it is recommended to use the default parameter.
- *Repair profile prediction:* according to specific models (Shen et al., 2018), it is possible to predict the DNA repair profile for a sgRNA. For this option, the cell type must be selected, but only animal cells are available (U2OS, HEK293, HCT116, K562), or the "*mESC*" mode, if the cell type to be edited is not known. For experiments with dCas9, the option "*Don't calculate (saves time)*" must be selected since it is time-consuming and is not applied to dCas9 approaches.
- *Requirements for the 5' end of the sgRNA:* depending on the polymerase used for the synthesis of sgRNA, it is possible to select its first two nucleotides aiming at the best transcription efficiency. For example, 5'-GN- (for the U6 promoter) or 5'-GG- (for T7 polymerase). Regarding the dCas9 system, if the purpose is base editing using RNP, in which sgRNA is transcribed *in vitro* by T7 polymerase, the option "GG" should be selected. In transcriptional modulation experiments, in which the sgRNA will be expressed *in vivo*, regulated by the U6 promoter, the option "GN or GG" must be selected.
- *Self-complementarity*: studies suggest that self-complementarity within the sgRNA or between the spacer and the scaffold RNA) can reduce the

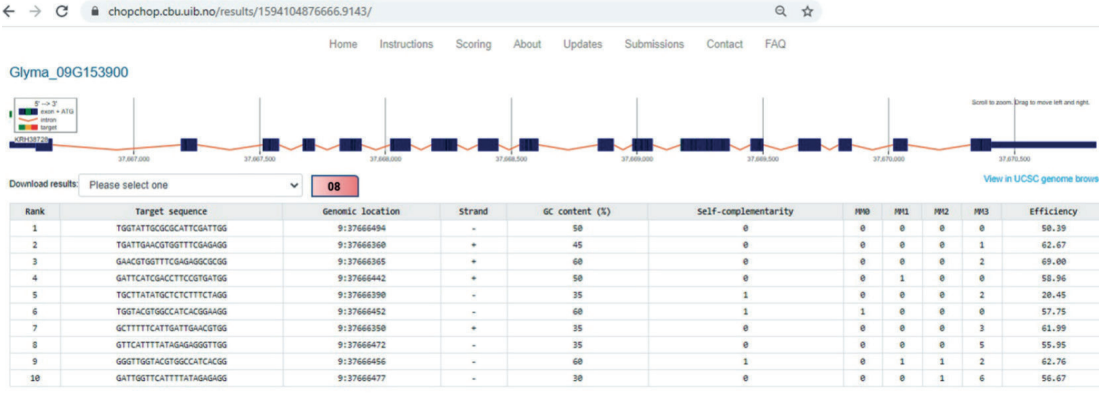
cleavage efficiency of sgRNA (Thyme et al., 2016). This option evaluates the complementarity within the sgRNA or between the spacer and a standard scaffold RNA (AGGCTAGTCCGT), an extended scaffold RNA (AGGCTAGTCCGT, ATGCTGGAA), or a custom scaffold RNA. Some users prefer to replace the first two nucleotides of sgRNA with "GG" for transcription via T7 polymerase. In this case, select the option "*I intend to replace the leading nucleotides with "GG"*. For experiments with dCas9, it is necessary to evaluate the options "*Check for self-complementarity*" and "*Check for complementarity versus backbone*". About the latter, if the experimental approach holds traditional RNA scaffold, you should select "*Standard backbone* (AGGCTAGTCCGT)", but if the approach uses RNA aptamers (for example, the SAM system), the "*Custom backbone*" option must be selected and the modified scaffold RNA sequence must be provided. Finally, the option "*I intend to replace the leading nucleotides with "GG"* should be selected exclusively for experiments with dCas9 aiming at base editing.

Step 7: Performing the analysis. After selecting all parameters, click on "*Find Target Sites!*" to submit the analysis. If everything is correct, the message "*Your job has been successfully submitted*" will be displayed. The analysis time may vary according to the selected parameters. CHOPCHOP displays the query results in a dynamic view and interactive table (Figure 4). The dynamic view displays all the on-target options for the region, color-coded according to a specific scoring system: green (best), amber (ok), and red (bad). In all cases, the gene is displayed in the 5'- 3' direction. All isoforms of the gene are displayed with their names and, at downstream, the ATG sites in the frame (green box). Simply click on a target in the visualization screen or on the options in the table and you will be redirected to a results page for each individual sgRNA, containing information about on-targets, off-targets, repair predictions (when applicable), micro-homology arms (knock-in mode), restriction sites, and primers designed for each region. In addition, CHOPCHOP lists the number of off-targets for each on-target with 0 (column "MM0"), 1 (column "MM1"), 2 (column "MM2"), or 3 (column "MM3") mismatches. Each on-target is ranked according to: (i) cleavage efficiency score ("Cas9" mode); (ii) number of off-targets and whether they have mismatches; (iii) presence of self-complementarity regions with more than three nucleotides. The number indicates the quantity of expected self-complementarity regions); (iv) GC content ("CRISPR/Cas9" mode): sgRNAs are more effective with a GC content between 40% and 70%; and (v) sgRNA location on the target gene (the closer to the 5' end, the better is the classification). Finally, each page of results provides (i) visualization of the target site of each sgRNA (with the cut site predicted in blue in the "CRISPR/Cas9" or "CRISPR/Cpf1" mode); (ii) primers

designed (in purple); (iii) restriction sites (in green unique sites in the amplicon, and in red non-exclusive sites); and (iv) details about the off-targets (genomic location, number of mismatches, and sequence).

Step 8: Download the results. The final result is available for download in four different formats: (i) BED (for viewing in the genome browser); (ii) GenBank (file with annotation and intron); (iii) TSV (results table; mostly used); and (iv) FASTA (Figure 4).

Other considerations: It is not wrong to suppose that the use of more than one sgRNA for different regions of the same promoter can, in association with the dCas9 complex and a ModT, amplify the desired levels of activation or repression (Deaner et al., 2017; Farzadfard et al., 2013; Gilbert et al., 2013). However, it is not valid in all cases, as discussed earlier. The presence/absence of nucleosomes close to the target gene can also interfere with the action of transcriptional regulators (Griesenbeck et al., 2003; Mao et al., 2011). In CRISPR-dCas9 system, in which ModTdepends on the interaction of the modulatory complex with the gene promoter region on DNA, the presence of nucleosomes is expected to be determinant to the efficiency of the technology in eukaryotes, considering that DNA is complexed with histones, making it more compact and less accessible (Horlbeck et al., 2016; Lee et al., 2007;



Rank	Target sequence	Genomic location	Strand	GC content (%)	Self-complementarity	MM0	MM1	MM2	MM3	Efficiency
1	TGGTATTGCGCGCATTCGATTGG	9:37666494	-	50	0	0	0	0	0	50.39
2	TGATTGAACGTGGTTTCGAGAGG	9:37666360	+	45	0	0	0	0	1	62.67
3	GAACGTGGTTTCGAGAGGCGCGG	9:37666365	+	60	0	0	0	0	2	69.00
4	GATTCATCGACCTTCCGTGATGG	9:37666442	+	50	0	0	1	0	0	58.96
5	TGCTTATATGCTCTCTTTCTAGG	9:37666390	-	35	1	0	0	0	2	20.45
6	TGGTACGTGGCCATCACGGAAGG	9:37666452	-	60	1	1	0	0	0	57.75
7	GCTTTTTCATTGATTGAACGTGG	9:37666350	+	35	0	0	0	0	3	61.99
8	GTTCATTTTATAGAGAGGGTTGG	9:37666472	-	35	0	0	0	0	5	55.95
9	GGGTTGGTACGTGGCCATCACGG	9:37666456	-	60	1	0	1	1	2	62.76
10	GATTGGTTCATTTTATAGAGAGG	9:37666477	-	30	0	0	0	1	6	56.67

Figure 4. CHOPCHOP v3 – Results. CHOPCHOP platform interface, in which the suggested sgRNA for dCas9-activation and a specific gene *of G. max* (ID: Glyma_09G153900) are presented. The ranking (column 1) is based on the number of off-targets presented in columns MM0, MM1, and MM2 (with one, two, or three mismatches, respectively). Even though it is not relevant for dCas9 technology, the platform provides cleavage efficiency values, in percentage, as shown in the last column. Other information is also presented, such as spacer sequence (column 2), position in the genome (column 3), DNA strand (sense or positive; antisense or negative; column 4), percentage of GC in sgRNA (column 5), and self-complementarity (column 6). The results can be downloaded (red square; step eight in Tutorial 1) according to the chosen format.

Radzisheuskaya et al., 2016; Rando; Chang, 2009; Rando; Winston, 2012; Smith et al., 2017, 2016). Thus, biochemical and *in vivo* pieces of evidence suggest that sgRNA design strategies should avoid regions close to the nucleosome core. Furthermore, the development of new tools for *in silico* design of specific and highly efficient sgRNAs should evaluate and include databases with DNA-accessibility maps and nucleosome positioning, becoming more reliable (Jiang; Pugh, 2009; Schep et al., 2015). Finally, as reported for native regulatory networks and some synthetic networks, the number of regulators associated with the regulon provides a fine-tuned adjustment of gene expression reprogramming mediated by CRISPR-dCas9. However, the use of several sgRNAs must be carefully analyzed, with a special focus on the position of regulatory elements and nucleosomes.

Transformation cassette design

The use of CRISPR technology for transcriptional or epigenetic modulation has great potential for biotechnological application in plants, especially for the manipulation of the expression of endogenous genes that may be associated with desirable phenotypes. Specifically to dCas9, few approaches can be performed in a non-transgenic manner (DNA free via RNP). The transcriptional modulation in a dCa9 system requires the constitutive expression of this endonuclease, sgRNA, and ModTs. A critical step in the application of this technology is the design of the plant-transformation cassette, which includes the selection of genetic elements for optimal production of all components required in this modulation system.

Tutorial 2: Design of transformation cassette for transcriptional modulation via dCas9

For didactic purposes, a step-by-step tutorial is proposed to guide the design of a transformation cassette, which should be divided into four basic modules: (i) expression of dCas9; (ii) sgRNA expression; (iii) expression of Adapter:ModT fusion protein (optional, according to the selected strategy); and (iv) expression of the selection marker. Considerations about the subject are presented based on both laboratory experiments and recent literature (Figure 5).

Module 1: Expression of dCas9. The design of this module is an important factor to be considered in transcriptional modulation strategies. This module has three elements: (i) promoter; (ii) dCas9; and (iii) terminator (Figure 5).

a) **Promoter:** in most cases, the promoter chosen to drive dCas9 expression is recognized by RNA polymerase II (pol II).It needs to be strong (high level of

expression) and constitutive (expressed in all tissues and at all times). In plants, the most obvious choice would be the *CaMV35S* promoter, even though some aspects must be considered. Regarding it is of viral origin, *CaMV35S* promoter can be silenced in plants. Since the transcriptional modulation of a target gene requires constant dCas9 expression, the expected phenotype will be lost if its silencing occurs. The second aspect of being considered is the constitutive characteristic of this, and other promoters used, for example, *A. thaliana's AtACT2-5*). In plants, the constitutive expression of Cas9/dCas9 can provide undesirable pleiotropic effects due to its action on possible off-targets. Finally, the function of the gene whose expression will be modulated

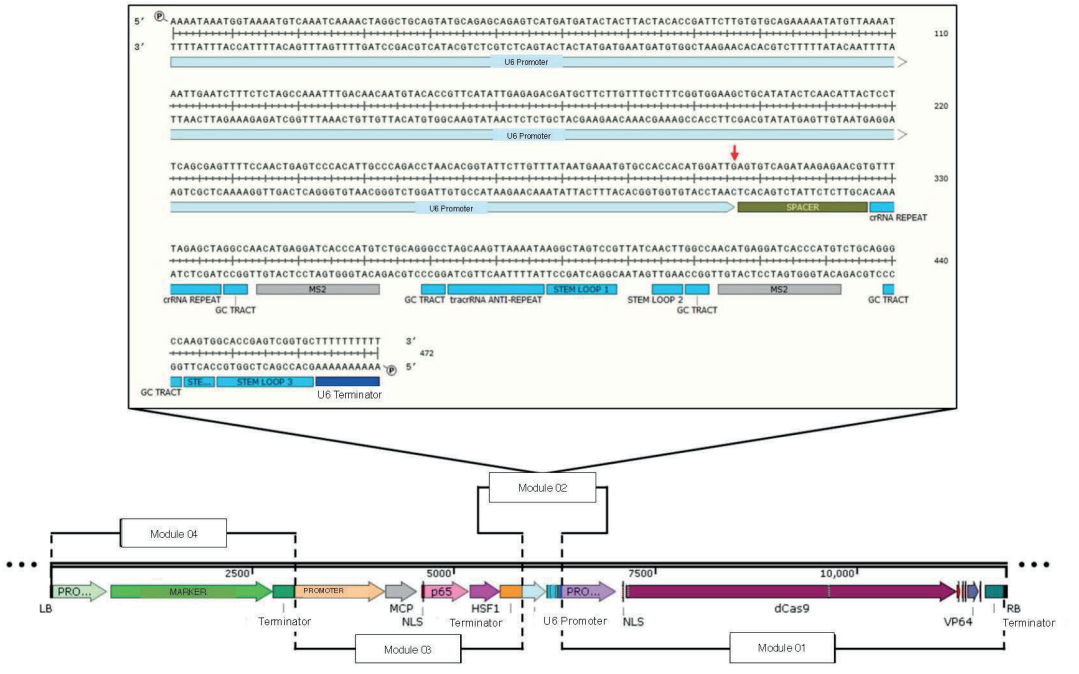


Figure 5. Design of transformation cassette – dCas9 technology (SAM system). As detailed in Tutorial 2, the transformation cassette is presented with the necessary elements for the activation of a specific gene, via dCas9 (SAM system), cloned into a binary vector. The cassette is didactically divided into four modules: (i) Module 1: Expression of dCas9; (ii) Module 2: Expression of sgRNA, with emphasis on main genetic elements (upper part of the figure). The scaffold RNA has the MS2 RNA aptamer sequence inserted in its structure both in the tetraloop and in the stem-loop 2. The red arrow highlights the importance of guanine (G) as the last nucleotide present in the 3' of the soybean *GmU6-10* promoter; (iii) Module 3: Expression of the Adapter:ModT fusion protein. In this module, the transcriptional modulators presented are protein domains from transcriptional activators p65 and HSF1, characteristic of the SAM system; and (iv) Module 04: Expression of the selection marker. *Acronyms:* LB - left border; MCP - MS2 Coat Protein; NLS - nuclear location site; RB - right border.

is the third aspect to be considered. Stress-related genes, when expressed constitutively, can lead to undesirable effects. In this way, if dCas9promoter is constitutive, genes with this characteristic would also be expressed at high levels constitutively. Thus, the choice of the promoter for the expression of dCas9 must be aligned with the function of the gene whose expression is intended to modulate. For example, if the desired phenotype is to improve plant resistance to a pathogen, an induced promoter-driven dCas9 expression can be considered.

b) **dCas9:** the nucleotide sequence of dCas9 must be carefully evaluated. The first aspect of being observed is whether the mutations necessary for the loss of catalytic activity are inserted in the sequence (D10A in the catalytic domain RuvC and H840A in the catalytic domain HNH). The second aspect is the presence of a nuclear localization signal (NLS) since the dCas9 activity will be performed exclusively in the eukaryotic nucleus. In plants, the main NLSs are (i) NLS from monkey virus SV40 (PKKKRKV); and (ii) Nucleoplasmin bipartite NLS (KRPAATKKAGQAKKKK). The number of repetitions and the region where they will be inserted is variable. Generally, the number of repetitions of each NLS range from 1 to 4, and it can be present in both amino and carboxyterminal of dCas9. Constructions with only one NLS are rare. NLSs can be inserted after ModT, with no loss of dCas9 activity. The third aspect of being considered is the fusion of dCas9 with ModT. Generally, this connection is mediated by a linker or bridge, corresponding to a sequence of 5 to 20 amino acids linking the dCas9 carboxy-terminal to ModT amino-terminal. These amino acids must provide flexibility and allow both proteins to show their correct activity. Some *in tandem* NLSs, spaced properly by at least five neutral amino acids, can be used as linkers between dCas9 and ModT. Moreover, the linker's length and amino acid composition between dCas9 and base editor domains can be crucial for the specificity of the edition. Short linkers will reduce the window of the nucleotides that can be edited. But be careful: exceedingly small linkers can interfere with the function of both proteins.

c) **Terminator:** the choice of the terminator is free. When using the *CaMV35S* promoter, it is suggested the use of 35S terminator (*t35S*), but any terminator recognized by an RNA pol II can be used, such as *tNOS* (NOS terminator of nopaline synthase gene from *Agrobacterium tumefaciens*). Experimental data showed that the presence of two different *in tandem* terminators (for example, *t35S* and *tNOS*) could increase transcription levels by improving the stability of transgene expression due to the reduction of post-transcriptional gene silencing (Basso et al., 2020).

Module 2: Expression of sgRNA. This module must be designed to efficiently delineate the expression of sgRNA and can be divided into four elements: (i) *U6* promoter; (ii) spacer; (iii) scaffold RNA; and (iv) *U6* terminator (*tU6*) (Figure 5).

a) **U6 Promoter:** the levels of sgRNA expression are directly correlated with the efficiency of CRISPR/Cas9-mediated genome engineering in eukaryotic cells. In general, RNA polymerase III (pol III) promoters are mostl used to drive sgRNA expression, since pol II promoters add extra nucleotides to the 5′ and 3′ ends (Hsu et al., 2013; Yoshioka et al., 2015). Originally, pol III *SNR52* and *RPR1* promoters were adopted for constitutive sgRNAs expression in yeasts (DiCarlo et al., 2013; Farzadfard et al., 2013; Gilbert et al., 2013). Then, to allow more flexibility in the design and higher expression levels of sgRNAs, two studies on constitutive delivery of sgRNAs tested the fusion of Delta Hepatitis Virus (HDV), and auto-cleavable ribozymes, allowing the expression of sgRNAs with pol II promoters (Gao; Zhao, 2014; Ryan et al., 2014). In addition to native pol III and inducible pol II promoters, different classes of promoters were used to regulate the expression of sgRNAs in eukaryotic systems (Farzadfard et al., 2013; Ferreira et al., 2018; Jensen et al., 2017; Nishimasu et al., 2014; Smith et al., 2016; Zhang et al., 2017). The *U6* stands out as pol III promoter that is naturally important in the endogenous expression of small non-coding nuclear RNAs, involved in intron splicing. Different *U6* promoters have been frequently used in plants and animals to express small RNAs at higher levels (Li et al., 2007; Miyagishi; Taira, 2002). Currently, this type of promoter is the preferred choice for regulating sgRNA expression in CRISPR/Cas9 (and variants) vectors (Friedland et al., 2013; Li et al., 2013). Moreover, the *U6* promoter has a highly conserved TSS, starting with a guanine nucleotide, which helps to improve the homogeneity of the transcribed sgRNA molecule and reduce the effects of off-targets (Li et al., 2007). CRISPR/Cas9 vectors with the *U6* promoter driven the sgRNA expression have been used successfully in several plant species. The *OsU6a*, *OsU6b,* and *OsU6c* promoters of rice are the most used for monocotyledons, and *AtU6-1* and *AtU6-29* from *Arabidopsis* are the preferred ones for dicotyledons (Jiang et al., 2013; Li et al., 2014; Mikami et al., 2015). However, there are limitations to its "universal" nature, since the *Arabidopsis U6* promoter was inefficient in wheat and rice (Shan et al., 2013). Thus, it is clear that the use of species-specific *U6* promoters can result in increased sgRNA expression and more efficient editing (Ng and Dean, 2017; Sun et al., 2015). In soybean, for example, the levels of sgRNA expressed by the endogenous promoter *GmU6* were twice higher than those obtained with the *AtU6-26* promoter, resulting in a considerable improvement in gene editing efficiency (14.7–20.2 % for *GmU*6 vs. 3.2–9.7% for *AtU6-26*) (Sun et al., 2015).

In this way, aiming at improving sgRNAs expression, at least one endogenous *U6* promoter has already been characterized for crops such as cotton, rice, barley, corn, sorghum, and wheat(Gasparis et al., 2018; Jiang et al., 2013; Long et al., 2018; Qi et al., 2016; Sun et al., 2015). It is also important to keep in mind that plant genomes contain several *U6* genes with different expression levels and not all *U6* promoters are equally efficient to drive gene expression (Domitrovich; Kunkel, 2003; Wang et al., 2008). For new plant species, it is important to characterize specific endogenous *U6* promoters to optimize its use and avoid waste of time. The main attention to be taken regarding the *U6* promoter experiments with Cas9 and dCas9 is the presence of a guanine (G) at the 3' end of the promoter (region closest to the first nucleotide of the sgRNA). If the available sequence of the *U6* promoter in question does not contain this nucleotide, the first nucleotide in the sgRNA must be a guanine. In some cases, some researchers add an extra guanine to the sgRNA sequence during design, to optimize its expression.

b) **Spacer :** this is the region of sgRNA that will effectively recognize the target sequence. The main aspects of its design have been thoroughly described previously. In the sgRNA expression cassette, the spacer comes immediately after the *U6* promoter and before the scaffold RNA.

c) **Scaffold RNA:** it is the constant sequence present in sgRNA, important in anchoring with Cas9/dCas9. As seen before, the structure of scaffold RNA from Cas9 and dCAS9 nucleases is formed by a tetraloop and three stem-loops. For dCas9 technology, the scaffold RNA can be presented in two ways: (i) classical form, with the standard sequence without any modification. In this approach, transcriptional or epigenetic modulation is mediated only by ModT fused with dCas9; (ii) modified form (SAM methodology), with RNA aptamers (for example, MS2, PP7, and COM) inserted in tetraloop and stem-loop 2 regions. The insertion of these RNA aptamers allows recognition by adapter protein, specific to each aptamer (for example, MCP, PCP, and COM) (Haimovich et al., 2016; Johansson et al., 1997; Lim and Peabody, 2002; Zhang et al., 2019c). These adapter proteins are fused with ModTs. Thus, the choice of the adapter protein needs to be in perfect agreement with the experimental strategy addressed. Regardless the strategy, the scaffold RNA comes immediately after the spacer and before the *tU6*.

d) **U6 terminator (tU6):** short and simple terminator (10 thymines added in 3' end of spacer) recognized by the most pol III (Gao et al., 2018). It should not be replaced by any other terminator used in the laboratory.

Module 3: Expression of Adapter:ModT fusion protein. This module is associated with SAM methodology (sgRNA with RNA aptamers). If using the classical approach (ModT fused only with dCas9 ,), this module must not be inserted in the vector design. Three main genetic elements can be highlighted: (i) promoter; (ii) fusion protein Adapter:ModT; and (iii) terminator.

a) **Promoter:** the choice of promoter for Adapter:ModT must follow the same criteria described in Module 1, for dCas9. If possible, both Modules should have the same promoter.

b) **Adapter**: ModT fusion protein: The adapter-aptamer coupling is specific. For example, if the MS2 aptamer is chosen, the adapter protein should be MCP. Due to its wide biotechnological application, the identification of new adapter-aptamer pairs is growing every day (Zhang et al., 2019c). Most of the constructions successfully used in dCas9 strategies presented ModTs fused *in tandem* with the carboxy-terminal MCP protein adapter, properly separated by linkers. There is no limit for the number of fused ModTs, but most constructions have only 1 or 2. Furthermore, the fusion protein needs to be transported to the nucleus to perform its activity. Thus, it is necessary to insert NLSs, following the same criteria presented in Module 1 for dCas9. Even in this case, NLS is commonly inserted between the adapter protein and the first ModT, also acting as a linker.

c) **Terminator:** similar to Module 1, the choice of the terminator is free. Any terminator that is recognized by a pol II can be used, for example, *tNOS*.

Module 04: Expression of Selection Marker. This module is specific to each plant species, considering the genetic transformation method that will be used. Similar to Modules 1 and 3, three elements are important: (i) promoter; (ii) selection marker, which in most cases is a gene that confers resistance to an antibiotic or an herbicide but can be any gene reporter, such as *gfp*, *uid*A, among others; and (iii) terminator. Since the design of this module can be variable, the promoters and/or terminators previously used in Modules 1 and 3 should not be repeated to avoid possible silencing in the plant genome (Figure 5).

Other considerations: Considering the vector design, some extra points need to be highlighted:

a) In several laboratories, the plasmid vectors are constructed by chemical synthesis. If this is the case, an important step to be considered is the codon usage optimization in Modules 1, 3, and 4. This care is extremely relevant considering the translation efficiency of some elements presented in these modules.

b) Keep all Modules in the same strand (direction), to avoid steric impediment during gene transcription.

c) It is suggested that unique sites for restriction enzymes be inserted between each module to facilitate *in vitro* manipulation, if necessary.

d) The sgRNAs can be expressed in simplex (1 sgRNA), or multiplex (2 or more sgRNAs). For multiplex systems, each sgRNA must have its *U6* promoter and terminator. In general, for the expression of different sgRNAs in the same plasmid vector, Module 2 must be repeated *in tandem* (preferably separated by unique restriction sites). Until now, there are no reports regarding the maximum number of sgRNAs that can be expressed in the same plasmid vector, but successful systems did not exceed four sgRNAs per vector.

e) The most used backbones are binary vectors based on the T-DNA from *A. tumefaciens*, but this can be replaced by another one, depending on the plant species and transformation method used. The cloning steps and subsequent amplification in bacteria can be hampered since the length of the final vector can reach 10.000 to 15.000 base pairs. If the chosen backbone is based on the T-DNA from *A. tumefaciens*, it is suggested to insert the selection marker closer to the left border (LB), to optimize the selection of transformed plants.

Final considerations

The CRISPR/dCas9 technology applied to the modulation of gene expression has as a main advantage the possibility of promoting transcriptional control of multiple gene targets simultaneously, simulating native networks of cellular metabolic pathways. It is a key tool for synthetic biology and deepening knowledge of plant pathways related to development, productivity, and response to biotic and abiotic stresses, which are frequent targets of genetic engineering for plant breeding. CRISPR/dCas9 is a functional technology and straightforward in transient transformation systems, allowing rapid simulation and visualization of the effects of gene expression modulation on cellular metabolic pathways. The different strategies available, varying elements such as ModT and nucleases, allow the adaptation and the use of this technique in different models.

An advantage of the CRISPR system for the transcriptional activation of an endogenous gene, compared to the overexpression strategy, using constitutive promoters to drive the complete coding sequence of a gene, is that this last strategy omits splicing patterns of endogenous genes and mask alternative transcripts. Therefore, the gene product may be unstable, interact with unconventional

partners, or have an incorrect subcellular localization (Park et al., 2017). However, when working with CRISPR/dCas9 technology, some topics deserve attention:

a) CRISPR/dCas9 technology does not induce definitive changes in the target genomic region, and the developed plants will always be transgenic. If the intention is to develop a biotechnological product, this can lead to a long and costly process for commercial release.

b) The effects on the modulation of target gene expression depend on transcriptional control's cellular processes and can vary between developed transgenic lines. Although inheritable throughout transgenic generations, gene modulation may also vary between mother and daughter plants or be influenced by environmental factors.

c) The technology requires minimal knowledge of the target promoter region sequence, which can be a limiting factor when working with species without detailed genomic information.

In this context, the constant improvements of CRISPR/dCas9 technology aiming at increasing its specificity and efficiency as, for example, with the use of other Cas orthologous, will increase the range of its use, especially in research with cultivated plants.

References

ABADI, S.; YAN, W.X.; AMAR, D.; MAYROSE, I. A machine learning approach for predicting {CRISPR}-Cas9 cleavage efficiencies and patterns underlying its mechanism of action. **PLoS Computational Biology**, 13, 2017. e1005807. https://doi.org/10.1371/journal.pcbi.1005807

ALKAN, F.; WENZEL, A.; ANTHON, C.; HAVGAARD, J. H.; GORODKIN, J. CRISPR-Cas9 off-targeting assessment with nucleic acid duplex energy parameters. **Genome Biology**, 19, n.177, 2018. https://doi.org/10.1186/s13059-018-1534-x

ANDERS, C.; NIEWOEHNER, O.; DUERST, A.; JINEK, M. Structural basis of PAM-dependent target {DNA} recognition by the Cas9 endonuclease. **Nature**, 513, p. 569–573, 2014. https://doi.org/10.1038/nature13579

BASSO, M. F.; ARRAES, F. B. M.; GROSSI-DE-SA, M.; MOREIRA, V. J. V.; ALVES-FERREIRA, M.; GROSSI-DE-SA, M. F. Insights into genetic and molecular elements for transgenic crop development. **Frontiers in Plant Science,** 2020. https://doi.org/10.3389/fpls.2020.00509

BEERLI, R. R.; SEGAL, D. J.; DREIER, B.; BARBAS, C. F. 1998. Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. **Proceedings of the National Academy of Sciences,** v. 95, p. 14628–14633, 1998. https://doi.org/10.1073/pnas.95.25.14628

CHAVEZ, A.; SCHEIMAN, J.; VORA, S.; PRUITT, B. W.; TUTTLE, M.; IYER, P. R.; LIN, S.; KIANI, S.; GUZMAN, C. D.; WIEGAND, D. J.; TER-OVANESYAN, D.; BRAFF, J. L.; DAVIDSOHN, N.; HOUSDEN, B. E.; PERRIMON, N., WEISS, R.; AACH, J.; COLLINS, J. J.; CHURCH, G. M. Highly efficient Cas9-mediated transcriptional programming. **Nature Methods**, v .12, p. 326–328, 2015. https://doi.org/10.1038/nmeth.3312

CHEN, B.; GILBERT, L. A.; CIMINI, B. A.; SCHNITZBAUER, J.; ZHANG, W.; LI, G.-W.; PARK, J.; BLACKBURN, E. H.; WEISSMAN, J. S.; QI, L. S., HUANG, B. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas System. **Cell Press Journal**, v. 155, n. 7 1479–1491, 2013. https://doi.org/10.1016/j.cell.2013.12.001

CHUAI, G.-H.; WANG, Q.-L.; LIU, Q. In silico meets in vivo: towards computational CRISPR-based sgRNA design. **Trends Biotechnology**. v. 35, p. 12–21, 2017. https://doi.org/10.1016/j.tibtech.2016.06.008

CHUAI, G.; MA, H.; YAN, J.; CHEN, M.; HONG, N.; XUE, D.; ZHOU, C.; ZHU, C.; CHEN, K.; DUAN, B.; GU, F.; QU, S.; HUANG, D.; WEI, J.; LIU, Q. DeepCRISPR: optimized CRISPR guide {RNA} design by deep learning. **Genome Biology**, v. 19, n. 80, 2018. https://doi.org/10.1186/s13059-018-1459-4

CONG, L.; RAN, F. A.; COX, D.; LIN, S.; BARRETTO, R.; HABIB, N.; HSU, P. D.; WU, X.; JIANG, W.; MARRAFFINI, L. A.; ZHANG, F. Multiplex genome engineering using CRISPR/Cas systems. **Science**, n. 339, 819–823, 2013. https://doi.org/10.1126/science.1231143

DANG, Y.; JIA, G.; CHOI, J., MA, H.; ANAYA, E.; YE, C.; SHANKAR, P.; WU, H. Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. **Genome Biology**, v. 16, n. 280, 2015. https://doi.org/10.1186/s13059-015-0846-3

DEANER, M.; ALPER, H. S. Systematic testing of enzyme perturbation sensitivities via graded {dCas9} modulation in *Saccharomyces cerevisiae*. **Metabolic Engineering**, v. 40, p. 14–22, 2017. https://doi.org/10.1016/j.ymben.2017.01.012

DEANER, M.; MEJIA, J.; ALPER, H. S. Enabling graded and large-scale multiplex of desired genes using a dual-mode dCas9 activator in *Saccharomyces cerevisiae*. **ACS Synthetic Biology**, v .6, p. 1931–1943, 2017. https://doi.org/10.1021/acssynbio.7b00163

DiCARLO, J. E.; NORVILLE, J. E.; MALI, P.; RIOS, X.; AACH, J.; CHURCH, G. M. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. **Nucleic Acids Research**, v. 41, p. 4336–4343, 2013. https://doi.org/10.1093/nar/gkt135

DOENCH, J. G.; FUSI, N.; SULLENDER, M.; HEGDE, M.; VAIMBERG, E. W.; DONOVAN, K. F.; SMITH, I.; TOTHOVA, Z.; WILEN, C.; ORCHARD, R.; VIRGIN, H. W.; LISTGARTEN, J.; ROOT, D. E. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. **Nature Biotechnology**, v. 34, p. 184–191, 2016. https://doi.org/10.1038/nbt.3437

DOENCH, J.G.; HARTENIAN, E.; GRAHAM, D.B.; TOTHOVA, Z.; HEGDE, M.; SMITH, I.; SULLENDER, M.; EBERT, B.L.; XAVIER, R.J.; ROOT, D.E. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. **Nature Biotechnology**, v. 32, p. 1262-1267, 2014. https://doi.org/10.1038/nbt.3026

DOMITROVICH, A. M.; KUNKEL, G. R. Multiple, dispersed human U6 small nuclear RNA genes with varied transcriptional efficiencies. **Nucleic Acids Research,** v. 31, p. 2344-2352, 2003. https://doi.org/10.1093/nar/gkg331

DREISSIG, S.; SCHIML, S.; SCHINDELE, P.; WEISS, O.; RUTTEN, T.; SCHUBERT, V.; GLADILIN, E.; METTE, M. F.; PUCHTA, H.; HOUBEN, A. Live-cell CRISPR imaging in plants reveals dynamic telomere movements. **The Plant Journal**, v. 91, p. 565–573, 2017. https://doi.org/10.1111/tpj.13601

FARZADFARD, F.; PERLI, S. D.; LU, T. K. Tunable and multifunctional eukaryotic transcription factors based on CRISPR/Cas. **ACS Synthetic Biology**, v. 2, p. 604–613, 2013. https://doi.org/10.1021/sb400081r

FERREIRA, R.; SKREKAS, C.; NIELSEN, J.; DAVID, F. Multiplexed CRISPR/Cas9 genome editing and gene regulation using Csy4 in *Saccharomyces cerevisiae*. **ACS Synthetic Biology**, v. 7, p. 10–15, 2018. https://doi.org/10.1021/acssynbio.7b00259

FONFARA, I.; LE RHUN, A.; CHYLINSKI, K.; MAKAROVA, K. S.; LÉCRIVAIN, A.-L.; BZDRENGA, J.; KOONIN, E. V.; CHARPENTIER, E. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type {II} CRISPR-Cas systems. **Nucleic Acids Research**, v. 42, p. 2577–2590, 2014. https://doi.org/10.1093/nar/gkt1074

FRIEDLAND, A. E.; TZUR, Y. B.; ESVELT, K. M.; COLAIÁCOVO, M. P.; CHURCH, G. M.; CALARCO, J. A. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. **Nature Methods**, v. 10, 741–743, 2013. https://doi.org/10.1038/nmeth.2532

FU, Y.; SANDER, J. D.; REYON, D.; CASCIO, V. M., JOUNG, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. **Nature Biotechnology,** v. 32, p. 279–284, 2014. https://doi.org/10.1038/nbt.2808

GAO, Y.; ZHAO, Y. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. **Journal of Integrative Plant Biology**, v. 56, p. 343–349, 2014. https://doi.org/10.1111/jipb.12152

GAO, Z.; HERRERA-CARRILLO, E.; BERKHOUT, B. Delineation of the exact transcription termination signal for type 3 polymerase III. **Molecular Therapy**-Nucleic Acids, v. 10, p. 36–44, 2018. https://doi.org/10.1016/j.omtn.2017.11.006

GASPARIS, S.; KAŁA, M.; PRZYBOROWSKI, M.; ŁYŻNIK, L. A.; ORCZYK, W.; NADOLSKA-ORCZYK, A. A simple and efficient CRISPR/Cas9 platform for induction of single and multiple, heritable mutations in barley (*Hordeum vulgare* L.). **Plant Methods**, v. 14, 111, 2018. https://doi.org/10.1186/s13007-018-0382-8

GAUDELLI, N. M.; KOMOR, A. C.; REES, H. A.; PACKER, M. S.; BADRAN, A. H.; BRYSON, D. I.; LIU, D. R. Programmable base editing of T to G C in genomic DNA without DNA cleavage. **Nature,** v. 551, p. 464–471, 2017. https://doi.org/10.1038/nature24644

GILBERT, L. A.; LARSON, M. H.; MORSUT, L.; LIU, Z.; BRAR, G.A.; TORRES, S.E.; STERN-GINOSSAR, N.; BRANDMAN, O.; WHITEHEAD, E. H.; DOUDNA, J. A.; LIM, W. A.; WEISSMAN, J. S.; QI, L. S. CRISPR-mediated modular {RNA}-guided regulation of transcription in eukaryotes. **Cell**, v. 154, p. 442–451, 2013. https://doi.org/10.1016/j.cell.2013.06.044

GRAHAM, D. B.; ROOT, D. E. Resources for the design of CRISPR gene editing experiments. **Genome Biology**, v. 16, 260, 2015. https://doi.org/10.1186/s13059-015-0823-x

GRATZ, S. J.; UKKEN, F. P.; RUBINSTEIN, C. D.; THIEDE, G.; DONOHUE, L. K. CUMMINGS, A. M.; O'CONNOR-GILES, K. M., 2014. Highly specific and efficient CRISPR/Cas9-

catalyzed homology-directed repair in Drosophila. **Genetics**, v. 196, p. 961–971, 2014. https://doi.org/10.1534/genetics.113.160713

GRIESENBECK, J.; BOEGER, H.; STRATTAN, J. S.; KORNBERG, R. D. Affinity purification of specific chromatin segments from chromosomal loci in yeast. **Molecular and Cellular Biology**, v. 23, p. 9275–9282, 2003. https://doi.org/10.1128/mcb.23.24.9275-9282.2003

HAEUSSLER, M.; SCHÖNIG, K.; ECKERT, H.; ESCHSTRUTH, A.; MIANNÉ, J.; RENAUD, J.-B.; SCHNEIDER-MAUNOURY, S.; SHKUMATAVA, A.; TEBOUL, L.; KENT, J.; JOLY, J.-S.; CONCORDET, J.-P. 2016. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. **Genome Biology**, v.17, p.148, 2016. https://doi.org/10.1186/s13059-016-1012-2

HAIMOVICH, G.; ZABEZHINSKY, D.; HAAS, B.; SLOBODIN, B.; PURUSHOTHAMAN, P.; FAN, L.; LEVIN, J. Z.; NUSBAUM, C.; GERST, J. E. Use of the MS2 aptamer and coat protein for RNA localization in yeast: a response to MS2 coat proteins bound to yeast mRNAs block 5′ to 3′ degradation and trap mRNA decay products: implications for the localization of mRNAs by MS2-MCP sys". **RNA**, v. 22, p. 660–666, 2016. https://doi.org/10.1261/rna.055095.115

HILTON, I. B.; D'IPPOLITO, A. M.; VOCKLEY, C. M.; THAKORE, P. I.; CRAWFORD, G. E.; REDDY, T. E.; GERSBACH, C. A. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. **Nature Biotechnology**, v. 33, p. 510–517, 2015. https://doi.org/10.1038/nbt.3199

HORLBECK, M. A.; GILBERT, L. A.; VILLALTA, J. E.; ADAMSON, B.; PAK, R. A.; CHEN, Y.; FIELDS, A.P.; PARK, C. Y.; CORN, J. E., KAMPMANN, M.; WEISSMAN, J. S. 2016. Compact and highly active next-generation libraries for {CRISPR}-mediated gene repression and activation. **eLife**, 2016. https://doi.org/10.7554/{eLife}.19760

HOWE, F. S.; RUSSELL, A.; LAMSTAES, A. R.; El-SAGHEER, A.; NAIR, A.; BROWN, T.; MELLOR, J. CRISPRi is not strand-specific at all loci and redefines the transcriptional landscape. **eLife**, 2017. https://doi.org/10.7554/{eLife}.29878

HSU, P. D.; LANDER, E. S.; ZHANG, F. Development and applications of CRISPR-Cas9 for genome engineering. **Cell,** v. 157, p.1262–1278, 2014. https://doi.org/10.1016/j.cell.2014.05.010

HSU, P. D.; SCOTT, D. A.; WEINSTEIN, J. A.; RAN, F. A.; KONERMANN, S.; AGARWALA, V.; LI, Y.; FINE, E. J.; WU, X.; SHALEM, O.; CRADICK, T. J.; MARRAFFINI, L. A.; BAO, G.; ZHANG, F. DNA targeting specificity of RNA-guided Cas9 nucleases. **Nature Biotechnology,** v. 31, p. 827–832, 2013. https://doi.org/10.1038/nbt.2647

HUA, K.; TAO, X.; YUAN, F.; WANG, D.; ZHU, J.-K. Precise AT to GC base editing in the rice genome. **Molecular Plant,** p. 627–630, 2018. https://doi.org/10.1016/j.molp.2018.02.007

HWANG, G.-H.; PARK, J.; LIM, K.; KIM, S.; YU, J.; YU, E.; KIM, S.-T.; EILS, R.; KIM, J.-S.; BAE, S. Web-based design and analysis tools for CRISPR base editing. **BMC Bioinforma**, v. 19, 542, 2018. https://doi.org/10.1186/s12859-018-2585-4

IKEDA, M.; OHME-TAKAGI, M. A Novel Group of Transcriptional Repressors in Arabidopsis. P **Plant and Cell Physiology**, v. 50, 970–975, 2009. https://doi.org/10.1093/pcp/pcp048

JACKSON, J. P.; LINDROTH, A. M.; CAO, X.; JACOBSEN, S. E. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. **Nature, v.** 416, p. 556–560, 2002. https://doi.org/10.1038/nature731

JENSEN, E. D.; FERREIRA, R.; JAKOČIŪNAS, T.; ARSOVSKA, D.; ZHANG, J.; DING, L.; SMITH, J. D.; DAVID, F.; NIELSEN, J.; JENSEN, M. K., KEASLING, J. D. Transcriptional reprogramming in yeast using dCas9 and combinatorial {gRNA} strategies. **Microbial Cell Factories,** v. 16, p. 46, 2017. https://doi.org/10.1186/s12934-017-0664-2

JIANG, C.; PUGH, B. F. A compiled and systematic reference map of nucleosome positions across the *Saccharomyces cerevisiae* genome. **Genome Biology**, 10, R109, 2009. https://doi.org/10.1186/gb-2009-10-10-r109

JIANG, F.; DOUDNA, J. A. CRISPR-Cas9 Structures and Mechanisms. **Annual Review of Biophysics**, v. 46, p .505–529, 2017. https://doi.org/10.1146/annurev-biophys-062215-010822

JIANG, W.; ZHOU, H.; BI, H.; FROMM, M.; YANG, B.; WEEKS, D. P. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. N **Nucleic Acids Research**, v. 41, p. e188, 2013. https://doi.org/10.1093/nar/gkt780

JINEK, M.; CHYLINSKI, K.; FONFARA, I.; HAUER, M.; DOUDNA, J. A.; CHARPENTIER, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. **Science**, v. 337, n. 6096 p. 816–821, 2012. https://doi.org/10.1126/science.1225829

JINEK, M.; JIANG, F.; TAYLOR, D. W.; STERNBERG, S. H.; KAYA, E.; MA, E.; ANDERS, C.; HAUER, M.; ZHOU, K.; LIN, S.; KAPLAN, M.; IAVARONE, A. T.; CHARPENTIER, E.; NOGALES, E.; DOUDNA, J. A. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. **Science**, v. 343, n. 6176, 1247997, 2014. https://doi.org/10.1126/science.1247997

JOHANSSON, H. E.; LILJAS, L.; UHLENBECK, O. C. RNA recognition by the MS2 phage coat protein. **Seminars in Virology**, v. 8, p. 176-185, 1997. https://doi.org/10.1006/smvy.1997.0120

KANG, B. C.; YUN, J. Y.; KIM, S. T.; SHIN, Y. J.; RYU, J.; CHOI, M.; WOO, J. W.; KIM, J. S. Precision genome engineering through adenine base editing in plants. **Nature Plants, v.** 4, p. 427-431, 2018. https://doi.org/10.1038/s41477-018-0178-x

KHOSRAVI, S.; SCHINDELE, P.; GLADILIN, E.; DUNEMANN, F.; RUTTEN, T.; PUCHTA, H.; HOUBEN, A. Application of aptamers improves CRISPR-based live imaging of plant telomeres. **bioRxiv**, 2020. https://doi.org/10.1101/2020.05.05.078246

KIANI, S.; CHAVEZ, A.; TUTTLE, M.; HALL, R. N.; CHARI, R.; TER-OVANESYAN, D.; QIAN, J.; PRUITT, B. W.; BEAL, J.; VORA, S.; BUCHTHAL, J.; KOWAL; E. J. K.; EBRAHIMKHANI, M. R.; COLLINS, J. J.; WEISS, R.; CHURCH, G. Cas9 gRNA engineering for genome editing, activation and repression. **Nature Methods**, v. 12, p. 1051–1054, 2015. https://doi.org/10.1038/nmeth.3580

KOMOR, A. C.; KIM, Y. B.; PACKER, M. S.; ZURIS, J. A.; LIU, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. **Nature**, v. 533, p. 420–424, 2016. https://doi.org/10.1038/nature17946

KONERMANN, S.; BRIGHAM, M. D.; TREVINO, A. E.; JOUNG, J.; ABUDAYYEH, O. O.; BARCENA, C.; HSU, P. D.; HABIB, N.; GOOTENBERG, J. S.; NISHIMASU, H.; NUREKI, O.; ZHANG, F. Genome-scale

transcriptional activation by an engineered CRISPR-Cas9 complex. **Nature**, v. 517, p. 583–588, 2015. https://doi.org/10.1038/nature14136

KUANG, Y.; LI, S.; REN, B.; YAN, F.; SPETZ, C.; LI, X.; ZHOU, X.; ZHOU, H. Base-editing-mediated artificial evolution of OsALS1 in planta to develop novel. **Molecular Plant**, v. 13, n. 4, p. 565–572, 2020. https://doi.org/10.1016/j.molp.2020.01.010

LABUN, K.; MONTAGUE, T. G.; GAGNON, J. A.; THYME, S. B.; VALEN, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. **Nucleic Acids Research**, v. 44, n. W1,8, p. W272-W276, 2016. https://doi.org/10.1093/nar/gkw398

LABUN, K.; MONTAGUE, T. G.; KRAUSE, M.; TORRES CLEUREN, Y. N.; TJELDNES, H.; VALEN, E. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. **Nucleic Acids Research**, v. 47, n. W1, 02, p. W171–W174, 2019. https://doi.org/10.1093/nar/gkz365

LANGMEAD, B.; SALZBERG, S. L. Fast gapped-read alignment with Bowtie 2. **Nature Methods**, n. 9, p. 357–359, 2012. https://doi.org/10.1038/nmeth.1923

LANGMEAD, B.; TRAPNELL, C.; POP, M.; SALZBERG, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. **Genome Biology**, v. 10, n. R25, 2009. https://doi.org/10.1186/gb-2009-10-3-r25

LEE, J. E.; NEUMANN, M.; DURO, D. I.; SCHMID, M. CRISPR-based tools for targeted transcriptional and epigenetic regulation in plants. **PLoS One**, v. 14, p. 1-17, 2019. https://doi.org/10.1371/journal.pone.0222778

LEE, W.; TILLO, D.; BRAY, N.; MORSE, R. H.; DAVIS, R. W.; HUGHES, T. R.; NISLOW, C. A high-resolution atlas of nucleosome occupancy in yeast. **Nature Genetics**, v. 39, p. 1235-1244, 2007. https://doi.org/10.1038/ng2117

LEI, Y.; LU, L.; LIU, H.-Y.; LI, S.; XING, F.; CHEN, L.-L. CRISPR-P: a web tool for synthetic single-guide RNA design of CRISPR-system in plants. **Molecular Plant Pathology**, v. 7, p.1494-1496, 2014. https://doi.org/10.1093/mp/ssu044

LI, H.; DURBIN, R. Fast and accurate short read alignment with Burrows-Wheeler transform. **Bioinformatics**, v. 25, n. 14 1754–1760, 2009. https://doi.org/10.1093/bioinformatics/btp324

LI, J.-F.; ZHANG, D.; SHEEN, J. Cas9-based genome editing in Arabidopsis and tobacco. **Methods in Enzymology**, v. 546, p. 459–472, 2014. https://doi.org/10.1016/B978-0-12-801185-0.00022-2

LI, J.; NORVILLE, J. E.; AACH, J.; MCCORMACK, M.; ZHANG, D.; BUSH, J.; CHURCH, G. M.; SHEEN, J. Multiplex and homologous recombination-mediated genome editing in Arabidopsis and *Nicotiana benthamiana* using guide {RNA} and Cas9. **Nature Biotechnology**, v. 31, p. 688–691, 2013. https://doi.org/10.1038/nbt.2654

LI, X.; JIANG, D.-H.; YONG, K.; ZHANG, D.-B. Varied transcriptional efficiencies of multiple Arabidopsis U6 small nuclear {RNA} genes. **Journal of Integrative Plant Biology**, v. 49, p. 222–229, 2007. https://doi.org/10.1111/j.1744-7909.2007.00393.x

LI, Y.; ZHU, J.; WU, H.; LIU, C.; HUANG, C.; LAN, J; ZHAO, Y.; XIE, C. Precise base editing of non-allelic acetolactate synthase genes confers sulfonylurea herbicide resistance in maize. **The Crop Journal**, v. 8, n. 3, p. 4–11, 2020. https://doi.org/10.1016/j.cj.2019.10.001

LI, Z.; ZHANG, D.; XIONG, X.; YAN, B.; XIE, W.; SHEEN, J.; LI, J. F. A potent Cas9-derived gene activator for plant and mammalian cells. **Nature Plants**, v. 3, p. 930–936, 2017. https://doi.org/10.1038/s41477-017-0046-0

LIM, F.; PEABODY, D. S. RNA recognition site of PP7 coat protein. **Nucleic Acids Research**, v. 30, p. 4138–4144, 2002. https://doi.org/10.1093/nar/gkf552

LIU, G.; ZHANG, Y.; ZHANG, T. Computational approaches for effective CRISPR guide RNA design and evaluation. **Computational and Structural Biotechnology**, v. 18, p. 35–44, 2020. https://doi.org/10.1016/j.csbj.2019.11.006

LIU, H.; DING, Y.; ZHOU, Y.; JIN, W.; XIE, K.; CHEN, L.-L. CRISPR-P 2.0: an improved CRISPR-Cas9 tool for genome editing in plants. **Molecular Plant**, v. 10, n. 3, p. 530–532, 2017. https://doi.org/10.1016/j.molp.2017.01.003

LIU, H.; WEI, Z.; DOMINGUEZ, A.; LI, Y.; WANG, X.; QI, L. S. CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation. **Bioinformatics**, v. 31, p. 3676–3678, 2015. https://doi.org/10.1093/bioinformatics/btv423

LONG, L.; GUO, D.-D.; GAO, W.; YANG, W.-W.; HOU, L.-P.; MA, X.-N.; MIAO, Y.-C.; BOTELLA, J. R.; SONG, C.-P. Optimization of CRISPR/Cas9 genome editing in cotton by improved sgRNA expression. **Plant Methods**, v. 14, n. 85, 2018. https://doi.org/10.1186/s13007-018-0353-0

LOWDER, L. G.; ZHANG, D.; BALTES, N. J.; PAUL, J. W.; TANG, X.; ZHENG, X.; VOYTAS, D. F.; HSIEH, T. F.; ZHANG, Y.; QI, Y. A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. **Plant Physiology**, v. 169, n. 2, p. 971–985, 2015. https://doi.org/10.1104/pp.15.00636

LOWDER, L. G.; ZHOU, J.; ZHANG, Y.; MALZAHN, A., ZHONG, Z., HSIEH, T. F., VOYTAS, D. F.; ZHANG, Y.; QI, Y. Robust transcriptional activation in plants using multiplexed CRISPR-Act2.0 and mTALE-Act Systems. **Molecular Plant**, v. 11, n. 2, p. 245-256, 2018. https://doi.org/10.1016/j.molp.2017.11.010

MAHFOUZ, M. M.; LI, L.; SHAMIMUZZAMAN, M.; WIBOWO, A.; FANG, X.; ZHU, J. K. De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. P **Proceedings of the National Academy of Sciences**, v. 108, n. 6, 2623–2628, 2011. https://doi.org/10.1073/pnas.1019533108

MANS, R.; van ROSSUM, H. M.; WIJSMAN, M.; BACKX, A.; KUIJPERS, N. G. A.; van den BROEK, M.; DARAN-LAPUJADE, P.; PRONK, J. T.; van MARIS, A. J. A., DARAN, J.-M. G. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. **FEMS Yeast Research**, v. 15, n. 2, 2015. https://doi.org/10.1093/femsyr/fov004

MAO, C.; BROWN, C. R.; GRIESENBECK, J.; BOEGER, H. Occlusion of regulatory sequences by promoter nucleosomes in vivo. **PLoS One**, v. 6, e17521, 2011. https://doi.org/10.1371/journal.pone.0017521

MARK, C.; ABRINK, M.; HELLMAN, L. 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**, v. 18, n. 5, 381-396, 1999. https://doi.org/10.1089/104454999315277

MIKAMI, M.; TOKI, S.; ENDO, M. Comparison of CRISPR/Cas9 expression constructs for efficient targeted mutagenesis in rice. **Plant Molecular Biology**, v. 88, p. 561–572, 2015. https://doi.org/10.1007/s11103-015-0342-x

MINKENBERG, B.; ZHANG, J.; XIE, K.; YANG, Y. CRISPR-PLANT v2: an online resource for highly specific guide RNA spacers based on improved off-target analysis. **Plant Biotechnology**. v. 17, n. 1, p. 5–8, 2019. https://doi.org/10.1111/pbi.13025

MIYAGISHI, M.; TAIRA, K. U6 promoter-driven siRNAs with four uridine 3' overhangs efficiently suppress targeted gene expression in mammalian cells. **Nature Biotechnology**, v. 20, n. 5, p. 497–500, 2002. https://doi.org/10.1038/nbt0502-497

MONTAGUE, T. G.; CRUZ, J. M.; GAGNON, J. A.; CHURCH, G. M.; VALEN, E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. **Nucleic Acids Research**, v. 42, n. W401-7. https://doi.org/10.1093/nar/gku410

MORENO-MATEOS, M. A., VEJNAR, C. E., BEAUDOIN, J.-D., FERNANDEZ, J. P., MIS, E. K., KHOKHA, M. K., GIRALDEZ, A. J. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. **Nature Methods**, v. 12, p. 982-988, 2015. https://doi.org/10.1038/nmeth.3543

NAMI, F.; BASIRI, M.; SATARIAN, L.; CURTISS, C.; BAHARVAND, H.; VERFAILLIE, C. Strategies for in vivo genome editing in nondividing cells. **Trends Biotechnology**, 36, n. 8, p. 770-786, 2018. https://doi.org/10.1016/j.tibtech.2018.03.004

NG, H.; DEAN, N. Dramatic improvement of CRISPR/Cas9 editing in *Candida albicans* by increased single guide RNA expression. **mSphere**, v. 19, n. 2:e00385-16, 2017. https://doi.org/10.1128/{mSphere}.00385-16

NISHIDA, K.; ARAZOE, T.; YACHIE, N.; BANNO, S.; KAKIMOTO, M.; TABATA, M.; MOCHIZUKI, M.; MIYABE, A.; ARAKI, M.; HARA, K. Y.; SHIMATANI, Z.; KONDO, A. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. **Science**, v. 353, n. 6305, 2016. https://doi.org/10.1126/science.aaf8729

NISHIMASU, H.; RAN, F.A.; HSU, P. D.; KONERMANN, S.; SHEHATA, S. I.; DOHMAE, N.; ISHITANI, R.; ZHANG, F.; NUREKI, O. Crystal structure of Cas9 in complex with guide RNA and target DNA. **Cell**, v. 156, p. 935–949, 2014. https://doi.org/10.1016/j.cell.2014.02.001

OHTA, M.; MATSUI, K.; HIRATSU, K.; SHINSHI, H.; OHME-TAKAGI, M. Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. **Plant Cell**, v. 13, p. 1959–1968, 2001. https://doi.org/10.1105/TPC.010127

PAPIKIAN, A.; LIU, W.; GALLEGO-BARTOLOMÉ, J.; JACOBSEN, S. E. Site-specific manipulation of Arabidopsis loci using CRISPR-Cas9 SunTag systems. **Nature Communications**, v. 10, p.1-11, 2019. https://doi.org/10.1038/s41467-019-08736-7

PARK, J.; BAE, S.; KIM, J.-S. Cas-Designer: a web-based tool for choice of CRISPR-Cas9 target sites. **Bioinformatics**, v. 15, n. 31, p. 4014–4016, 2015. https://doi.org/10.1093/bioinformatics/btv537

PARK, J.; DEMPEWOLF, E.; ZHANG, W.; WANG, Z. RNA-guided transcriptional activation via CRISPR / dCas9 mimics overexpression phenotypes in Arabidopsis. **Plos One**, v. 12, n. 6, p.1-13. 2017. https://doi.org/10.1371/journal.pone.0179410

PENG, D.; TARLETON, R. EuPaGDT: a web tool tailored to design CRISPR guide RNAs for eukaryotic pathogens. **Microbial Genomics**, v. 30, n. 1(4):e000033, 2015. https://doi.org/10.1099/mgen.0.000033

PEREZ, A. R.; PRITYKIN, Y.; VIDIGAL, J. A.; CHHANGAWALA, S., ZAMPARO; L., LESLIE; C. S., VENTURA, A. GuideScan software for improved single and paired CRISPR guide RNA design. **Nature Biotechnology**, v. 35, p. 347–349, 2017. https://doi.org/10.1038/nbt.3804

PFLUEGER, C.; TAN, D.; SWAIN, T.; NGUYEN, T.; PFLUEGER, J.; NEFZGER, C.; POLO, J. M.; FORD, E.; LISTER, R. A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs. **Genome Research**, v. 28, p. 1193–1206, 2018. https://doi.org/10.1101/gr.233049.117

PIATEK, A.; ALI, Z.; BAAZIM, H.; LI, L.; ABULFARAJ, A.; AL-SHAREEF, S.; AOUIDA, M., MAHFOUZ, M. M. RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. **Plant Biotechnology,** 13, n. 4, p. 578–589, 2015. https://doi.org/10.1111/pbi.12284

PULIDO-QUETGLAS, C.; APARICIO-PRAT, E.; ARNAN, C.; POLIDORI, T.; HERMOSO, T.; PALUMBO, E.; PONOMARENKO, J.; GUIGO, R.; JOHNSON, R. Scalable design of paired CRISPR guide RNAs for genomic deletion. **PLOS Computational Biology**, v. 13, n. 3:e1005341, 2017. https://doi.org/10.1371/journal.pcbi.1005341

QI, L. S.; LARSON, M. H.; GILBERT, L. A.; DOUDNA, J. A.; WEISSMAN, J. S.; ARKIN, A. P.; LIM, W. A. Repurposing CRISPR as an RNA-γuided platform for sequence-specific control of gene expression. **Cell,** v. 152, p.1173–1183, 2013. https://doi.org/10.1016/j.cell.2013.02.022

QI, W.; ZHU, T.; TIAN, Z.; LI, C.; ZHANG, W.; SONG, R. High-efficiency CRISPR/Cas9 multiplex gene editing using the glycine tRNA-processing system-based strategy in maize. **BMC Biotechnology**, v. 16, n. 58. 2016. https://doi.org/10.1186/s12896-016-0289-2

RADZISHEUSKAYA, A.; SHLYUEVA, D.; MÜLLER, I.; HELIN, K. Optimizing sgRNA position markedly improves the efficiency of CRISPR/dCas9-mediated transcriptional repression. **Nucleic Acids Research**, v. 44, n.18:e141, 2016. https://doi.org/10.1093/nar/gkw583

RANDO, O. J.; CHANG, H. Y. Genome-wide views of chromatin structure. **Annual Review of Biochemistry**, v. 78, p. 245–271, 2009. https://doi.org/10.1146/annurev.biochem.78.071107.134639

RANDO, O. J.; WINSTON, F. Chromatin and transcription in yeast. **Genetics**, v.190, p. 351–387, 2012. https://doi.org/10.1534/genetics.111.132266

ROCA PAIXÃO, J. F.; GILLET, F. X.; RIBEIRO, T. P.; BOURNAUD, C.; LOURENÇO-TESSUTTI, I. T.; NORIEGA, D. D.; MELO, B. P. DE; DE ALMEIDA-ENGLER, J., GROSSI-DE-SA, M. F. Improved drought stress tolerance in Arabidopsis by CRISPR/dCas9 fusion with a Histone AcetylTransferase. **Scientific Reports, v.** 9, p. 1-9, 2019. https://doi.org/10.1038/s41598-019-44571-y

RYAN, O. W.; SKERKER, J. M.; MAURER, M. J.; LI, X.; TSAI, J. C.; PODDAR, S.; LEE, M. E.; DeLOACHE, W.; DUEBER, J. E.; ARKIN, A. P.; CATE, J. H. D. Selection of chromosomal DNA libraries using a multiplex CRISPR system. **eLife**, n. 3 p. e03703, 2014. https://doi.org/10.7554/{eLife}.03703

SARASUA, B. S. Synthetic transciptional repressors design based on the CRISPR-Cas technology in *N. benthamiana.* 2020. 70 f. Dissertação (Mestrado em Biotecnologia Molecular e Celular de Plantas) – Universidade Politécnica de Valência, Espanha, 2020.

SCHEP, A. N.; BUENROSTRO, J. D.; DENNY, S. K.; SCHWARTZ, K.; SHERLOCK, G.; GREENLEAF, W. J. Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture

within regulatory regions. **Genome Research**, v. 25, p. 1757-1770, 2015. https://doi.org/10.1101/gr.192294.115

SELMA, S.; BERNABÉ-ORTS, J. M.; VAZQUEZ-VILAR, M.; DIEGO-MARTIN, B.; AJENJO, M.; GARCIA-CARPINTERO, V.; GRANELL, A.; ORZAEZ, D. Strong gene activation in plants with genome-wide specificity using a new orthogonal CRISPR/Cas9-based programmable transcriptional activator. **Plant Biotechnology Journal**, v. 17, p.1703–1705, 2019. https://doi.org/10.1111/pbi.13138

SHAN, Q.; WANG, Y.; LI, J.; ZHANG, Y.; CHEN, K.; LIANG, Z.; ZHANG, K.; LIU, J.; XI, J. J.; QIU, J.-L.; GAO, C. Targeted genome modification of crop plants using a CRISPR-Cas system. **Nature Biotechnology**, v. 31, p. 686–688, 2013. https://doi.org/10.1038/nbt.2650

SHEN, M. W.; ARBAB, M.; HSU, J. Y.; WORSTELL, D.; CULBERTSON, S. J.; KRABBE, O.; CASSA, C. A.; LIU, D. R.; GIFFORD, D. K. Sherwood, R. I. Predictable and precise template-free CRISPR editing of pathogenic variants. **Nature**, v. 563, p. 646-651, 2018. https://doi.org/10.1038/s41586-018-0686-x

SINGH, R.; KUSCU, C.; QUINLAN, A.; QI, Y.; ADLI, M. Cas9-chromatin binding information enables more accurate CRISPR off-target prediction. **Nucleic Acids Research**, v. 43, n. 18, p. e118, 2015. https://doi.org/10.1093/nar/gkv575

SMITH, J. D.; SCHLECHT, U.; XU, W.; SURESH, S.; HORECKA, J.; PROCTOR, M. J.; AIYAR, R. S.; BENNETT, R.A.O.; CHU, A.; LI, Y.F.; ROY, K.; DAVIS, R. W.; STEINMETZ, L. M.; HYMAN, R. W.; LEVY, S. F.; ST ONGE, R. P. A method for high-throughput production of sequence-verified DNA libraries and strain collections. **Molecular Systems Biology**, v. 13, n. 2, p. 913, 2017. https://doi.org/10.15252/msb.20167233

SMITH, J. D.; SURESH, S.; SCHLECHT, U.; WU, M.; WAGIH, O.; PELTZ, G.; DAVIS, R. W.; STEINMETZ, L. M.; PARTS, L.; St ONGE, R. P. Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design. **Genome Biology**, v. 17, n. 45, 2016. https://doi.org/10.1186/s13059-016-0900-9

STEMMER, M.; THUMBERGER, T.; DEL SOL KEYER, M.; WITTBRODT, J.; MATEO, J. L. Cctop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. **PLoS One**, v. 10, p. e0124633, 2015. https://doi.org/10.1371/journal.pone.0124633

SULLENGER, B. A. RGEN editing of RNA and DNA: the long and winding road from catalytic RNAs to CRISPR to the clinic. **Cell**, v. 181, p. 955–960, 2020. https://doi.org/10.1016/j.cell.2020.04.050

SUN, X.; HU, Z.; CHEN, R.; JIANG, Q.; SONG, G.; ZHANG, H.; XI, Y. Targeted mutagenesis in soybean using the CRISPR-Cas9 system. **Scientific Reports, v.** 5, n. 10342, 2015. https://doi.org/10.1038/srep10342

TANENBAUM, M. E.; GILBERT, L. A.; QI, L. S.; WEISSMAN, J. S.; VALE, R. D. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. **Cell**, v. 159, n. 3, p. 635–646, 2014. https://doi.org/10.1016/j.cell.2014.09.039

TANG, X.; LIU, G.; ZHOU, J.; REN, Q.; YOU, Q.; TIAN, L.; XIN, X.; ZHONG, Z.; LIU, B.; ZHENG, X.; ZHANG, D.; MALZAHN, A.; GONG, Z.; QI, Y.; ZHANG, T.; ZHANG, Y. A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. **Genome Biology, v.** 19, n. 84, 2018. https://doi.org/10.1186/s13059-018-1458-5

TANG, X.; LOWDER, L. G., ZHANG, T.; MALZAHN, A. A.; ZHENG, X.; VOYTAS, D. F.; ZHONG, Z.; CHEN, Y.; REN, Q.; LI, Q.; KIRKLAND, E. R.; ZHANG, Y.; QI, Y. A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. **Nature Plants**, v. 3, n. 17018, 2017. https://doi.org/10.1038/nplants.2017.18

THAKORE, P. I.; D'IPPOLITO, A. M.; SONG, L.; SAFI, A.; SHIVAKUMAR, N. K.; KABADI, A. M.; REDDY, T. E.; CRAWFORD, G. E.; GERSBACH, C. A. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. **Nature Methods**, v. 12, p. 1143–1149, 2015. https://doi.org/10.1038/nmeth.3630

THYME, S. B.; AKHMETOVA, L.; MONTAGUE, T. G.; VALEN, E.; SCHIER, A. F. Internal guide RNA interactions interfere with Cas9-mediated cleavage. **Nature Communications**, v. 7, p. 11750, 2016. https://doi.org/10.1038/ncomms11750

TIAN, S.; JIANG, L.; CUI, X.; ZHANG, J.; GUO, S.; LI, M.; ZHANG, H.; REN, Y.; GONG, G.; ZONG, M.; LIU, F.; CHEN, Q.; XU, Y. Engineering herbicide-resistant watermelon variety through CRISPR/Cas9-mediated base-editing. **Plant Cell Reports**, v. 37, p. 1353–1356, 2018. https://doi.org/10.1007/s00299-018-2299-0

TIWARI, S. B.; BELACHEW, A.; MA, S. F.; YOUNG, M.; ADE, J.; SHEN, Y.; MARION, C. M.; HOLTAN, H. E.; BAILEY, A.; STONE, J. K.; EDWARDS, L.; WALLACE, A. D.; CANALES, R. D.; ADAM, L.; RATCLIFFE, O. J.; REPETTI, P. P. The EDLL motif: a potent plant transcriptional activation domain from AP2/ERF transcription factors. **Plant Journal**, v. 70, n. 5, p. 855-865, 2012. https://doi.org/10.1111/j.1365-313X.2012.04935.x

TSAI, S. Q.; ZHENG, Z.; NGUYEN, N. T.; LIEBERS, M.; TOPKAR, V. V.; THAPAR, V.; WYVEKENS, N.; KHAYTER, C.; IAFRATE, A. J.; LE, L. P.; ARYEE, M. J.; JOUNG, J. K. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. **Nature Biotechnology**, v. 33, p. 187–197, 2015. https://doi.org/10.1038/nbt.3117

VAZQUEZ-VILAR, M.; BERNABÉ-ORTS, J. M.; FERNANDEZ-DEL-CARMEN, A.; ZIARSOLO, P.; BLANCA, J.; GRANELL, A.; ORZAEZ, D. A modular toolbox for gRNA-Cas9 genome engineering in plants based on the GoldenBraid standard. **Plant Methods**, v. 12, p. 1–12, 2016. https://doi.org/10.1186/s13007-016-0101-2

WANG, M.-B.; HELLIWELL, C. A.; WU, L.-M.; WATERHOUSE, P. M.; PEACOCK, W. J.; DENNIS, E. S. Hairpin RNAs derived from RNA polymerase II and polymerase III promoter-directed transgenes are processed differently in plants. **RNA**, v. 14, n. 5, p. 903-913, 2008. https://doi.org/10.1261/rna.760908

WANG, T.; WEI, J. J.; SABATINI, D. M.; LANDER, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. **Science**, v. 343, n. 6166, p. 80-84, 2014. https://doi.org/10.1126/science.1246981

WILSON, L.O.W.; O'BRIEN, A. R.; BAUER, D. C. The current state and future of CRISPR-Cas9 gRNA design tools. **Frontiers in Pharmacology**, v. 9, p. 749, 2018. https://doi.org/10.3389/fphar.2018.00749

WU, X.; SCOTT, D. A.; KRIZ, A. J.; CHIU, A. C.; HSU, P. D.; DADON, D. B.; CHENG, A.W.; TREVINO, A. E.; KONERMANN, S.; CHEN, S.; JAENISCH, R.; ZHANG, F.; SHARP, P.A. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. **Nature Biotechnology**, v. 32, p. 670–676, 2014. https://doi.org/10.1038/nbt.2889

XIE, X.; MA, X.; ZHU, Q.; ZENG, D.; LI, G.; LIU, Y.-G. CRISPR-GE: a convenient software toolkit for CRISPR-based genome editing. **Molecular Plant Pathology**, v. 10, p. 1246–1249, 2017. https://doi.org/10.1016/j.molp.2017.06.004

XU, H.; XIAO, T.; CHEN, C.-H.; LI, W.; MEYER, C. A.; WU, Q.; WU, D.; CONG, L.; ZHANG, F.; LIU, J. S.; BROWN, M.; LIU, X. S. Sequence determinants of improved CRISPR sgRNA design. **Genome Research**, v. 25, p. 1147–1157, 2015. https://doi.org/10.1101/gr.191452.115

YOSHIOKA, S.; FUJII, W.; OGAWA, T.; SUGIURA, K.; NAITO, K. Development of a mono-promoter-driven CRISPR/Cas9 system in mammalian cells. **Scientific Reports**, v. 5, p. 18341, 2015. https://doi.org/10.1038/srep18341

ZALATAN, J. G.; LEE, M. E.; ALMEIDA, R.; GILBERT, L. A.; WHITEHEAD, E. H.; LA RUSSA, M.; TSAI, J. C.; WEISSMAN, J. S.; DUEBER, J. E.; QI, L. S.; LIM, W. A. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. **Cell**, v. 160, p. 339-350, 2015. https://doi.org/10.1016/j.cell.2014.11.052

ZETSCHE, B.; GOOTENBERG, J. S.; ABUDAYYEH, O. O.; SLAYMAKER, I. M.; MAKAROVA, K. S.; ESSLETZBICHLER, P.; VOLZ, S. E.; JOUNG, J.; van der OOST, J.; REGEV, A.; KOONIN, E. V.; ZHANG, F. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. **Cell**, v. 163, p. 759–771, 2015. https://doi.org/10.1016/j.cell.2015.09.038

ZHANG, D.; HURST, T.; DUAN, D.; CHEN, S.-J. Unified energetics analysis unravels SpCas9 cleavage activity for optimal gRNA design. **Proceedings of the National Academy of Sciences**, v. 116, p. 8693–8698, 2019a. https://doi.org/10.1073/pnas.1820523116

ZHANG, J.-H.; PANDEY, M.; KAHLER, J. F.; LOSHAKOV, A.; HARRIS, B.; DAGUR, P. K., MO, Y.-Y.; SIMONDS, W. F. Improving the specificity and efficacy of CRISPR/CAS9 and gRNA through target specific DNA reporter. **Journal of Biotechnology**, v. 189, p. 1-8, 2014. https://doi.org/10.1016/j.jbiotec.2014.08.033

ZHANG, R., LIU, J., CHAI, Z.; CHEN, S.; BAI, Y.; ZONG, Y.; CHEN, K.; LI, J.; JIANG, L.; GAO, C. Generation of herbicide tolerance traits and a new selectable marker in wheat using base editing. **Nature Plants**, v. 5, 4p. 80–485, 2019b. https://doi.org/10.1038/s41477-019-0405-0

ZHANG, T.; GAO, Y.; WANG, R.; ZHAO, Y. Production of guide RNAs in vitro and in vivo for CRISPR using ribozymes and RNA polymerase II promoters. **Bio-protocol**, v. 7, n. 4, p.e2148, 2017. https://doi.org/10.21769/BioProtoc.2148

ZHANG, X.-H.; TEE, L.Y.; WANG, X.-G.; HUANG, Q.-S.; YANG, S.-H. Off-target effects in CRISPR/Cas9-mediated genome engineering. **Molecular Therapy:** Nucleic Acids, v. 4, p. e264, 2015. https://doi.org/10.1038/mtna.2015.37

ZHANG, Y.; LAI, B. S., JUHAS, M. Recent advances in aptamer discovery and applications. **Molecules**. v. 24, 2019c. https://doi.org/10.3390/molecules24050941

ZONG, Y.; WANG, Y.; LI, C.; ZHANG, R.; CHEN, K.; RAN, Y.; QIU, J. L.; WANG, D.; GAO, C. Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. **Nature Biotechnology**, v. 35, p. 438-440, 2017. https://doi.org/10.1038/nbt.3811