

Modulação da expressão gênica em plantas via tecnologia CRISPR/dCas9

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

Introdução

A regulação da expressão gênica inclui uma diversidade de processos celulares que ocorrem de forma coordenada e em múltiplas etapas para desencadear o aumento ou a redução de um produto gênico específico. A expressão gênica pode ser induzida por estímulos endógenos e ambientais e modulada em diferentes níveis celulares, como na iniciação da transcrição, no processamento do RNA e na modificação pós-traducional da proteína.

A manipulação de genes-alvo é de primordial importância para o entendimento da função gênica e reprogramação das atividades celulares, tanto para o aprofundamento dos conhecimentos básicos sobre processos bioquímicos e moleculares, como para a intensificação de características de interesse agrônomo. Neste ponto, a precisão é essencial para que se obtenha o êxito necessário em aplicações da engenharia genética e da biologia sintética.

Nas últimas décadas, tecnologias envolvendo nucleases sítio específicas para a manipulação precisa do DNA sofreram um profundo avanço, surgindo como alternativas promissoras para a indução de mutações sítio-dirigidas e controle fino da expressão gênica. Entre essas tecnologias, destacam-se as de edição de genomas, como a da nuclease dedo de zinco (ZFN, do inglês *Zinc Finger Nuclease*), a de nucleases com efetores do tipo ativador transcricional (TALENs, do inglês *Transcription Activator-Like Effector Nucleases*) e, mais recentemente, a tecnologia CRISPR/Cas (do inglês *Clustered Regularly Interspaced Short Palindromic Repeats*) associada à nuclease Cas. Esta última tem seu caráter revolucionário, sobretudo pela sua especificidade, universalidade e relativa simplicidade (Pickar-Oliver; Gersbach, 2019). Além disso, CRISPR/Cas é uma ferramenta flexível, passível de modificações, o que contribui para seu contínuo aprimoramento e diversifica suas aplicações nos estudos das funções celulares e na biotecnologia.

Este capítulo aborda a tecnologia CRISPR pelo uso da variante *dead Cas9* (CRISPR/dCas9), com ênfase no seu uso para a modulação da expressão gênica em plantas. São apresentadas e discutidas as principais estratégias atualmente utilizadas, para as quais foi proposto um roteiro, considerando os principais aspectos para um bom desenho experimental.

O sistema dCas9

O sistema CRISPR/Cas9 foi identificado primeiramente em *Streptococcus pyogenes* e, posteriormente, em muitas outras bactérias, assim como na maioria das arqueias, como um sofisticado sistema imunológico adaptativo, guiado por RNAs codificados pelo locus CRISPR e pelas proteínas Cas (do inglês *CRISPR associated protein*), a fim de fornecer “imunidade adquirida”, principalmente, contra bacteriófagos (Hsu et al., 2014; Zhang et al., 2014). Sua aplicação na engenharia genética foi possível, principalmente, pela profunda elucidação da estrutura da Cas9, em conjunto com o RNA guia (gRNA).

Estruturalmente, a Cas9 de *S. pyogenes* pode ser considerada uma proteína multifuncional grande (1.368 resíduos de aminoácidos) formada por cinco domínios e com atividade nucleásica (Jiang; Doudna, 2017). A estrutura da apoenzima Cas9 é formada pelo lóbulo de reconhecimento alfa-helicoidal (REC) e pelo lóbulo nucleásico (NUC). Este último contém os domínios nucleásicos HNH conservado e RuvC tripartido, além de um domínio C-terminal mais variável (CTD) (Figura 1) (Jinek et al., 2014; Nishimasu et al., 2014).

No sistema CRISPR/Cas9 nativo, o gRNA maduro é composto por duas moléculas independentes de pequenos RNAs não codificadores, que interagem entre si por complementaridade de bases: 1) crRNA (do inglês *CRISPR RNA*), região onde ocorre o reconhecimento da sequência de DNA-alvo; e 2) tracrRNA (do inglês *transactivating crRNA*), importante para o ancoramento do gRNA na Cas9 (Jinek et al., 2012). Estudos in vitro determinaram a sequência mínima necessária, tanto do crRNA quanto do tracrRNA (sendo denominado também de RNA scaffold), para sua utilização em modelos sintéticos. Essas duas sequências mínimas foram unidas, criando assim o sgRNA (do inglês *single guide RNA*) (Figura 1) (Anders et al., 2014; Dang et al., 2015; Hsu et al., 2013; Jinek et al., 2012; Nishimasu et al., 2014). Após o reconhecimento da sequência de DNA-alvo, por meio da sua complementariedade ao gRNA/sgRNA, a Cas9 cliva o DNA dupla fita (dsDNA) em três pares de base a montante da sequência PAM (do inglês *Protospacer Adjacent Motif*), por meio de seus dois domínios nucleásicos distintos.

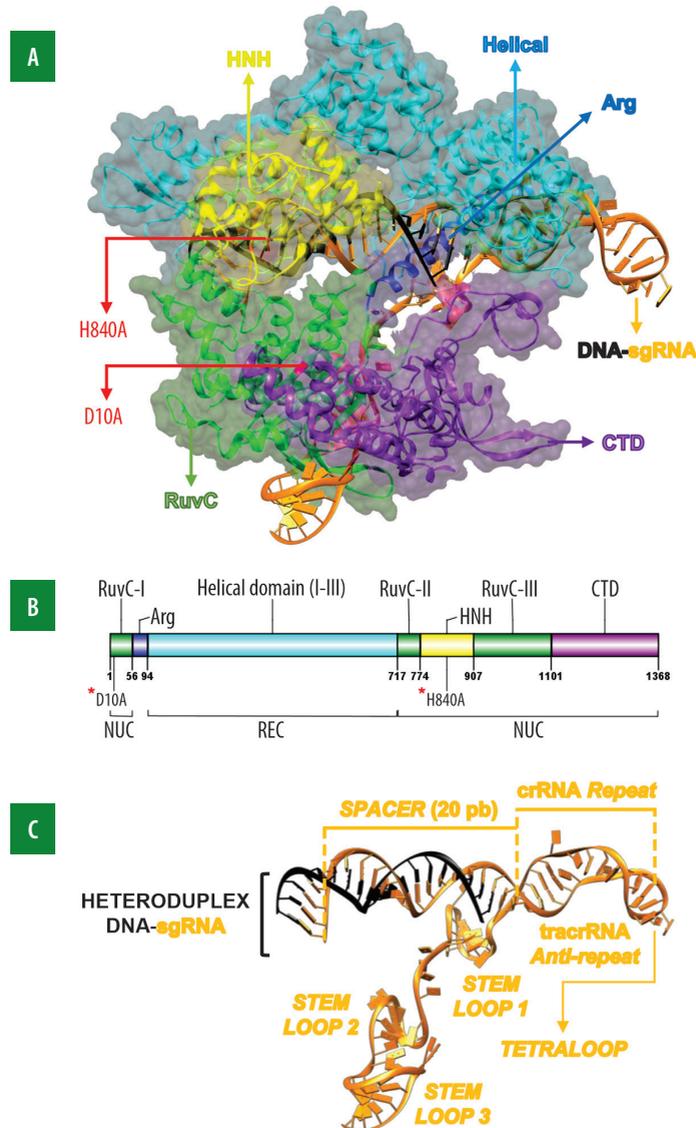


Figura 1. Estrutura da Cas9/dCas9 e do sgRNA. Modelo estrutural (A) e representação em diagrama (B) dos domínios funcionais da Cas9/dCas9 associada com sgRNA interagindo com a fita do DNA-alvo (fita complementar somente, representada em preto). Além do domínio C-terminal variável CTD, são apresentados no lóbulo NUC (sigla para *nucleásico*), os domínios RuvC (tripartido) e HNH, essenciais para a clivagem das duas fitas do DNA-alvo. São representados em vermelho (A) e com asterisco (B) mutações tanto no domínio RuvC (D10A) quanto no domínio HNH (H840A), características da dCas9. Isoladas, essas mutações conferem atividade nicase para a Cas9, mas juntas abolem completamente a sua atividade nucleásica. O lóbulo REC (sigla para *reconhecimento*) é representado exclusivamente pelo domínio Helical. A estrutura do sgRNA (C) é composta pelo crRNA (*spacer* + *crRNA repeat*) e RNA scaffold ou *tracrRNA* (neste esquema, o *tracrRNA anti-repeat* é ligado a *crRNA repeat* por um *tetraloop*, seguido de três *stem loops*). As representações estruturais apresentadas são baseadas em modelos depositados em banco de dados públicos (PDB ID: 4008), disponível em: <https://www.rcsb.org/>

Enquanto o domínio HNH cliva a fita de DNA complementar à sequência de sgRNA (fita-alvo), o domínio RuvC é responsável pela clivagem da fita de DNA oposta à fita complementar (fita não alvo) (Jiang; Doudna, 2017). Estudos mostraram que mutações específicas nos domínios HNH (a exemplo, H840A) ou RuvC (a exemplo, D10A) são capazes de converter a Cas9 em uma nicase (enzima que cliva um dsDNA em somente uma das fitas), enquanto mutações em ambos os domínios extinguem a sua atividade endonucleásica, sem interferir na sua capacidade de reconhecimento específico do DNA guiado por RNA, sendo assim designada *dead Cas9* ou dCas9 (Tabela 1) (Jiang; Doudna, 2017; Jinek et al., 2014, 2012).

Aplicações do sistema dCas9 em plantas

As aplicações do sistema CRISPR vão além daquelas relacionadas à indução de mutações sítio-dirigidas no DNA pela clivagem da fita dupla do DNA e ativação do sistema de reparo. A proteína dCas9 oferece uma plataforma única, multifuncional e dinâmica para o recrutamento de proteínas com funções diversas a locais específicos do genoma para promover, por exemplo, o controle da regulação da transcrição, a edição epigenômica, a edição de bases e o imageamento do genoma.

A edição de bases vem se tornando uma potente ferramenta na biotecnologia vegetal para a indução de mutações pontuais direcionadas. O desenvolvimento de variedades tolerantes a herbicidas para as culturas do trigo, arroz, milho e melancia, por exemplo, foi possível pela introdução de mutações pontuais nos genes que codificam as enzimas acetolactato sintase e/ou da acetil-coenzima A, de forma que o produto final não é transgênico (Kuang et al., 2020; Li et al., 2019; Tian et al., 2018; Zhang et al., 2019b). Sistemas editores de bases requerem a fusão de variantes da Cas9, dCas9 ou nCas9, guiadas por um sgRNA, fusionadas a uma citosina (C) desaminase, que induz a conversão de C-G (Guanina) para T-A (Timina-Adenina) (CBE, do inglês *cytosine base editor*), ou a uma A desaminase, que promove a conversão de A-T para G-C (ABE, do inglês *adenine base editor*) (Gaudelli et al., 2017; Komor et al., 2016; Nishida et al., 2016). Esses sistemas foram adaptados com sucesso para uso em plantas, inclusive em sistemas multiplex, em que mais de um gene é alvo para indução da mutação simultaneamente, com alta eficiência, especificidade e sem ocorrência de *indels* no local da edição. Resultados mais satisfatórios foram obtidos com o uso da nCas9 (Hua et al., 2018; Kang et al., 2018; Zong et al., 2017).

Outra potente aplicação de CRISPR/dCas9 é o imageamento de genoma, usado para a visualização da organização espacial e de interações temporais da cromatina em tempo real. Para tanto, a proteína dCas9 é fusionada a proteínas fluorescentes, como a GFP (do inglês *Green Fluorescent Protein*, proteína verde fluorescente),

Tabela 1. Sumário de ensaios de mutagênese com a nuclease Cas9 de *Streptococcus pyogenes* (SpCas9) (UniProtKB - Q99ZW2).

Posição	Mutação ⁽¹⁾	Consequência(s)	Referência
10	D → A	DNA-alvo não complementar ao crRNA não é clivado; atividade nicase	Nishimasu et al. (2014)
15	S → A	Redução da eficiência de clivagem do DNA-alvo	Nishimasu et al. (2014)
66	R → A	Redução significativa da eficiência de clivagem do DNA-alvo	Nishimasu et al. (2014)
70	R → A	Ausência de clivagem do DNA-alvo	Nishimasu et al. (2014)
74	R → A	Redução significativa da eficiência de clivagem do DNA-alvo	Nishimasu et al. (2014)
78	R → A	Redução moderada da eficiência de clivagem do DNA-alvo	Nishimasu et al. (2014)
97 – 150	Deleção	Ausência de atividade nucleásica	Nishimasu et al. (2014)
165	R → A	Redução moderada da eficiência de clivagem do DNA-alvo	Nishimasu et al. (2014)
175 – 307	Deleção	Perda de 50% da atividade nucleásica	Nishimasu et al. (2014)
302 – 409	Deleção	Ausência de atividade nucleásica	Nishimasu et al. (2014)
475 – 477	PWN → AAA	Leve redução na clivagem e na interação com o DNA-alvo	Jinek et al. (2014)
762	E → A	Clivagem de somente uma das fitas do DNA-alvo, provavelmente a não complementar ao crRNA	Nishimasu et al. (2014)
840	H → A	DNA-alvo complementar ao crRNA não é clivado; atividade nicase	Anders et al. (2014)
854	N → A	Redução da eficiência de clivagem do DNA-alvo	Nishimasu et al. (2014)
863	N → A	Clivagem de somente uma das fitas do DNA-alvo, provavelmente a complementar ao crRNA	Nishimasu et al. (2014)
982 – 983	HH → AA	Processamento correto o sgRNA	Fonfara et al. (2014)
982	H → A	Redução da eficiência de clivagem do DNA-alvo	Nishimasu et al. (2014)
983	H → A	Clivagem de somente uma das fitas do DNA-alvo, provavelmente a não complementar ao crRNA	Nishimasu et al. (2014)
986	D → A	Clivagem de somente uma das fitas do DNA-alvo, provavelmente a não complementar ao crRNA	Nishimasu et al. (2014)

Continua...

Tabela 1. Continuação.

Posição	Mutação ⁽¹⁾	Consequência(s)	Referência
1099 – 1368	Deleção	Nenhuma atividade nucleásica é detectada	Nishimasu et al. (2014)
1125 – 1127	DWD → AAA	Nenhuma mudança na eficiência de clivagem, mas uma leve redução na interação com o DNA-alvo	Jinek et al. (2014)
1132	G → C	Inativação da proteína (provavelmente)	Jiang et al. (2013)
1133 – 1135	RKR → AKA	Quase perda completa da atividade nucleásica	Anders et al. (2014)
1133	R → A	Redução drástica da interação com o DNA-alvo, mas uma leve redução na eficiência de clivagem	Anders et al. (2014)
1135	R → A	Redução drástica da interação com o DNA-alvo, mas uma leve redução na eficiência de clivagem	Anders et al. (2014)

⁽¹⁾ Abreviação dos aminoácidos: A – alanina; D – ácido aspártico; E – ácido glutâmico; G – glicina; H – histidina; K – lisina; N – asparagina; P – prolina; R – arginina; S – serina; W – triptofano.

resultando, por exemplo, em uma visualização robusta de repetições teloméricas e suas movimentações ou de múltiplos locus genômicos em células vivas, além de permitir a visualização de interações proteínas-DNA (Chen et al., 2013; Dreissig et al., 2017). Em uma otimização do método, Khosravi et al. (2020) fusionaram proteínas fluorescentes a proteínas capazes de se ligar a aptâmeros inseridos no sgRNA, de forma a recrutar um maior número de proteínas repórteres, o que resultou em maior qualidade da marcação quando comparada à do sistema original.

O sistema CRISPR/dCas9 é uma ferramenta em potencial no contexto de regulação transcricional de genes codificadores e não codificadores de proteínas. Trata-se de uma ferramenta flexível, que pode ser usada tanto para a ativação (CRISPRa, ativador) como para a repressão (CRISPRi, interferente) da transcrição, reversível, pois não modifica permanentemente o genoma, e com capacidade de regulação em escala multigênica, uma vez que diversos genes podem ter sua expressão modulada simultaneamente e em sentidos opostos. Nessa metodologia, a dCas9 e/ou o sgRNA são fusionados a moduladores transcricionais (ModT), que atuam tanto na ativação quanto na repressão da transcrição. São proteínas ou domínios de proteínas capazes de se ligar ao DNA e recrutar elementos transcricionais-chave para o controle da expressão de genes.

A alternativa de fusão de ModT à dCas9 surgiu como necessidade de potencializar os efeitos da tecnologia CRISPRi (CRISPR interferente), inicialmente desenvolvida a partir da obstrução da transcrição, pela interferência física na ligação da RNA polimerase e de fatores de transcrição ao sítio de iniciação da transcrição (TSS,

do inglês *Transcription Start Site*), ou no processo de alongação (Qi et al., 2013) (Figura 2A). A estratégia de fusão de ModT à dCas9 não só permitiu a otimização da repressão transcricional, como também viabilizou o desenvolvimento de metodologia para a ativação (Gilbert et al., 2013). Dessa forma, CRISPR/dCas9 pode ser vista como uma plataforma genérica e universal, uma vez que é capaz de promover a ativação e repressão transcricional em diferentes níveis, além de ser transponível a diferentes espécies. Nesse sentido, um mesmo repressor ou ativador transcricional pode ser usado para a regulação da transcrição em células de diferentes espécies, com eficiências variando em função da região-genômica-alvo, desenho do sgRNA, estratégia de entrega (método de transformação) e do sistema biológico estudado.

Quando ModT atuam na regulação epigenética da cromatina, como é o caso de acetil e metiltransferases, a técnica é denominada de edição epigenômica (Hilton et al., 2015; Thakore et al., 2015). Em plantas, esse método foi utilizado para a indução da tolerância à seca em *Arabidopsis thaliana*, pela superexpressão do gene *AtAREB1* (do inglês *Abscisic Acid-Responsive Element Binding Protein 1*), utilizando a dCas9 fusionada a uma acetiltransferase 1 (Roca Paixão et al., 2019) e também para a alteração do tempo de floração, por meio da regulação epigenética do gene *AtFT* (do inglês *Flowering Locus T*) em *A. thaliana*, utilizando o sistema CRISPR/dCas9 associado a uma acetiltransferase ou a uma metiltransferase (Lee et al., 2019).

A seguir são apresentados os principais ModT, incluindo ativadores transcricionais (AtvT) e repressores transcricionais (RepT), e as principais estratégias para a modulação da expressão gênica por meio do sistema CRISPR/dCas9.

Estratégias para a modulação da expressão gênica via CRISPR/dCas9

Aumento da expressão gênica (CRISPRa)

Fusão de ativadores transcricionais à proteína dCas9 (dCas9:AtvT)

Esta pode ser considerada a estratégia pioneira para a ativação da transcrição em plantas pelo sistema CRISPR/dCas9. A sequência de um ativador transcricional (AtvT) é fusionada in tandem à região C-terminal da proteína dCas9. Esse complexo é direcionado especificamente a uma região promotora do gene-alvo pelo sgRNA (Figura 2B). Ativadores como o VP64, EDLL, TAD e HAT (Tabela 2) foram utilizados

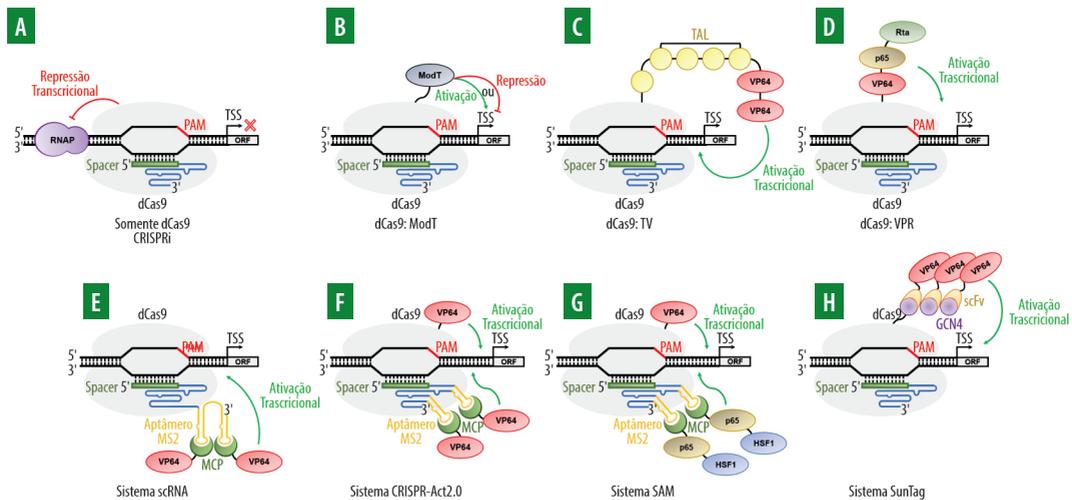


Figura 2. Estratégias de uso da tecnologia CRISPR/dCas9 na modulação da expressão gênica. (A) Repressão (CRISPRi, interferente) via obstrução da RNA Polimerase (RNAP) pelo complexo dCas9-sgRNA; (B) Ativação ou repressão pela fusão em tandem de moduladores transcricionais (ModT) à dCas9; (C) Ativação pela combinação de ModTs (VP64 e TAD) fusionados à dCas9 (dCas9:VT); (D) Ativação pela combinação de ModTs (VP64:p65:Rta, VPR) fusionados à dCas9; (E) Sistema RNA *scaffold* (scRNA), em que o sgRNA contendo um aptâmero MS2 em sua extremidade 3' recruta ModT (VP64) via fusão a MCP (MS2 Coat Protein); (F) Sistema CRISPR-Act2.0 em que dCas9:VP64 é utilizada em combinação a um sgRNA modificado com dois aptâmeros MS2 que recrutam ModT via MCP; (G) Sistema SAM, otimização do sistema CRISPR-Act2.0, em que dCas9 é fusionada a uma combinação de ModTs (p65-HSF1); (H) Sistema SunTag que consiste na fusão da dCas9 a repetições in tandem do peptídeo GCN4 que recruta ModT (VP64) via anticorpo (scFv) que se liga a GCN4. Siglas: TSS – *transcription start site*; ORF – *open reading frame*; PAM – *protospacer adaptor motif*.

com sucesso para o aumento da expressão de genes codificadores e não codificadores (microRNAs) em plantas, podendo ser potencializada com o uso de múltiplos sgRNAs, para diferentes regiões de uma mesma região promotora (Lowder et al., 2015; Piatek et al., 2015; Roca Paixão et al., 2019; Vazquez-Vilar et al., 2016). No entanto, o número ideal de sgRNAs varia em função do sistema dCas9:AtvT utilizado, e seu uso excessivo pode levar ao aumento de alvos não específicos (*off-targets*). Ademais, o número de sgRNAs e dCas9 pode chegar ao ponto de saturação e levar à atenuação da transcrição devido à interferência estrutural da ligação física de múltiplos complexos dCas9-sgRNA a uma mesma região do DNA-alvo (Piatek et al., 2015; Roca Paixão et al., 2019).

Outro ponto a ser considerado é a posição dos sgRNAs. Em geral, os sgRNAs direcionados à região a montante do TATA box e do TSS correlacionam-se positivamente com a ativação gênica mediada por dCas9, provavelmente devido à interferência com o complexo de iniciação transcricional, enquanto o posicionamento da dCas9 a jusante ou nas proximidades do TATA box afeta negativamente a expressão de

Tabela 2. Principais moduladores transcricionais utilizados no sistema CRISPR/dCas em plantas.

Modulador	Tipo ⁽¹⁾	Origem	Referência
VP64	AtvT	Repetição tetramérica artificial do domínio mínimo de ativação da proteína VP16 do Herpes-Vírus Simples	Beerli et al. (1998)
EDLL	AtvT	Domínio do ativador transcricional AtERF98 (fator de resposta ao etileno) de <i>Arabidopsis thaliana</i>	Tiwari et al. (2012)
TAD	AtvT	Domínio de ligação ao DNA do tipo ativador transcricional Hax3 de <i>Xanthomonas campestris</i>	Mahfouz et al. (2011)
TV	AtvT	6x TAD + 2x VP64	Li et al. (2017)
VPR	AtvT	Ativador tripartido contendo VP64, p65 (subunidade <i>trans</i> -ativadora NF-kB) e Rta	Chavez et al. (2015)
HAT	AtvT	Núcleo catalítico da histona acetiltransferase 1 de <i>A. thaliana</i>	Roca Paixão et al. (2019)
p65-HSF	AtvT	Subunidade <i>trans</i> -ativadora NF-kB fusionada ao domínio de ativação do fator heat-shock (HSP1) de humanos	Konermann et al. (2015)
SRDX	RepT	Domínio de repressão do motivo EAR (motivo anfífilo de repressão associado a ERF)	Ohta et al. (2001)
KYP	RepT	Domínio SET da metiltransferase H3K9 KRYPTONITE (KYP) de <i>A. thaliana</i>	Jackson et al. (2002)
BRD	RepT	Domínio de repressão B3 de fatores de transcrição de <i>A. thaliana</i>	Ikeda e Ohme-Takagi (2009)
KRAB	RepT	Domínio repressor de fatores de transcrição que contém o motivo do tipo <i>Krüppel</i> dedo de zinco de tetrápodes	Mark et al. (1999)

⁽¹⁾ AtvT – Ativador Transcricional; RepT- Repressor Transcricional

genes, provavelmente devido ao impedimento físico da maquinaria de transcrição pela dCas9 (Deaner; Alper, 2017; Farzadfard et al., 2013). No entanto, a distância ótima do sgRNA a montante do TSS para o máximo de ativação gênica pode variar em função do tipo de AtvT utilizado (Piatek et al., 2015).

A escolha da fita de DNA (senso ou antisenso) também influencia a eficiência do método. Apesar de não haver um consenso, sgRNAs situados na fita senso e próximos ao sítio de início da transcrição geralmente são mais eficientes em ativar a transcrição, uma vez que aqueles localizados na fita antisenso podem impedir a ligação da RNA polimerase e o início da transcrição ou levarem a uma terminação prematura da transcrição (Howe et al., 2017; Piatek et al., 2015).

Dessa forma, quanto maior for o conhecimento teórico e experimental a respeito da estrutura e funcionalidade da região promotora-alvo (composição e participação de *cis*-elementos, recrutamento de fatores de transcrição, posição do TSS, entre outros), maior será a probabilidade de sucesso experimental da tecnologia dCas9. Por isso, sugere-se fortemente que, em um primeiro momento, diferentes sgRNAs, em diferentes combinações, posições e orientações, sejam testados em sistemas de expressão transiente.

Outra variação dessa estratégia é a fusão de múltiplos ativadores à sequência dCas9. No sistema dCas9:VP, por exemplo, a proteína dCas9 foi fusionada a seis módulos TAD seguidos de dois VP64, o que proporcionou um aumento de expressão de até 190 vezes em comparação à expressão basal do gene endógeno, tanto em monocotiledôneas como em dicotiledôneas (Figura 2C). O sistema dCas9:VP mostrou-se eficiente em um sistema multiplex, promovendo o aumento da expressão simultânea de três genes, e também foi eficaz em um sistema *DNA-free*, utilizando complexos de ribonucleoproteínas (RNP), porém com menor nível de ativação gênica (Tabelas 2 e 3) (Li et al., 2017). No entanto, os efeitos da fusão de mais de um ativador à dCas9 nem sempre são aditivos. A fusão dCas9:VP64:EDLL, por exemplo, resultou em uma modesta ativação de genes endógenos, inferior ao clássico sistema dCas9:VP64, muitas vezes considerado fraco e ineficiente em células vegetais (Tabela 3) (Li et al., 2017; Lowder et al., 2018). Deve-se levar em consideração que a fusão de um número excessivo de ativadores pode desencadear a degradação de mRNAs em função do alto número de sequências repetitivas (Li et al., 2017).

Uma observação interessante é a de que existe uma correlação negativa entre a expressão endógena basal do gene e a magnitude do aumento da expressão via dCas9:AtvT, uma vez que genes endógenos com expressão basal menor são mais passíveis de serem mais fortemente induzidos por esse sistema, em comparação com aqueles com transcrição basal mais intensa (Li et al., 2017; Lowder et al., 2018).

Modificação da estrutura de sgRNAs para o recrutamento de ativadores

Mesmo com todos os aperfeiçoamentos, a estratégia de fusão de ModT à proteína dCas9 pode não ser capaz de modular satisfatoriamente os níveis de expressão de genes-alvo. Além disso, a relação um para um, inerente em um sistema dCas9-sgRNA, restringe a programação de circuitos de genes, baseados em transcrição multigênica a apenas uma direção de regulação: repressão ou ativação. Isso não corresponde à complexidade e sofisticação subjacentes às redes de transcrição nativas. No entanto, em analogia à fusão de ModTs com a dCas9, a engenharia do próprio sgRNA provou ser uma plataforma modular e regulável para diversificar

Tabela 3. Estratégias de uso do sistema CRISPR/dCas9 para a ativação da expressão gênica em plantas.

Espécie vegetal	Gene-alvo	Estratégia CRISPR/dCas9	Sistema de transformação	Nível de ativação	Referência
<i>A. thaliana</i>	AtPAP1	dCas9:VP64		7,0 x	
		dCas9:VP64-EDLL		4,0 x	Lowder et al. (2015)
		dCas9:VP64-MS2-VP64	Plantas transgênicas	45,0 x	Lowder et al. (2018)
		dCas9:VP64-MS2-EDLL		30,0 x	Park et al. (2017)
		dCas9:VP64- MS2-p65-HSF		3,0 x	
	AtFIS2	dCas9:VP64		200,0 x	
		dCas9:VP64-EDLL		2,5 x	
		dCas9:VP64-MS2-VP64	Plantas transgênicas	1500,0 x	Lowder et al. (2018)
		dCas9:VP64-MS2-EDLL		30,0 x	
	miR319	dCas9-VP64		7,5 x	Lowder et al. (2015);
		dCas9:VP64-MS2-VP64	Plantas transgênicas	6,0 x	Lowder et al. (2018)
		dCas9:VP64-MS2-VP64 + 3sgRNA		2,5 x	
AtUULC1	dCas9:VP64-MS2-VP64 + 3 sgRNAs	Plantas transgênicas	40,0 x	Lowder et al. (2018)	
	dCas9:TV	Plantas transgênicas	139,0 x	Li et al. (2017)	
AtWRKY	dCas9:TV + RNP	Protoplasto/PEG	11,7 x		
	dCas9:TV		510,0 x		
pWRKY::luciferase	dCas9:VP64-MS2-TV	Protoplasto/PEG	170,0 x	Li et al. (2017)	
	dCas9:VP64-		6,7 x		
AtRLP23	dCas9:TV	Plantas transgênicas	30,0 x		
	dCas9:TV	Protoplasto/PEG	44,0 x	Li et al. (2017)	
	dCas9:TV + RNP	Protoplasto/PEG	9,3 x		

Continua...

Tabela 3. Continuação.

Espécie vegetal	Gene-alvo	Estratégia CRISPR/dCas9	Sistema de transformação	Nível de ativação	Referência
<i>A. thaliana</i>	<i>AtWRKY</i>			80,0 x	
	<i>AtRLP23</i>	dCas9:TV multiplex	Protoplasto/PEG	37,0 x	Li et al. (2017)
	<i>AtCDG1</i>			192,0 x	
	<i>AtAREB1</i>	dCas9:HAT	Plantas transgênicas	1,7 x	Roca Paixão et al. (2019)
	<i>AtFWA</i> <i>AtAPETALA3</i> <i>AtCLAVATA3</i> <i>Evadé transposon</i>	dCas9:SunTag-VP64	Plantas transgênicas	1,5 x 500,0 x 100,0 x 4000,0 x	Papikian et al. (2019)
<i>N. benthamiana</i>	<i>AtAVP1</i>	dCas9:VP64-MS2-p65-HSF SAM	Plantas transgênicas	5,0 x	Park et al. (2017)
	<i>AtFT</i>	Cas9-MS2-VP64	Plantas transgênicas	30,0 x	Lee et al. (2019)
	<i>NbPDS</i>	dCas9:EDLL + 3 sgRNAs dCas9:TAD + 3 sgRNAs	Agroinfiltração	3,4 x 4,0 x	Piatek et al. (2015)
	<i>pNOS::luciferase</i>	dCas9:VP64	Agroinfiltração	3,0 x	Vazquez-Vilar et al. (2016)
		dCas9:EDLL		3,0 x	
	dCas9:EDLL-MS2-VP64 dCas9:SunTag-VP64 dCas9:SunTag -EDLL		4,0 x 3,5 x 3,0 x	Selma et al. (2019)	

Continua...

Tabela 3. Continuação.

Espécie vegetal	Gene-alvo	Estratégia CRISPR/dCas9	Sistema de transformação	Nível de ativação	Referência
<i>O. sativa</i>	<i>OsGW7</i>	dCas9:TV	Protoplasto/PEG	79,0 x	Li et al. (2017)
	<i>OsER1</i>	dCas9:TV	Protoplasto/PEG	62,0 x	Li et al. (2017)
		dCas9:TV + RNP		13,0 x	
	<i>Os03g01240</i>	dCas9:VP64	Protoplasto/PEG	2,0 x	Lowder et al. (2018)
		dCas9:VP64-MS2-VP64 + 2 sgRNAs		3,0 x	
	<i>Os04g39780</i>	dCas9:VP64	Protoplasto/PEG	0,0	Lowder et al. (2018)
		dCas9:VP64-MS2-VP64 + 2 sgRNAs		4,0 x	
	<i>Os03g01240</i>	dCas9:VP64 multiplex	Protoplasto/PEG	2,0 x	Lowder et al. (2018)
				2,0 x	
				2,0 x	
<i>Os04g39780</i>	dCas9:VP64-MS2-VP64 multiplex	Protoplasto/PEG	3,0 x	Lowder et al. (2018)	
			6,5 x		
			2,5 x		

não apenas as posições a serem reconhecidas nos genomas-alvo, mas também diversificar a função da regulação transcricional mediada pelo sistema CRISPR. Essa estratégia é baseada na modificação da estrutura do sgRNA pela adição de aptâmeros de RNA, tornando-o capaz de recrutar à região promotora-alvo proteínas adaptadoras fusionadas a ModTs. A ação desse maior número de ModT ampliou a magnitude dos níveis de ativação propiciada pela modulação clássica mediada pela dCas9:ModTs (Konermann et al., 2015; Zalatan et al., 2015).

Os primeiros sistemas que desenharam sgRNAs com aptâmeros de RNA que interagem com proteínas aproveitaram sua extremidade 3' para inserir sequências reconhecidas por proteínas específicas como a MCP, que, devidamente fusionadas com ModTs, interagem de maneira ortogonal com os aptâmeros como o MS2 fusionados à porção *RNA scaffold* de sgRNAs, permitindo a especificidade do alvo no genoma, juntamente com a função moduladora. Esse sistema foi denominado *scaffold RNA (scRNA)* (Figura 2E) (Jensen et al., 2017; Kiani et al., 2015; Zalatan et al., 2015).

Posteriormente, no desenvolvimento de um segundo sistema, foi mostrado que a região do *RNA scaffold* assume uma estrutura secundária específica, formada por um *tetraloop* e 3 *stem loops*, importante para a ancoragem do sgRNA na Cas9/dCas9 (Figura 1). Ficou evidente que o *tetraloop* e o *stem loop 2* se projetam para fora do complexo da dCas9-sgRNA, com os quatro pares de bases distais de cada haste completamente livres de interações com as cadeias laterais dos aminoácidos da proteína. Estudos de funcionalidade demonstraram que substituições e deleções nas regiões do *tetraloop* e do *stem loop 2* não afetam a função de reconhecimento do DNA-alvo pela dCas9 e que podem tolerar a adição de aptâmeros de RNA (por exemplo, MS2), que, por sua vez, são reconhecidos por proteínas específicas (por exemplo, MCP). Quando fusionadas com ModTs, essas proteínas específicas recrutam um maior número de ModT para a região-alvo no genoma, reconhecida pelo complexo dCas9-sgRNA. A tecnologia de combinação de ModTs fusionados à dCas9 e à estrutura secundária do sgRNA foi nomeada como mediador de ativação sinérgica (SAM, do inglês *Synergistic Activation Mediator*). Um sistema similar adaptado para plantas com utilização do AtvT VP64 foi denominado CRISPR-Act2.0 (Figura 2 F e G)(Konermann et al., 2015; Li et al., 2017; Lowder et al., 2018; Nishimasu et al., 2014).

Em plantas, esses sistemas já tiveram eficácia comprovada com o uso de variadas combinações de AtvT (Tabela 2) (Lee et al., 2019; Lowder et al., 2018; Selma et al., 2019). O sistema CRISPR-Act2.0 apresentou capacidade superior de ativação de três genes simultaneamente, em esquema multiplex, quando comparado ao sistema clássico dCas9:VP64 (Tabela 3) (Lowder et al., 2018). A eficiência do siste-

ma, no entanto, pode variar em função do gene-alvo, tipo de ativadores utilizados e quantidade de sgRNAs. Por exemplo, ao mesmo tempo em que o sistema dCas9:VP64-MS2:VP64 foi capaz de ativar a expressão do gene *AtFIS2* de *A. thaliana* em até 1.500 vezes, estratégia similar, com utilização de dCas9:VP64-MS2:EDLL para a ativação desse mesmo gene, apresentou resultado inferior ao obtido pelo sistema clássico dCas9:VP64 (Tabela 3) (Lowder et al., 2018). Esses mesmos autores lançaram a hipótese de que para muitos genes deve haver um limite máximo de aumento de expressão e que níveis superiores podem desencadear o silenciamento gênico pós-transcricional, pela produção de pequenos RNAs interferentes. Dessa forma, estratégias que deveriam otimizar o sistema de ativação gênica via CRISPR/dCas9 nem sempre resultam no efeito desejado, como exemplificado na ativação de miR319 em *A. thaliana* (Tabela 3) (Lowder et al., 2018, 2015).

Sistema SuperNova Tag (SunTag) para ativação gênica

Esta estratégia foi delineada primeiramente em células de mamíferos na tentativa de otimizar o aumento da expressão gênica via dCas9, visando potencializar o número de ativadores na região promotora-alvo. Neste método, a dCas9 é fusionada a repetições in tandem ao peptídeo GCN4, que recruta múltiplas cópias de ModT, via ligação a um anticorpo (scFv), que se liga a GCN4 (Figura 2H) (Tanenbaum et al., 2014). Em plantas, essa estratégia foi adaptada com utilização do AtvT VP64, tornando-se capaz de promover o aumento da transcrição de genes endógenos e elementos transponíveis em até 4.000 vezes (Tabela 3) (Papikian et al., 2019). Uma variação da técnica é a substituição de VP64 pelo domínio catalítico de metiltransferases para estabelecer a metilação em regiões promotoras e o controle da transcrição. Dessa forma, o sistema dCas9-SunTag pode ser uma potente ferramenta para a modulação epigenômica sítio-específica. Trata-se de um sistema altamente específico, capaz de promover a ativação transcricional de genes, tanto na eucromatina como na heterocromatina, e com efeitos herdáveis ao longo de múltiplas gerações, mesmo na ausência do transgene (Papikian et al., 2019). Além disso, o uso do sistema dCas9-SunTag associado a um domínio catalítico de metiltransferases minimiza significativamente a metilação não específica generalizada, comum quando esses domínios são fusionados diretamente à dCas9 (Pflueger et al., 2018).

Redução da expressão gênica (CRISPRi)

A dCas9 sozinha é capaz de inibir fisicamente a elongação ou o início da transcrição de um gene-alvo, pelo impedimento da ligação ao complexo transcricional, quando direcionada à sua região promotora, nas proximidades do TATA-Box ou TSS,

ou à sua região codificadora (Figura 2A). A estratégia CRISPRi é uma plataforma eficiente para a regulação negativa direcionada da expressão de genes, com efeitos reversíveis e pouca incidência de *off-targets*, podendo ser aplicada em escala multi-gênica (Qi et al., 2013). Trata-se de uma plataforma funcional em eucariotos e uma estratégia alternativa para a modulação da expressão gênica em procariotos, que não possuem maquinaria de RNA interferente (RNAi).

Em plantas, a eficiência do método foi demonstrada em sistemas transientes de transformação, atingindo até 80% de repressão da expressão de um gene repórter (Tabela 4) (Piatek et al., 2015; Vazquez-Vilar et al., 2016). Estes últimos autores demonstraram a interferência direta da posição do complexo dCas9-sgRNA, de forma que a expressão gênica detectada foi menor para sgRNAs localizados mais próximos ao TSS. Observaram também um efeito aditivo modesto e não significativo do uso simultâneo de dois ou três sgRNAs na redução da expressão gênica (Tabela 4) (Vazquez-Vilar et al., 2016).

Fusão de repressores transcricionais à proteína dCas9 (dCas9:RepT)

Seguindo metodologia semelhante à descrita anteriormente, nesta estratégia, a sequência da dCas9 é fusionada in tandem a um ou mais repressores transcricionais (RepTs) (Tabela 2). Em plantas, o mais utilizado até o momento é o SRDX, que, além de ser compatível com um sistema multiplex, teve sua máxima eficiência quando três domínios foram fusionados à dCas9, chegando a reduzir a expressão gênica em até 80% (Lowder et al., 2018, 2015; Sarasua, 2020; Vazquez-Vilar et al., 2016). O RepT KRAB, apesar de muito usado em modelos animais, foi ineficiente e instável na modulação da expressão gênica em plantas, mesmo quando testado em diversos contextos, variando a posição do sgRNA em relação ao TSS e até a nuclease Cas (Tabela 4). Esse fenômeno pode ser explicado pelo fato de KRAB ser um motivo exclusivo de tetrápodes, o que limitaria sua capacidade em recrutar proteínas moduladoras da cromatina em células vegetais (Sarasua, 2020).

Interessantemente, o sistema dCas9:SRDX foi capaz de mascarar os efeitos de ativação gênica de dCas9:EDLL e dCas9:TAD, mantendo a expressão gênica em níveis basais quando esses ModTs foram utilizados conjuntamente (Piatek et al., 2015). Apesar de a estratégia dCas9:RepT ser uma ferramenta útil para a modulação da expressão gênica, em alguns casos, sua eficiência foi semelhante ou inferior à observada quando a redução da transcrição foi induzida somente pela proteína dCas9 (Tabela 4) (Sarasua, 2020; Vazquez-Vilar et al., 2016).

Outro ponto a destacar é o fato de, até o momento, esse sistema ter sido observado em plantas apenas quando RepTs foram utilizados. Isso porque a utilização de ortólogos específicos da proteína Cas parece tornar o sistema mais eficiente.

Tabela 4. Estratégias de uso do sistema CRISPR/dCas9 para a repressão da expressão gênica em plantas.

Espécie vegetal	Gene-alvo	Estratégia CRISPR/dCas9	Nível de repressão	Sistema de transformação	Referência
<i>A. thaliana</i>	<i>AtCSTF64</i>	dCas9:3x (SRDX) + 3 sgRNAs	60	Plantas transgênicas	Lowder et al. (2015)
	<i>miR159A</i>	dCas9: 3x (SRDX) multiplex	80	Plantas transgênicas	Lowder et al. (2015)
	<i>miR159B</i>		70		
	<i>miR159B</i>	dCas12a:3x (SRDX)	90	Plantas transgênicas	Tang et al. (2017)
<i>AtFT</i>	dCas9-MS2-SRDX dCas9-MS2-KYP	0 40	Plantas transgênicas	Lee et al. (2019)	
<i>N. benthamiana</i>	<i>NbPDS</i>	dCas9	20	Agroinfiltração	Piatek et al. (2015)
		dCas9:SRDX	50		
		dCas9	80	Agroinfiltração	VazquezVilar et al. (2016)
		dCas9 + 2 sgRNA	85		
		dCas9 + 3 sgRNAs	85		
		dCas9:SRDX	50		
		dCas9:SRDX + 3 sgRNAs	50		
		dCas9:BRD	40		
		dCas9:BRD + 3 sgRNAs	70		
		dCas9:BRD		Agroinfiltração	Sasuna (2020)
		dCas9:BRD + 2 sgRNAs	30		
		dCas9:SunTag – BRD	30		
		dCas9:SunTag – BRB + 2 sgRNA	30		
		dCas9:SRDX	40		
<i>pNOS:</i>	dCas9:SRDX + 2sgRNAs	40			
<i>luciferase</i>	dCas9:SunTag - SRDX	65			
	dCas9:KRAB	40			
	dCas12a:BRD	30			
	dCas12a:BRD + 2sg RNA	80			
	dCas12a:SunTag – BRD	60			
	dCas12a:SunTag – BRD + 2sgRNAs	80 90			
	dCas12a:SRDX	80			
	dCas12a:SRDX + 2sg RNA	80			
	dCas12a:SunTag – SRDX	80			
	dCas12a:SunTag – SRDX+ 2 sgRNA	90 30			
	dCas12a:KRAB				

Tang et al. (2017), por exemplo, ao utilizarem a nuclease dCas12a (também denominada dCpf1) fusionada a três domínios do RepT SRDX [dCas12a:3x (SRDX)], observaram redução de até 90% na transcrição do *miR159B* de *A. thaliana*. Em sistema semelhante, utilizando dCas9:3x (SRDX), esse valor foi de aproximadamente 70% (Tabela 4) (Lowder et al., 2018). Sarasua (2020) também observou a superioridade da nuclease dCas12a, comparada à dCas9, quando ambas foram fusionadas aos RepTs SRDX e BRD para a redução da expressão de um gene repórter, em um sistema transiente de transformação. Também foi observado um efeito aditivo variável e significativo do uso de dois sgRNAs apenas para a combinação dCas9:SRDX, não havendo diferença quando um ou dois sgRNAs foram usados nas combinações dCas9:BRD e dCas12a:SRDX. Já para dCas12a:BRD, o uso de dois sgRNAs foi menos eficiente (Tabela 4). O fato de a dCas12a atuar com mais eficiência na repressão da transcrição gênica, em comparação à dCas9, pode estar relacionado ao tamanho dessas nucleases, o que facilitaria seu acesso à região genômica-alvo e a modificação da cromatina (Sarasua, 2020). Enquanto dCas12a possui, aproximadamente, 1.250 aminoácidos, dCas9 possui 1.368 aminoácidos.

Modificação da estrutura de sgRNAs para o recrutamento de repressores

Apesar de ser uma estratégia bastante utilizada para a ativação da transcrição gênica, há na literatura até o momento apenas um estudo utilizando a metodologia CRISPR-Act2.0 para a repressão da transcrição gênica em plantas (Lee et al., 2019). Os autores relataram a ineficiência do método com utilização do RepT SRDX, porém foi possível atingir níveis de repressão mediana com outro RepT (KYP) (Tabela 4).

Sistema SuperNova Tag (SunTag) para repressão gênica

Esta estratégia foi bastante explorada em um estudo detalhado realizado por Sarasua (2020), que variou diversos parâmetros do método, como posição do sgRNA em relação ao TSS, número de sgRNAs (1 ou 2), tipo de nuclease Cas (dCas9 ou dCas12a), tipo de RepT (SRDX ou BRD) e tamanho dos espaçadores (5 ou 22 aminoácidos) entre os epítomos GCN4, em um sistema transiente de expressão gênica para o monitoramento da expressão do gene repórter da luciferase guiado pelo promotor da nopalina (*pNOS*). Para fins de otimização do método em plantas, foi observada maior eficiência com a utilização da nuclease dCas12a (Tabela 4), provavelmente devido ao seu menor tamanho em comparação à dCas9, conforme discutido anteriormente, e de espaçadores maiores, de 22 aminoácidos entre epítomos, uma vez que espaçadores menores podem oferecer impedimentos estéricos à ligação de anticorpos. No geral, o sistema SunTag promoveu níveis de repressão gênica similares aos observados com a estratégia de dCas:RepT, mostrando-se pou-

co superior quando foram utilizados dois sgRNAs, com a dCas12a (dCas12a:Suntag – SRDX/BRD + 2 sgRNA). Não foi observada diferença marcante entre os dois RepTs testados (SRDX e BRD) (Tabela 4).

Desenho experimental: importantes aspectos a serem considerados

Como visto anteriormente, a modulação epigenética transcricional mediada pela tecnologia CRISPR contribui significativamente para o entendimento e controle ortogonal das regulações transcricionais e pós-transcricionais, visando avaliar e manipular eventos celulares pela modulação da expressão de moléculas desejadas. Em comparação com outros métodos, como a tecnologia de RNAi, a utilização da tecnologia CRISPR é fácil e oferece um direcionamento mediado por RNA de genes individuais ou múltiplos, o que a torna uma ferramenta poderosa no estudo da regulação transcricional nativa multifatorial de diversos organismos, incluindo plantas (Deaner et al., 2017; Jensen et al., 2017; Zalatan et al., 2015).

Além disso, conjuntos únicos de experimentos de transformação (multiplex) podem ser facilmente implementados, e a possibilidade da utilização de mais de um sgRNA permite uma rápida progressão por meio de ciclos de engenharia genética interativos. Sendo assim, é possível aplicar uma rápida avaliação dos efeitos combinatórios de variações na expressão gênica, a fim de identificar alvos primários e secundários que não podem ser conhecidos a priori com a variação na expressão de um único gene. Mesmo assim, existem muitas considerações a respeito do desenho experimental que devem ser levadas em conta ao se usar o sistema CRISPR para edição de bases ou modulação da transcrição de genes. Dessa forma, serão abordados por meio de tutoriais, nos tópicos a seguir, dois aspectos importantes: 1) desenho do sgRNA; e 2) desenho do cassete de transformação (vetores plasmidiais).

Desenho dos sgRNAs

Embasamento teórico para a escolha do melhor software/plataforma

Dois principais desafios para a aplicação biotecnológica do sistema CRISPR/Cas estão associados com o desenho correto do sgRNA: 1) predição confiável de *off-targets*; e 2) eficiência no reconhecimento de *on-targets* (Xu et al., 2015; Zhang et al., 2015).

Os parâmetros para o desenho correto dos sgRNAs, utilizados na tecnologia CRISPR/dCas9, são basicamente os mesmos a serem considerados para proteínas Cas nativas, mesmo que a eficiência de clivagem não seja um fator importante nessa tecnologia. Dessa forma, três passos principais devem ser seguidos para o desenho de um sgRNA: 1) identificação das sequências PAM na região a ser editada; 2) seleção de uma sequência de 20 nucleotídeos a montante de PAM (excluindo-o), nomeada como *spacer*; e 3) adição do *spacer* à extremidade 5' do *RNA scaffold*, importante na ancoragem do sgRNA nas proteínas Cas9/dCas9. O desenho experimental bem-sucedido do sgRNA, mais especificamente do *spacer*, engloba a predição tanto dos *on-targets* quanto dos *off-targets* (Doench et al., 2016; Tang et al., 2018).

Para sgRNAs desenhados para a Cas9, guaninas são preferidas nas posições 1 e 2 antes da sequência PAM (NGG), enquanto a presença de timinas nas posições +4 e -4, que cercam PAM, desfavorecem a eficiência de clivagem (Wang et al., 2014). Além disso, as regiões a jusante de PAM podem influenciar a eficiência de clivagem e a especificidade do sgRNA, enquanto as sequências a montante geralmente não apresentam efeito significativo a respeito desses parâmetros (Doench et al., 2014). A citosina é preferida no local de clivagem pela Cas9 (posição -3 próximo ao PAM) (Cong et al., 2013; Xu et al., 2015), e um alto conteúdo de GC na região que compreende as bases 4-13, a jusante da sequência PAM, contribui para a eficiência de clivagem e especificidade do sgRNA. O tamanho do sgRNA é também um fator crucial para a especificidade do alvo da Cas9, uma vez que o seu tamanho mínimo para experimentos visando o nocaute genético é de 17 nucleotídeos (Fu et al., 2014). Já para a modulação transcricional e epigenética mediada por dCas9, o melhor tamanho do sgRNA é de 20 nucleotídeos, uma vez que a redução no tamanho dos sgRNAs diminui significativamente a eficácia da regulação transcricional mediada por dCas9 (Kiani et al., 2015; Smith et al., 2016). Existe certo grau de flexibilidade no desenho das posições distais de sgRNAs desenhados para reconhecer promotores de genes em regiões ricas em nucleossomos ou sequências de ativação a montante (UASs, do inglês *upstream-activating sequences*) (Kiani et al., 2015; Smith et al., 2016).

Ferramentas computacionais poderosas são continuamente desenvolvidas e atualizadas para facilitar cada vez mais o desenho experimental *in silico* de sgRNAs para fins biotecnológicos (Chuai et al., 2017; Graham; Root, 2015; Wilson et al., 2018). Porém, como escolher o melhor software/plataforma para o desenho de sgRNAs diante da diversidade de opções? Para facilitar a escolha, é preciso saber o que cada um oferece, destacando seus pontos positivos e limitações. Sendo assim, vários modelos de previsão de eficiência *in silico* foram desenvolvidos, baseados nas principais observações experimentais na área. O *Conjunto de Regras 1* é um modelo preditivo baseado em um sistema de pontuação associado a um método

de aprendizado *in silico*, que classifica os dados de maneira linear e generalizada, a fim de avaliar a eficiência de corte da Cas9, com alta correlação a resultados experimentais (Doench et al., 2014; Liu et al., 2020). Para melhorar a precisão, estudos independentes adaptaram mais conjuntos de dados e construíram um novo modelo, denominado *Conjunto de Regras 2*, no qual análises de nucleotídeos, independentemente da posição e da localização do sgRNA dentro da região a ser editada, podem melhorar a previsão de *off-targets* e *on-targets* (Doench et al., 2016; Liu et al., 2020).

Os softwares e algoritmos desenvolvidos para prever a especificidade dos sgRNAs (determinação de *off-targets*) possuem variados tipos de entrada de dados (*input*) (Tabela 5), mas todos são baseados em dois métodos principais: 1) método baseado em alinhamento, usado principalmente para descobrir *in silico* todos os possíveis *off-targets*; e, 2) método baseado em pontuação (*score*), no qual os sgRNAs são pontuados e classificados de acordo com uma matriz de *off-targets* identificados por alinhamento.

Em teoria, os potenciais *off-targets* podem ser identificados ao se alinhar as sequências de sgRNAs ao genoma de referência, com base na similaridade da sequência. As ferramentas *Bowtie* (Langmead et al., 2009) são capazes de detectar *off-targets*, mas não conseguem identificar pequenos PAMs, uma vez que foram desenvolvidas para o alinhamento de *reads* provenientes de sequenciamento de nova geração (NGS) (Haeussler et al., 2016; Wilson et al., 2018). Além disso, essas ferramentas permitem um número muito limitado de *mismatches*, tornando impraticável a identificação de todos os *off-targets* em potencial (Langmead e Salzberg, 2012). Deve-se considerar também que nem todas as posições de nucleotídeos que contêm *mismatches* têm o mesmo efeito decisivo na clivagem de *off-targets* e que a previsão baseada em alinhamento sempre identifica *off-targets* redundantes, dos quais muitos são falso-positivos. Levando em consideração esses fatores, entre as ferramentas baseadas exclusivamente em alinhamentos, os softwares CasOFFinder e FlashFry são as melhores opções (até o momento) para identificar os possíveis *off-targets* para qualquer nuclease Cas (Liu et al., 2020). Vale a pena destacar a alta velocidade de processamento e *outputs* de fácil compreensão do FlashFry.

Para melhorar a forma de detecção de *off-targets*, o método baseado em pontuação (*score*) foi desenvolvido utilizando-se basicamente duas abordagens: 1) orientada por hipóteses, em que os *off-targets* são pontuados com base na contribuição de fatores específicos do contexto do genoma para a especificidade do sgRNA; e 2) baseada na aprendizagem, em que os sgRNAs são pontuados e previstos a partir de um modelo de treinamento, que considera os diferentes recursos que afetam a especificidade.

Tabela 5. Ferramentas on-line mais usadas para o design de sgRNAs⁽¹⁾.

Ferramenta	Sítio na Web	Espécies	Entrada de Dados (Input)
Benchling	https://www.benchling.com	Diversas	Sequência de DNA; ID do gene
CCTop	https://crispr.cos.uni-heidelberg.de	Diversas	Sequência de DNA
CHOPCHOP	https://chopchop.cbu.uib.no	Diversas	Sequência de DNA; TSS; localização ou ID do gene
CRISPETa	http://crispeta.crg.eu	Espécies Modelo	Sequência de DNA; localização ou ID do gene
CRISPOR	http://crispor.tefor.net	Diversas	Sequência de DNA; localização ou ID do gene
CRISPR-ERA	http://CRISPR-ERA.stanford.edu	Diversas	Sequência de DNA; TSS; ID do gene
CRISPR-GE	http://skl.scau.edu.cn	Plantas	Sequência de DNA; localização ou ID do gene
CRISPR-P 2.0	http://crispr.hzau.edu.cn/CRISPR2	Plantas	Sequência de DNA; localização ou ID do gene
CRISPR-PLANT	https://www.genome.arizona.edu/crispr2	Plantas	Sequência de DNA; localização ou ID do gene
CRISPR RGEN Tools	http://www.rgenome.net	Diversas	Sequência de DNA; localização ou ID do gene; sgRNA
CRISPRscan	https://www.crisprscan.org	Espécies Modelo	Sequência de DNA; localização ou ID do gene
E-CRISP	http://www.e-crisp.org/E-CRISP	Diversas	Sequência de DNA; ID do gene
EuPaGDT	http://grna.ctegd.uga.edu	Patógenos Eucarióticos	Sequência de DNA
FlyCRISPR	https://flycrispr.org/	Insetos	Sequência de DNA
Yeastrixtion	http://yeastrixtion.tnw.tudelft.nl/#!/	<i>Saccharomyces cerevisiae</i>	ID do gene

⁽¹⁾ As informações aqui apresentadas estão sujeitas a atualizações dos sítios na Web.

Na abordagem orientada por hipóteses, a pontuação MIT (Hsu-Zhang) avalia as contribuições de *mismatches* em diferentes posições ao longo de todo o sgRNA com a utilização de uma matriz de peso para identificação de possíveis *off-targets*. A pontuação MIT foi integrada a ferramentas de desenho de sgRNA, tais como CHOPCHOP (Labun et al., 2016) e CRISPOR (Haeussler et al., 2016). De forma complementar, mas independente do sistema MIT, a determinação da frequência de corte (CFD, do inglês *cutting frequency determination*) é outro sistema de pontuação popular para avaliação de *off-targets*, em que a presença de PAMs não canônicas (NAG, NCG e NGA) são inseridas como fatores importantes na pontuação, assim como *indels* e *mismatches* nas sequências-alvo. O *score* CFD está inserido nas ferramentas GUIDE-Seq (Tsai et al., 2015), CRISPRscan (Moreno-Mateos et al., 2015), GuideScan (Perez et al., 2017), CRISPOR (Haeussler et al., 2016) e Benchling (Doench et al., 2016), provando ter um desempenho melhor que o *score* MIT. Posteriormente, a previsão da especificidade do sgRNA com base nas características estruturais do complexo Cas9-sgRNA provou ser superior à previsão baseada somente nas características da sequência. As CRISPROff (Alkan et al., 2018) e uCRISPR (Zhang et al., 2019a) utilizam propriedades energéticas para avaliação de *off-targets*. Ambas, quando comparadas com ferramentas que utilizam outros métodos de pontuação, como MIT e CFD, se mostraram mais precisas na predição de *off-targets*. No entanto, nenhuma delas foi avaliada sistematicamente por experimentos em larga escala, e deve-se ter cuidado ao utilizá-las.

Algoritmos empíricos orientados por hipóteses podem não prever *off-targets* de forma eficiente, uma vez que consideram poucas fontes de dados, enquanto métodos baseados em aprendizado constroem modelos complexos, usando combinações de vários recursos. O algoritmo *CRISPR Target Assessment* (CRISTA) usa o *BWA* como ferramenta de busca de *off-targets* e implementa vários recursos (PAM, composição nucleotídica, conteúdo de GC, estrutura da cromatina, metilação do DNA, estrutura secundária do sgRNA, etc.) para prever a propensão à clivagem (Abadi et al., 2017). O CRISTA apresenta melhor desempenho que o MIT e CFD. Em paralelo, o DeepCRISPR é uma plataforma computacional de última geração que unifica a previsão de sítios para sgRNA, tanto para *on-targets* quanto para *off-targets*, em uma estrutura com aprendizado profundo (Chuai et al., 2018). Essa ferramenta, além de identificar sequências-alvo e características epigenéticas possíveis que podem afetar a eficácia do sgRNA, sendo alimentada por diferentes bases de dados, foi capaz de superar outras ferramentas disponíveis de previsão de *off-targets* (Doench et al., 2016; Hsu et al., 2013; Liu et al., 2020; Singh et al., 2015; Stemmer et al., 2015).

Muitas ferramentas on-line combinam vários dos métodos de pontuação acima mencionados ou até implementam seu próprio sistema de avaliação (Tabela 5), a fim de fornecer opções alternativas para os usuários, como o CHOPCHOP (Labun et

al., 2019, 2016; Montague et al., 2014) e o CRISPRscan (Moreno-Mateos et al., 2015). O E-CRISP utiliza seu próprio sistema de pontuação para determinar a qualidade de cada sgRNA, denominado de *Especificidade, Anotação, Eficácia* (SAE, do inglês *Specificity, Annotation, Efficacy*). O CCTop (Stemmer et al., 2015) atribui empiricamente pontuações para a predição de *off-targets*. Já o CRISPOR (Haeussler et al., 2016) é uma plataforma versátil que classifica os sgRNAs de acordo com diferentes sistemas de pontuação que visam avaliar potenciais *off-targets* no genoma de interesse e prever a atividade *on-target*, além de oferecer uma lista definida com diferentes sequências PAM para serem selecionadas pelo usuário. A plataforma CRISPR RGEN (Sullenger, 2020) disponibiliza diversas ferramentas computacionais, assim como informações para o desenho de sgRNAs, para várias endonucleases ortólogas a Cas9. A plataforma possui nove ferramentas úteis, dentre as quais se destacam o Cas-Designer (Park et al., 2015). Comparada com outras ferramentas, o Cas-Designer permite *mismatches* ao detectar *off-targets* no híbrido DNA/RNA. Além disso, esse método de detecção é mais rápido do que outros, devido ao algoritmo CasOFFinder associado. Sequências ou coordenadas genômicas e formatos de arquivo FASTA são permitidos como entradas. Mais de 350 genomas e 20 tipos de sequências PAM são disponibilizados para os usuários, e os resultados incluem, além dos *on-targets*, pontuação *out-of-frame* (calculada por micro-homologia), número de incompatibilidade (0–2) e sequências *off-targets* com *mismatches* de até 1 par de bases. Sítios para *on/off-targets* são redirecionados para o navegador de genomas *Ensembl1*, para realização de uma avaliação adicional. Ao contrário do Cas-Designer, o BE-Designer é implementado principalmente para edição de bases. Nessa ferramenta, quatro métodos de edição de bases são especificados, e a região de edição também é ajustável. O CRISPR-ERA (Liu et al., 2015) e o CHOPCHOP v3 (Labun et al., 2019) suportam o desenho de sgRNA para o sistema dCas9, sendo o CHOPCHOP uma plataforma que inclui 30 diferentes espécies de plantas. CRISPETA (Pulido-Quetglas et al., 2017) é usado principalmente para a deleção em genoma com a utilização de dois sgRNAs. Além disso, a escolha da ferramenta adequada deve ser feita com cautela, uma vez que diferentes ferramentas são desenvolvidas, preferencialmente, para diferentes genomas ou tipos de células. Por exemplo, o Yeaststriction (Mans et al., 2015) é direcionado para leveduras, o FlyCRISPR (Gratz et al., 2014) é voltado principalmente para espécies de inseto do gênero *Drosophila* (e alguns outros insetos), o EuPaGDT (Peng; Tarleton, 2015) é recomendado para patógenos eucarióticos, enquanto CRISPR-P/CRISPR-P 2.0 (Lei et al., 2014; Liu et al., 2017), CRISPR-PLANT (Minkenberg et al., 2019) e CRISPR-GE (Xie et al., 2017) são implementados para a edição de genomas de plantas. O CRISPR-P 2.0 é uma das

¹ Disponível em: <https://www.ensembl.org/index.html>

plataformas web mais utilizadas para edição de genomas em plantas na atualidade e auxilia tanto no desenho do sgRNA para *on-targets* quanto para a predição de *off-targets*. Ele suporta o desenho de sgRNAs para 49 genomas de plantas (até o momento), cobrindo quase todas as espécies para as quais existe um genoma bem montado, mas suporta também a identificação de sgRNAs em sequências fornecidas pelo usuário. Utiliza, também, um sistema de pontuação modificado para classificar *off-targets* e a eficiência de direcionamento de sgRNAs para a Cas9 de *S. pyogenes*, a mais usada até o momento, e também de outros sistemas CRISPR, incluindo Cpf1 (Zetsche et al., 2015). Essa plataforma avalia o conteúdo do GC, a presença de possíveis sítios para endonuclease de restrição, pontua a micro-homologia da sequência que flanqueia a região a ser editada e avalia a estrutura secundária do sgRNA.

Visto que existem poucos softwares/plataformas on-line específicos para o desenho de guias direcionados para a tecnologia CRISPR/dCas9, o CHOPCHOP v3² (Labun et al., 2019) foi o software escolhido para a elaboração de um tutorial, principalmente por apresentar funcionalidade específica para dCas9, por conter dados referentes a mais de 250 espécies e passar por atualizações constantes. No entanto, caso já se tenham informações *in silico* e/ou *in vivo* a respeito da região promotora-alvo, outros softwares/plataformas podem ser utilizados, tais como CRISPR RGEN³ ou CRISPR-P 2.0⁴, que são versáteis e de fácil uso.

Tutorial 1: Desenho do sgRNA visando modulação transcricional via dCas9

Passo 1: Acessar a URL <https://chopchop.cbu.uib.no/>

Passo 2: Escolha da espécie de interesse (Seção *In*). As espécies eucarióticas e procarióticas já inseridas em seu banco de dados encontram-se divididas, até o momento, nas seguintes categorias: Arthropoda (37), Bacteria (58), Chordata (46), Cnidaria (1), Ctenophora (1), Echinodermata (2), Fungi (30), Mollusca (2), Nematoda (2), Outros (12), Parasitas (11), Plantas (30), Virus (20). No presente tutorial, foi escolhida a espécie de planta *Glycine max* (GCA_000004515.3). Caso o genoma do organismo de interesse não esteja inserido, é possível solicitar aos desenvolvedores a sua inserção, desde que o usuário tenha em mãos um arquivo FASTA com a montagem do genoma-alvo, assim como um arquivo GFF3 válido referente à sua anotação funcional (validado no programa *gff3ToGenePred*, disponível na própria plataforma CHOPCHOP) (Figura 3). O ID atribuído no arquivo GFF3 deve

² Disponível em: <https://chopchop.cbu.uib.no/>

³ Disponível em: <http://www.rgenome.net/about/>

⁴ Disponível em: <http://crispr.hzau.edu.cn/CRISPR2/>

chopchop.cb.uib.no

01 Home Instructions Scoring About Updates Submissions Contact FAQ

CHOPCHOP

03 Target: Glyma_09G153900
RefSeq/ENSEMBL/gene name or genomic coordinates.

02 In: Glycine max (GCA_00000451)
Add new species.

04 Using: CRISPR/Cas9
Change default PAM and guide length in Options.

05 For: activation
Presets can be adjusted in Options.

Paste Target Options Reset Options

06a Find Target Sites! 07

06b

ABA "GENERAL"

Target specific region of gene:
 Coding region All exons (inc. UTRs) Splice sites 5' UTR 3' UTR Promoter 300 0
 Only target exon(s): e.g. 1,2

Restrict targeting:
 Search exons and immediate short flanking regions. Only search within the exon.

Isoform consensus determined by:
 Intersection (only searches regions present in all isoforms) Union (searches all exons in all isoforms)

Pre-filtering:
 Minimum required GC [%] content has to be between min: 10 and max: 90
 Self-complementarity has to be below: -1

Restriction enzymes:
 Company preference: NEB
 Minimum size of restriction enzyme binding site: 4

Fasta input:
 Color scoring should ignore one off-target without mismatches.
 Displayed flanking sequence length in detailed view: 300

ABA "Cas9"

sgRNA length without PAM: 20

PAM-3':
 NGG NAG NGA NRG (R = A or G) NNAGAAW (W = A or T) NNNGMTT (M = A or C) NNGRRT (R = A or G)
 Custom PAM: e.g. NGAG

Method for determining off-targets in the genome:
 Off-targets with up to 3 mismatches in protospacer (Hsu et al., 2013)
 Off-targets may have no more than 0 mismatches in the protospacer seed region (Cong et al., 2013)

Efficiency score:
 Doench et al. 2014 - only for NGG PAM
 Doench et al. 2016 - only for NGG PAM
 Chari et al. 2015 - only NGG and NNAGAAW PAM's in hg19 and mm10
 Xu et al. 2015 - only for NGG PAM, but can be used with other PAMs
 Moreno-Mateos et al. 2015 - only for NGG PAM
 G20

Repair profile prediction (Shen et al. 2018):
 mESC (recommended when you don't know which cell type)
 U2OS
 HEK293
 HCT116
 K562
 Don't calculate (saves time)

5' requirements for sgRNA:
 GN or NG
 GG
 No requirements

Self-complementarity (Thyme et al.):
 Check for self-complementarity
 I intend to replace the leading nucleotides with "GG"
 Check for complementarity versus backbone:
 Standard backbone (AGGCTAGTCCGT)
 Extended backbone (AGGCTAGTCCGT,ATGCTGGAA)
 Custom backbone: e.g. ATGCTGGAA

Figura 3. CHOPCHOP v3 – Parâmetros. Apresentação do passo a passo, descrito no Tutorial 01, para a escolha dos principais parâmetros para o desenho de sgRNAs, visando a ativação via dCas9 de um gene específico de *G. max* (ID: Glyma_09G153900). A numeração apresentada em cada retângulo vermelho representa os passos de 1 a 7 do Tutorial 1.

corresponder ao identificador que se deseja usar para procurar os genes, assim como se deve ter certeza de que os nomes dos cromossomos sejam idênticos, tanto no arquivo GFF3 quanto no FASTA, e as coordenadas dos genes correspondam ao arquivo FASTA do genoma. Após o cumprimento de todas essas exigências, deve-se enviar uma requisição por e-mail (disponível na seção “Add new species”, abaixo do local de seleção da espécie de interesse). Vale a pena ressaltar que a inserção de um novo genoma é feita pelos mantenedores da plataforma. Não é permitida ao usuário a inserção manual do genoma-alvo.

Passo 3: Escolha do gene-alvo (Seção *Target*). É nesse campo onde se insere o ID do gene-alvo. O ID do gene escolhido para este tutorial é Glyma_09G153900. É importante que o ID corresponda exatamente com o do genoma depositado no banco de dados da plataforma CHOPCHOP. Caso não exista a correspondência, a própria plataforma enviará a mensagem de erro “Error status: 501”, na qual se orienta a forma correta de se escrever o ID para aquele gene da espécie selecionada. Esse campo age de forma sincronizada com o ícone cinza “Paste Target”. Com essa opção ativa, é possível encontrar o gene de acordo com o ID (formato RefSeq ou ENSEMBL), nome do gene ou coordenadas do gene no genoma (por exemplo, chrX:15,560,138-15,602,945). Ao clicar no ícone cinza “Paste Target”, o usuário muda a função do ícone para a opção “Gene Target”, onde é possível colocar a sequência de nucleotídeos que se deseja analisar. Esta última opção não é muito utilizada nessa plataforma para o desenho de sgRNAs visando a utilização da tecnologia dCas9, pois, na maior parte dos casos, com essa abordagem, se perde o contexto genômico da região promotora do gene-alvo, o que é essencial para a modulação transcricional via dCas9 (Figura 3).

Passo 4: Escolha da nuclease (Seção *Using*). Neste campo, a plataforma disponibiliza até o momento cinco diferentes tipos de nucleases para edição de genomas: 1) CRISPR/Cas9; 2) CRISPR/Cas9 nicase; 3) CRISPR/Cpf1 (CasX); 4) CRISPR/Cas13 (c2c2); e 5) TALEN. Para este tutorial, a nuclease selecionada é “CRISPR/Cas9” (Figura 3).

Passo 5: Escolha da estratégia (Seção *For*). Neste campo, é selecionada a abordagem ou contexto no qual serão utilizados os sgRNAs que serão selecionados (para este tutorial, foi selecionada a opção “Activation” (Figura 3):

- a) *Knockout* (para mutações em que se pretende mudar a fase de leitura do gene de interesse; disponível para quase todas as nucleases disponibilizadas na plataforma, com exceção da CRISPR/Cas13): para aplicações com a Cas9, pode-se prever a taxa de mudança na fase de leitura propiciada por cada sgRNA. Para a maioria das aplicações, algumas recomendações são sugeridas: 1) usar a Cas9 tradicional (20 nucleotídeos com sequência PAM 5'-NGG-3'); 2) ter certeza de que o sgRNA possui o menor número possível

de *off-targets*, de preferência nenhum; 3) garantir que o sgRNA selecionado pode reconhecer todas as isoformas do gene-alvo (disponível apenas para genomas selecionados nos quais há informações no nível de isoformas); e 4) selecionar sgRNAs a jusante de quaisquer ATG *in frame* dentro da estrutura (caixas verdes na região codificadora azul, apresentados na tabela de resultados finais) para evitar a expressão de proteínas truncadas.

- b) *Knock-down* (para edição mRNA com CRISPR/Cas13; disponível somente para humanos e camundongos): neste modo é possível procurar *off-targets* em transcriptomas. É possível selecionar o sistema Cas13 em "Options" (após selecionar "CRISPR/Cas13" na seção *For*), especificando o *protospacer flanking site* (PFS) e o tamanho do sgRNA.
- c) *Knock-in* (para inserção de sequências de DNA em uma região de interesse; indisponíveis para CRISPR/Cpf1, CRISPR/Cas13 e TALEN): é altamente recomendável a familiarização com os diferentes tipos de *knock-ins* disponíveis para a escolha do método mais apropriado (Nami et al., 2018). O *knock-in* pela tecnologia CRISPR permite a engenharia genética específica do local por meio da quebra de fita dupla do DNA (DSB) induzida por nucleases e, posteriormente, o reparo do DSB por vias de reparo celular: tanto por junção final não homóloga (NHEJ) quanto por reparo direcionado por homologia (HDR). Um aspecto importante a ser considerado ao decidir o tipo de técnica usar é se a modificação precisará ocorrer em células que se dividem ou que não se dividem. A HDR é restrita às fases S e G2 do ciclo celular e, portanto, não é adequada para células que não se encontrem em divisão. As sequências referentes aos braços de homologia são fornecidas na página de resultados detalhados do sgRNA. É possível ajustar a posição dos braços da micro-homologia em relação à extremidade 5' do sgRNA (padrão: nucleotídeo -3 a partir da sequência PAM) e especificar o tamanho do braço. Pode ser aconselhável verificar a complementaridade entre a sequência inserida e seus braços de micro-homologia. Nas configurações *default* para *knock-in*, podem ser obtidos braços de homologia de até 800 pares de bases. Caso se deseje um braço de homologia com tamanho maior (até 2.000 pares de bases), é necessário alterar o tamanho da sequência de flanqueamento mostrada na página inicial ("Options" > "General" > "Displayed flanking sequence length in detailed view").
- d) *Activation/Repression* (para sgRNAs que têm como alvo regiões promotoras; disponível para as nucleases CRISPR/Cas9 e CRISPR/Cpf1): nesta opção o default para o desenho dos sgRNAs é: 1) modo de ativação: 300 pares de bases a montante do TSS; 2) modo de repressão: 200 pares de bases a

jusante e a montante do TSS. É recomendável usar mais de um sgRNA para essas aplicações.

- e) *Nanopore Enrichment*: utilizado para o enriquecimento de fragmentos de DNA específicos para sequenciamento por meio da plataforma da Oxford Nanopore com o uso de CRISPR. Está disponível somente para as nucleases CRISPR/Cas9 e CRISPR/Cpf1.

Passo 6: Escolha das opções avançadas (seção *Options*). Esta seção encontra-se no ícone cinza no centro da página inicial e é extremamente relevante para refinar os resultados de acordo com a característica tanto do genoma-alvo quanto da sequência-alvo (Figura 3). Ao clicar nesse ícone, com a nuclease CRISPR/Cas9 selecionada na Seção *Using*, novas opções surgem, divididas em três abas: 1) “General”, na qual é possível ajustar parâmetros gerais para a tecnologia CRISPR; 2) “Cas9”, opção de ajuste de parâmetros exclusivos para a nuclease Cas9; e 3) “Primers”, para ajustar parâmetros relativos ao desenho de primers que serão utilizados em experimentos de PCR para detecção de mutações (utilizado basicamente em experimentos de *knockout*). Para experimentos com dCas9, somente serão ajustados parâmetros presentes nas abas “General” e “Cas9”:

a) Aba “General”:

- *Target a specific region of the gene*: pode-se especificar a região onde os sgRNAs serão desenhados: 1) somente a região codificadora (*default*); 2) toda a sequência exônica (incluindo 5' e 3'-UTR); 3) locais de *splicing*; 4) somente 5'-UTR; 5) somente 3'-UTR; 6) o promotor (é possível especificar quantos pares de bases a montante e a jusante do TSS serão analisados); ou 7) um éxon específico ou um subconjunto específico de éxons. Se desejar direcionar a análise para um íntron, especifique as coordenadas genômicas do íntron (tamanho máximo, 40.000 pares de bases).
- *Restrict targeting*: ao pesquisar *on-targets* em uma região de interesse, no modo *default*, o CHOPCHOP permite que o sgRNA se vincule apenas fora dessa região, para que o corte (na maioria dos casos) ainda ocorra na região-alvo. Você pode desativar essa opção em experimentos com dCas9.
- *Isoform Consensus*: ao analisar genes com múltiplas isoformas, os modos “Intersection” ou “Union” podem ser ativados. No modo “Intersection” o CHOPCHOP procura somente por sgRNAs presentes em todas as isoformas. No modo “Union”, o CHOPCHOP procura por sgRNAs em todos os éxons de todas as isoformas. Portanto, pode-se usar esse modo para

direcionar uma (ou mais) isoformas específicas. Você pode desativar essa opção em experimentos com dCas9.

- *Pre-filtering*: é possível pré-filtrar os sgRNAs com base no conteúdo do GC e na pontuação de autocomplementaridade (definido como -1 para desativar, o que não é recomendado). Os sgRNAs que não atendem a esses requisitos não serão relatados na tabela de resultados finais.
- *Restriction enzymes*: a avaliação da mutagênese pode ser feita por meio da análise da digestão com enzimas de restrição. O CHOPCHOP exibe sítios de restrição no local de destino e permite restringir a pesquisa a sítios de restrição de enzimas comercializadas por uma empresa específica (NEB™, Fermentas™, Promega™, Roche™, Sigma™ e Stratagene™), assim como o número mínimo de bases que o sítio de restrição deve ter. Este parâmetro é irrelevante em experimentos de modulação transcricional mediada por dCas9.

b) Aba “Cas9”:

- *sgRNA length without PAM*: uma vez que alguns estudos sugerem que sgRNAs truncados podem apresentar maior eficiência de clivagem, é possível utilizar valores menores de 20 nucleotídeos (*default*), o que não é recomendado para experimentos com dCas9.
- *PAM-3' sequence*: o *default* para o PAM referente a Cas9 é NGG. Alternativamente é possível selecionar um PAM de um sistema ortólogo, ou ainda inserir um PAM personalizado. Para a dCas9 de *S. pyogenes*, recomenda-se usar o PAM *default*.
- *Method for determining off-targets in the genome*: existem duas opções: 1) *mismatches* de uma base até a décima primeira base ao 5' da sequência PAM abolem completamente a clivagem por Cas9 (Cong et al. 2013). No entanto, mutações mais a montante do PAM mantêm a atividade de clivagem. Este método procura *mismatches* apenas nos primeiros nove nucleotídeos, pois se prevê que um *mismatch* posterior ao PAM não cause clivagem; e 2) *mismatches* podem ser tolerados em qualquer posição, exceto no PAM. Este método procura *mismatches* apenas nos 20 pares de bases a montante do PAM (Hsu et al., 2013). Este é o método *default* e o mais utilizado para experimentos com dCas9.
- *Efficiency score*: existem vários métodos para se avaliar e calcular a eficiência de clivagem pela nuclease Cas9. A forma mais simples de pontuação de eficiência é a “G20” (Doench et al., 2016), em que se prioriza uma guanina na posição 20, logo após o PAM, sendo este valor normalizado na plataforma

CHOPCHOP para se encontrar no intervalo entre 0 e 1. Visto que a eficiência de clivagem é um fator irrelevante em experimentos com dCas9, recomenda-se usar o default para este parâmetro.

- *Repair profile prediction*: de acordo com modelos específicos (Shen et al., 2018), é possível prever o perfil de reparo do DNA para um determinado sgRNA. Para esta opção, deve-se selecionar o tipo de célula, sendo disponíveis somente células animais (U2OS, HEK293, HCT116, K562), ou o modo “mESC” se o tipo celular a ser editado não for conhecido. Para experimentos com dCas9, a opção “Don’t calculate (saves time)” deve ser selecionada, pois, além de consumir muito tempo computacional, é irrelevante para esse tipo de análise.
- *Requirements for the 5’ end of the sgRNA*: dependendo da polimerase usada para a síntese do sgRNA, é possível selecionar os seus dois primeiros nucleotídeos visando a melhor eficiência de transcrição. Por exemplo, 5’-GN- (para o promotor U6) ou 5’-GG- (para T7 polimerase). Com relação ao sistema dCas9, caso o propósito seja a edição de bases, em que geralmente se utiliza RNPs (abordagem na qual o sgRNA será sintetizado in vitro, com auxílio da T7 polimerase), a opção “GG” deve ser selecionada. Já em experimentos de modulação transcricional, em que o sgRNA será expresso in vivo, regulado pelo promotor U6, a opção “GN or GG” deve ser selecionada.
- *Self-complementarity*: estudos sugerem que a autocomplementaridade dentro do sgRNA ou entre o *spacer* e o *RNA scaffold (backbone)* (em que a sequência descrita na literatura é 5’-GTTTTAGAGCTAGAAATAGCAAGTTAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC-3’) pode reduzir a eficiência de clivagem do sgRNA (Thyme et al., 2016). Essa opção procura complementaridade dentro do sgRNA, ou entre o *spacer* e um *RNA scaffold* padrão (AGGCTAGTCCGT), um *RNA scaffold* estendido (AGGCTAGTCCGT, ATGCTGGAA) ou um *RNA scaffold* personalizado. Alguns usuários optam por substituir os primeiros dois nucleotídeos de seu sgRNA por “GG” para transcrição via T7 polimerase, e nesse caso selecione a opção “I intend to replace the leading nucleotides with “GG”. Para experimentos com dCas9, é necessário avaliar tanto a opção “Check for self-complementarity” quanto a opção “Check for complementarity versus backbone”. Com relação a esta última, caso a abordagem experimental seja a clássica, na qual o RNA scaffold é o tradicional, deve-se selecionar “Standard backbone (AGGCTAGTCCGT)”, mas, se a abordagem for com a utilização de aptâmeros de RNA (por exemplo, o sistema SAM), a opção “Custom backbone” deve ser selecionada, e a sequência do RNA scaffold modificado deve ser fornecida. Por fim, a opção “I

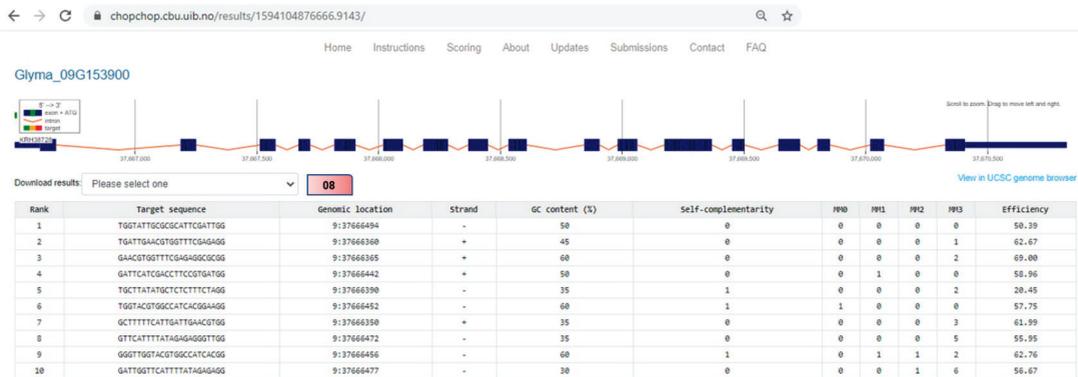


Figura 4. CHOPCHOP v3 – Resultados. Interface da plataforma CHOPCHOP, na qual são apresentados os sgRNA sugeridos para a ativação e um gene específico de *G. max* (ID: Glyma_09G153900), via dCas9. O ranking (coluna 1) é baseado no número de *off-targets* apresentados nas colunas MM0, MM1 e MM2 (com um, dois ou três *mismatches*, respectivamente). Mesmo sendo irrelevante para a tecnologia dCas9, a plataforma fornece valores de eficiência de clivagem (em porcentagem), conforme apresentado na última coluna. Também são apresentadas outras informações, como sequência do *spacer* (coluna 2), posição no genoma (coluna 3), fita de DNA (senso ou positiva; antissenso ou negativa; coluna 4), percentual de GC no sgRNA (coluna 5) e autocomplementaridade (coluna 6). O download dos resultados (quadrado vermelho; passo 8 no Tutorial 1) pode ser feito de acordo com o formato escolhido.

intend to replace the leading nucleotides with “GG” só pode ser marcada em experimentos com dCas9 visando edição de bases.

Passo 7: Execução da análise. Após a seleção de todos os parâmetros, clicar em “Find Target Sites!” para submeter a análise. Caso tudo esteja correto, a mensagem “Your job has been successfully submitted” será exibida. O tempo de análise pode variar de acordo com os parâmetros selecionados. CHOPCHOP exibe os resultados da consulta em uma visualização dinâmica e tabela interativa (Figura 4). A visualização dinâmica exibe todas as opções do *on-target* para a região, especificada com código de cores de acordo com um sistema de pontuação específico: verde (melhor), âmbar (ok) e vermelho (ruim). Em todos os casos, o gene é exibido na direção 5' - 3'. Todas as isoformas do gene são exibidas com seus nomes e, a jusante, os sítios de ATG *in frame* (caixa verde). Basta clicar em um alvo na visualização ou em uma opção na tabela para ser direcionado a uma página de resultados para cada sgRNA individual, contendo informações sobre *on-targets*, *off-targets*, previsões de reparo (quando aplicável), braços de micro-homologia (modo *knock-in*), assim como os sítios de restrição e *primers* desenhados para cada região. Além disso, o CHOPCHOP lista quantos *off-targets* cada *on-target* possui com 0 (coluna “MM0”), 1 (coluna “MM1”), 2 (coluna “MM2”) ou 3 (coluna “MM3”) *mismatches*. Clicar no

resultado revelará mais informações sobre os *off-targets*. Cada *on-target* é ranqueado de acordo com: 1) *score* de eficiência de clivagem (modo “Cas9”); 2) número de *off-targets* e se eles apresentam *mismatches*; 3) existência de regiões de autocomplementaridade com mais de três nucleotídeos (o número indica quantas regiões de autocomplementaridade são previstas); 4) conteúdo de GC (modo “CRISPR/Cas9”): os sgRNAs são mais eficazes com um conteúdo de GC entre 40% e 70% ; e 5) localização do sgRNA dentro de um gene (quanto mais próximo da extremidade 5', melhor será a classificação). Por fim, cada página individual de resultados fornece 1) uma visualização do sítio-alvo de cada sgRNA (com o local de corte previsto em azul no modo “CRISPR/Cas9” ou “CRISPR/Cpf1”); 2) primers desenhados (em roxo); 3) sítios de restrição (em verde – sítios únicos no amplicon, e em vermelho – sítios não exclusivos); e 4) detalhes sobre os *off-targets* (localização genômica, número de *mismatches* e sequência).

Passo 8: Download dos resultados. As análises finais são disponibilizadas para download em quatro diferentes formatos: 1) BED (para visualização em *genome browser*); 2) GenBank (arquivo com anotação e íntron); 3) TSV (tabela de resultados; mais utilizado); e 4) FASTA (Figura 4).

Outras considerações: Não é errado imaginar que a utilização de mais de um sgRNA para diferentes regiões de um mesmo promotor possa, em associação com o complexo dCas9 com um modulador transcricional (ModT), amplificar os níveis de ativação ou repressão desejados (Deaner et al., 2017; Farzadfard et al., 2013; Gilbert et al., 2013). Mesmo que essa afirmação seja considerada verdadeira, ela não é válida para todos os casos, conforme discutido anteriormente. Além disso, é sabido que a presença/ausência de nucleossomos nas mediações do gene-alvo também pode interferir efetivamente com a ação dos reguladores transcricionais (Griesenbeck et al., 2003; Mao et al., 2011). Especificamente para o sistema CRISPR-dCas9, que depende inerentemente da interação com o DNA, é de se esperar que a presença de nucleossomos seja determinante na eficiência da tecnologia em eucariotos, visto que o DNA presente nessa estrutura é amplamente enovelado em histonas, o que o torna mais compactado e menos acessível (Horlbeck et al., 2016; Lee et al., 2007; Radzishewska et al., 2016; Rando; Chang, 2009; Rando; Winston, 2012; Smith et al., 2017, 2016). Sendo assim, evidências bioquímicas e *in vivo* sugerem que as estratégias de desenho de sgRNAs devem evitar o seu direcionamento para próximo ao core do nucleossomo. Associado a essa ideia, o desenvolvimento de futuras ferramentas para o desenho *in silico* de sgRNA específicos e altamente eficientes deve avaliar e incluir em seus algoritmos bancos de dados disponíveis com mapas de acessibilidade ao DNA e posicionamento de nucleossomos para que as análises *in silico* sejam mais fidedignas à realidade *in vivo* (Jiang; Pugh, 2009; Schep et al., 2015). Dessa forma, como nas redes reguladoras de transcrição nativa e em

outras redes sintéticas, o número de reguladores associados ao *regulon* de destino oferece uma válvula modular para ajustar o impacto da reprogramação mediada por CRISPR-dCas9. No entanto, o uso de vários sgRNAs deve ser cuidadosamente analisado, com foco especial na posição dos elementos reguladores e nucleossomos existentes, a fim de ajustar o potencial regulatório por simples aumento no número de sgRNA direcionados a essas regiões.

Desenho do cassete de transformação

A utilização da tecnologia CRISPR visando a modulação transcricional ou epigenética apresenta grande potencial de aplicação biotecnológica em plantas, principalmente a manipulação da expressão de genes endógenos que podem estar associados com algum fenótipo de interesse. Especificamente com relação a dCas9, poucas abordagens são capazes de serem executadas de forma não transgênica (*DNA free*, via ribonucleoproteína). Dessa forma, a modulação transcricional aqui apresentada necessita da expressão constante da dCas9, sgRNA e ModTs. Um passo crítico na aplicação dessa tecnologia é o desenho do cassete para transformação da planta de interesse, o que inclui a seleção de elementos genéticos capazes de produzir de forma otimizada cada um dos componentes do sistema de modulação.

Tutorial 2: Desenho do cassete de transformação para modulação transcricional via dCas9

Para fins didáticos, é proposto neste tutorial um passo a passo que direciona o desenho de um cassete de transformação, dividido em quatro módulos: 1) expressão da dCas9; 2) expressão do sgRNA; 3) expressão da proteína de fusão Adaptador:ModT; e 4) expressão do marcador de seleção. São apresentadas considerações baseadas tanto em experiências laboratoriais quanto na literatura recente a respeito do assunto (Figura 5).

Módulo 1: Expressão da dCas9. A expressão da dCas9 é um fator muito importante no desenho do vetor para modulação transcricional. Este módulo apresenta três elementos: 1) promotor; 2) dCas9; e 3) terminador (Figura 5).

- a) **Promotor:** na maioria dos casos, o promotor escolhido para a expressão da dCas9 é reconhecido por uma RNA polimerase II (pol II), além de ser forte (nível de expressão elevado) e constitutivo (expresso em todos os tecidos e a todo momento). Em plantas, uma boa escolha seria o promotor *CaMV35S*. Mesmo sendo o mais usado e, em princípio, a escolha mais óbvia, alguns aspectos devem ser considerados. O primeiro deles é a origem desse pro-

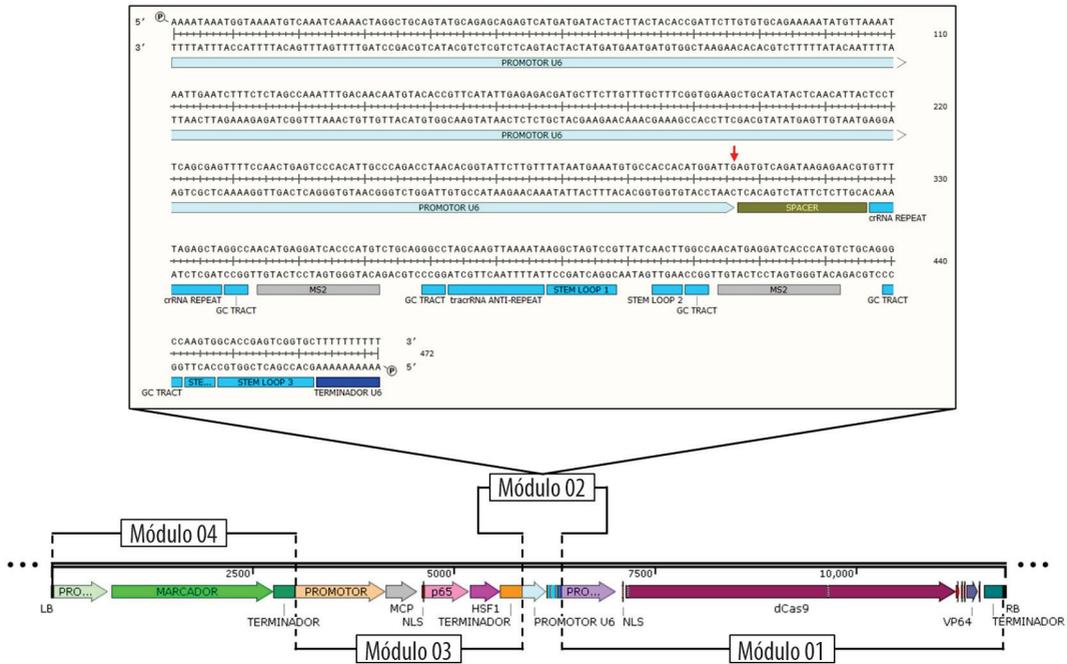


Figura 5. Desenho do cassete de transformação – Tecnologia dCas9 (sistema SAM). Conforme detalhado no Tutorial 2, é apresentado o cassete de transformação com os elementos necessários para a ativação de um gene específico, via dCas9 (sistema SAM), clonado em um vetor binário, baseado no T-DNA de *A. tumefaciens*. O cassete é dividido didaticamente em quatro módulos: *Módulo 1*: Expressão da dCas9; *Módulo 2*: Expressão do sgRNA, com destaque para cada elemento deste módulo (parte superior da figura). O *RNA scaffold* apresenta inserida em sua estrutura a sequência do aptâmero de RNA MS2 tanto no *tetraloop* quanto no *stem loop 2*. A seta vermelha enfatiza a importância da guanina como último nucleotídeo presente no 3' do promotor *GmU6-10* de soja); *Módulo 3*: Expressão da proteína de fusão Adaptador:ModT. Neste módulo, os ModTs apresentados são p55 e HSF1, característicos do sistema SAM; e *Módulo 4*: Expressão do marcador de seleção. Siglas: LB – left border; MCP – MS2 Coat Protein; NLS – sítio de localização nuclear; RB – right border.

motor. Por ser de origem viral, ele pode ser passível de silenciamento em plantas. Visto que a modulação da transcrição do gene-alvo necessita da atividade da dCas9 constante, caso ocorra seu silenciamento, o fenótipo gerado pela modulação do gene-alvo será perdido. O segundo aspecto a ser considerado é a característica constitutiva desse e de outros promotores também usados (por exemplo *AtACT2-5* de *A. thaliana*). É sabido que em plantas, a expressão constitutiva da Cas9 pode propiciar efeitos pleiotrópicos indesejáveis devido a sua ação em possíveis *off-targets*. E o terceiro aspecto a ser considerado é a função do gene cuja expressão será modulada. Muitas vezes, genes que são importantes em resposta a diferentes tipos de estresses apresentam efeitos indesejáveis, se expressos de forma constitutiva. Por isso, sua expressão é naturalmente regulada e restrita ao

momento em que é necessário. Caso o promotor que regulará a expressão da dCas9 seja constitutivo, genes com esta característica também seriam expressos em altos níveis de forma constitutiva. Dessa forma, a escolha do promotor para a expressão da dCas9 deve estar alinhada com a função do gene cuja expressão se pretende modular. Por exemplo, se o fenótipo desejado for a resistência da planta a um patógeno, pode-se pensar em usar um promotor para regular a expressão da dCas9 que seja induzido somente durante a infecção pelo patógeno-alvo.

- b) **dCas9:** deve-se avaliar com cuidado a sequência da dCas9 a ser expressa a fim de modular a transcrição de genes. O primeiro aspecto a ser observado é se as mutações necessárias para a perda da atividade catalítica estão inseridas na sequência (D10A no domínio catalítico RuvC e H840A no domínio catalítico HNH). O segundo aspecto é a presença de sinal de localização nuclear (NLS, do inglês *nuclear localization site*), visto que a atividade da dCas9 será desempenhada exclusivamente no núcleo eucariótico. Em plantas, os principais NLSs escolhidos são: 1) NLS do vírus de macaco SV40 (PKKKRKV); e 2) NLS bi-partido da nucleoplasmina (KRPAATKKAGQAKKKK). O número de repetições e a região onde serão inseridos é variável. Geralmente o número de repetições de cada NLS varia de 1 a 4, podendo estar presente tanto na porção amino quanto na carboxiterminal da dCas9. Dificilmente se encontram construções com apenas um NLS. Além disso, existem construções desenhadas com sucesso que apresentam NLSs também após ModT. Deve-se ter muito cuidado para que a adição desses NLSs não interfira na função da dCas9 no ModT. O terceiro aspecto a ser considerado é a fusão com o ModT. Geralmente essa fusão é mediada por um *linker* ou ponte, correspondente a uma sequência de 5 a 20 resíduos de aminoácidos que liga a região carboxiterminal da dCas9 à aminoterminal do ModT. Esses aminoácidos devem conferir flexibilidade e permitir que ambas as proteínas apresentem sua atividade correta. Inclusive, alguns NLSs inseridos praticamente in tandem, devidamente espaçados por pelo menos cinco aminoácidos neutros, podem ser utilizados como *linkers* entre a dCas9 e o ModT. Vale a pena ressaltar que o tamanho e composição de aminoácidos do *linker* entre a dCas9 e os domínios editores de bases é crucial para a especificidade da edição. *Linkers* menores reduzirão a janela de bases que poderão ser editadas. Mas cuidado: *linkers* muito pequenos podem interferir na função de ambas as proteínas.
- c) **Terminador:** a escolha do terminador é mais livre. Ao se utilizar o promotor *CaMV35S*, sugere-se que seja usado o terminador *35S*, mas qualquer terminador que seja reconhecido por uma pol II pode ser utilizado, como, por exemplo, o terminador *NOS* do gene nopalina sintase de *Agrobacterium tumefaciens*.

Dados experimentais mostram que a presença de dois terminadores diferentes in tandem (por exemplo, terminador 35S e NOS) podem aumentar os níveis de transcrição por melhorar a estabilidade da expressão do transgene, causar a terminação da transcrição mais eficiente e reduzir o silenciamento pós-transcricional do gene-alvo (Basso et al., 2020).

Módulo 2: Expressão do sgRNA. O principal objetivo aqui é delinear de forma eficiente a expressão do sgRNA. Este módulo apresenta quatro elementos: 1) promotor *U6*; 2) *spacer*; 3) *RNA scaffold*; e 4) terminador *U6* (Figura 5).

- a) **Promotor U6:** foi demonstrado que os níveis de expressão de sgRNAs estão diretamente correlacionados com a eficiência da engenharia do genoma mediada por CRISPR/Cas9 em células eucarióticas. Em geral, os promotores da RNA polimerase III (pol III) são mais usados para direcionar a expressão de sgRNAs, visto que os promotores de pol II adicionam nucleotídeos extras às extremidades 5' e 3' dos sgRNAs e, portanto, podem interferir em sua função (Hsu et al., 2013; Yoshioka et al., 2015). Originalmente, os promotores pol III SNR52 e RPR1 foram adotados para expressão constitutiva de sgRNAs em leveduras (DiCarlo et al., 2013; Farzadfard et al., 2013; Gilbert et al., 2013). Em seguida, para permitir maior flexibilidade ao desenho e à força de expressão dos sgRNAs, dois estudos sobre entrega constitutiva de sgRNAs testaram a fusão do Vírus Delta da Hepatite (HDV) e ribozimas autocliváveis do tipo martelo de cabeça, permitindo a expressão de sgRNAs com promotores da pol II (Gao e Zhao, 2014; Ryan et al., 2014). Além dos promotores pol III nativos e pol II induzíveis, diferentes classes de promotores foram utilizadas para regular a expressão de sgRNAs em sistemas eucarióticos (Farzadfard et al., 2013; Ferreira et al., 2018; Jensen et al., 2017; Nishimasu et al., 2014; Smith et al., 2016; Zhang et al., 2017). Além desses, o promotor *U6* se destaca quando o assunto é expressão de sgRNAs. É um promotor de pol III de classe III importante na expressão de pequenos RNAs nucleares não codificadores, envolvidos no *splicing* de íntrons durante a produção de moléculas maduras de mRNA em células eucarióticas. Diferentes promotores *U6* já vêm sendo frequentemente usados em plantas e animais para expressar pequenos RNAs em níveis mais elevados (Li et al., 2007; Miyagishi; Taira, 2002). Atualmente, esse tipo de promotor tem sido a escolha preferida para regular a expressão de sgRNA em vetores CRISPR/Cas9 (e variantes) (Friedland et al., 2013; Li et al., 2013). Além disso, o promotor *U6* possui um TSS altamente conservado, começando com um nucleotídeo de guanina, o que ajuda a melhorar a homogeneidade da molécula de sgRNA transcrita e reduzir os efeitos de *off-targets* (Li et al., 2007). Os sistemas vetoriais CRISPR/Cas9 com o promotor *U6* para direcionar a expressão do sgRNA foram utilizados com sucesso em várias espécies de

plantas. Os promotores *OsU6a*, *OsU6b* e *OsU6c* do arroz são os mais utilizados para monocotiledôneas, e *AtU6-1* e *AtU6-29* de *Arabidopsis*, os preferidos para dicotiledôneas (Jiang et al., 2013; Li et al., 2014; Mikami et al., 2015). No entanto, existem limitações à sua natureza “universal”, pois o promotor *U6* de *Arabidopsis* foi considerado ineficiente em trigo e arroz (Shan et al., 2013). Dessa forma, fica claro que o uso de promotores *U6* específicos da espécie pode resultar em aumento da expressão de sgRNA e, assim, em maior eficiência de edição (Ng; Dean, 2017; Sun et al., 2015). Em soja, por exemplo, os níveis de sgRNA expressos pelo promotor endógeno *GmU6* foram duas vezes maiores que os obtidos com o promotor *AtU6-26* de *Arabidopsis*, resultando em um aumento considerável na eficiência da edição de genes (14,7–20,2% para *GmU6* vs. 3,2–9,7% para *AtU6-26*) (Sun et al., 2015). Tendo isso em vista, além da soja, diversas espécies vegetais de interesse agrônômico (tais como algodão, arroz, cevada, milho, sorgo e trigo, entre outras) já tiveram pelo menos um promotor *U6* endógeno caracterizado e testado em experimentos visando a expressão de sgRNAs (Gasparis et al., 2018; Jiang et al., 2013; Long et al., 2018; Qi et al., 2016; Sun et al., 2015). Também é importante ter em mente que os genomas vegetais contêm vários genes *U6* com diferentes níveis de expressão, mas nem todos os promotores *U6* são igualmente eficientes na condução da expressão gênica (Domitrovich; Kunkel, 2003; Wang et al., 2008). Ao desenvolver sistemas CRISPR/Cas9 para novas espécies, é importante investir tempo caracterizando promotores *U6* exógenos e endógenos a fim de escolher o ideal, especialmente se o sistema de transformação for difícil ou demorado. Em termos práticos, o principal cuidado que se deve tomar a respeito do promotor *U6* durante o desenho do vetor, em experimentos com Cas9 e dCas9, é a presença de uma guanina (G) na extremidade 3’ do promotor (região mais próxima do primeiro nucleotídeo do sgRNA). Caso a sequência disponível do promotor *U6* em questão não apresente esse nucleotídeo, o primeiro nucleotídeo do sgRNA deve ser uma guanina. Em alguns casos, alguns pesquisadores adicionam uma guanina extra à sequência do sgRNA durante o desenho, a fim de otimizar sua expressão.

- b) **Spacer:** esta é a porção do sgRNA que efetivamente irá reconhecer uma sequência no genoma-alvo. Os principais aspectos do seu desenho foram minuciosamente descritos anteriormente. No cassete de expressão do sgRNA, o *spacer* vem imediatamente após o promotor *U6* e antes do *RNA scaffold*.
- c) **RNA scaffold:** é a porção constante presente no sgRNA, importante na ancoragem e interação com a Cas9/dCas9. Como visto antes, estruturalmente, o *RNA scaffold* da nuclease Cas9 (e também da dCas9) é formado por um *tetraloop* e três *stem loops*. Para a tecnologia dCas9, o *RNA scaffold* pode se

apresentar de duas formas: 1) forma tradicional, com a sequência padrão, sem modificação alguma, usado principalmente em abordagens em que não se utiliza aptâmeros de RNA para amplificar a modulação transcricional. Nessa abordagem, a modulação transcricional ou epigenética é mediada somente pelo ModT fusionado com a dCas9; 2) forma modificada (metodologia SAM), em que são inseridas sequências de aptâmeros de RNA (por exemplo, MS2, PP7 e COM) tanto na região do *tetraloop* quanto no *stem loop 2*. A inserção desses aptâmeros de RNA é o que possibilita o reconhecimento por proteína adaptadora (por exemplo, MCP, PCP e COM), sendo cada uma específica para cada aptâmero (Haimovich et al., 2016; Johansson et al., 1997; Lim e Peabody, 2002; Zhang et al., 2019c). Essas proteínas adaptadoras, por sua vez, estão fusionadas com domínios modulares da transcrição. Assim, a escolha de qual forma deve ser usada necessita estar em perfeita consonância com a estratégia experimental abordada. Independentemente da forma a ser utilizada, no cassete de expressão do sgRNA, o *RNA scaffold* vem imediatamente após o *spacer* e antes do terminador *U6*.

- d) **Terminador U6:** a maioria das pol III que reconhecem promotores como o *U6* reconhecem terminadores curtos e simples (repetição de 10 timinas adicionados em sequência ao 3' do último nucleotídeo da sequência que será expressa) (Gao et al., 2018). Não se deve substituí-lo por qualquer outro terminador utilizado no laboratório.

Módulo 3: Expressão da proteína de fusão Adaptador:ModT. A presença deste módulo está vinculada com a metodologia SAM (sgRNA com aptâmeros de RNA). Caso a abordagem seja a clássica (apenas a dCas9 fusionada com ModT), ele não deve ser inserido no desenho do vetor. Este módulo apresenta três elementos: 1) promotor; 2) proteína de fusão Adaptador:ModT; e 3) terminador.

- a) **Promotor:** a escolha do promotor para a expressão da proteína deve ser alinhada com a escolha do promotor descrito no Módulo 1, seguindo exatamente os mesmos critérios de escolha, visto que a expressão da dCas9 e da proteína de fusão Adaptador:ModT devem ser sincronizadas. Se possível, ambos os Módulos devem apresentar a mesma região promotora.
- b) **Proteína de fusão Adaptador:ModT:** a escolha do adaptador deve ser alinhada com a escolha do aptâmero de RNA inserido no *RNA scaffold* do sgRNA, uma vez que o acoplamento adaptador-aptâmero é específico. Por exemplo, se o aptâmero escolhido for o MS2, a proteína adaptadora deve ser a MCP. Devido à ampla aplicação biotecnológica, a identificação de novos pares adaptador-aptâmero vem crescendo a cada dia (Zhang et al., 2019c). Com relação aos ModTs escolhidos, a maioria das construções que obtiveram sucesso

com a tecnologia dCas9 apresentavam tais domínios fusionados in tandem com o carboxiterminal da proteína adaptadora, devidamente separados por *linkers*. Não existe nenhum relato que limite o número de ModTs fusionados, mas a maioria das construções apresentam somente 1 ou 2. Além disso, de forma similar à proteína dCas9, a proteína de fusão necessita ser transportada para o núcleo para que desempenhe sua atividade. Dessa forma, é necessário inserir NLSs, seguindo os mesmos critérios apresentados no Módulo 1 para a dCas9. Inclusive, neste caso, o NLS é comumente inserido entre a proteína adaptadora e o primeiro ModT, funcionando também como um *linker*.

- c) **Terminador:** de forma similar ao que foi apresentado no Módulo 1, a escolha do terminador é mais livre. Qualquer terminador que seja reconhecido por uma pol II pode ser utilizado, como, por exemplo, o terminador *NOS*.

Módulo 4: Expressão do Marcador de Seleção. Este módulo é específico para cada espécie de planta, levando sempre em consideração a metodologia de transformação que será utilizada. Similar aos Módulos 1 e 3, três elementos são importantes: 1) promotor; 2) marcador de seleção, que na maior parte das vezes é um gene que confere resistência a um antibiótico ou herbicida, mas pode ser qualquer um gene repórter, tal como *gfp*, *gus*, entre outros; e 3) terminador. Visto que o desenho deste módulo pode variar, sugere-se somente que não se repitam promotores e/ou terminadores que já tenham sido utilizados anteriormente nos Módulos 1 e 3, a fim de evitar um possível silenciamento na planta (Figura 5).

Outras considerações: além das acima mencionadas, algumas informações extras são importantes no desenho do vetor:

- a) Na maioria dos laboratórios, a construção do vetor se dá por síntese química. Se for esse o caso, uma consideração importante a ser levada em conta é a otimização do *codon usage* dos genes que serão expressos nos Módulos 1, 3 e 4. Esse cuidado é extremamente relevante para que a tradução dos genes presentes nesses módulos seja também otimizada na planta de interesse.
- b) Mantenha todos os Módulos na mesma direção, a fim de evitar impedimentos físicos durante a transcrição.
- c) Sugere-se que sejam inseridos sítios únicos para enzimas de restrição entre cada módulo a fim de facilitar a manipulação do vetor final (se necessário).
- d) Como visto anteriormente, os sgRNAs podem ser expressos em simplex (1 sgRNA) ou em multiplex (2 ou mais sgRNAs). Para sistemas multiplex, cada sgRNA deve ser expresso individualmente, isto é, cada sgRNA deve ter seu promotor *U6* e seu terminador *U6*. Em vias de regra, para expressão de diferentes sgRNAs no mesmo vetor em que se encontra a dCas9, o Módulo 2 deve ser

repetido in tandem (de preferência separadas por sítios únicos de restrição). Cada cópia do Módulo irá apresentar um sgRNA diferente. Mesmo que não existam relatos do número máximo de sgRNAs que podem ser expressos em uma mesma construção, sistemas que obtiveram sucesso não ultrapassaram quatro sgRNAs por construção.

- e) Os *backbones* mais utilizados para a clonagem do cassete de transformação são vetores binários baseados no T-DNA de *A. tumefaciens*, mas este pode ser substituído por outro, dependendo da planta de interesse e da metodologia de transformação utilizada. Além disso, a clonagem do cassete de transformação no *backbone* escolhido e posterior amplificação do vetor final em bactéria pode ser dificultada pelo tamanho do cassete de transformação, que pode chegar a ter de 10.000 a 15.000 pares de bases. Caso o vetor escolhido como *backbone* seja baseado no T-DNA de *A. tumefaciens*, sugere-se que o Módulo 4 (referente ao marcador de seleção) se localize mais próximo à borda esquerda (LB, do inglês *Left Border*), a fim de otimizar a seleção de plantas transformadas nas quais o cassete de transformação foi inserido completamente no genoma-alvo.

Considerações finais

A tecnologia CRISPR/dCas9 aplicada à modulação da expressão gênica tem como principal diferencial a possibilidade de promover o controle transcricional em escala multigênica, simulando as redes de interação gênica inerentes às principais vias metabólicas celulares. Trata-se de uma ferramenta-chave para a biologia sintética e para o aprofundamento do conhecimento de vias relacionadas ao desenvolvimento e produtividade em plantas, bem como à resposta a estresses bióticos e abióticos, alvos frequentes da engenharia genética para o melhoramento de plantas. Além disso, CRISPR/dCas9 é uma tecnologia simples e funcional em sistemas transientes de transformação, o que permite uma rápida simulação e visualização dos efeitos da modulação da expressão do gene-alvo em vias metabólicas celulares. As diversas estratégias disponíveis, variando elementos como ModT e nucleases, permitem a adaptação e uso da técnica em diferentes modelos de estudos.

Como vantagem do sistema CRISPRa para a ativação transcricional de um gene endógeno, em comparação à estratégia de superexpressão, usando promotores constitutivos e a sequência codificadora completa do gene, tem-se que, muitas vezes, esse tipo de expressão ectópica omite os padrões de *splicing* do gene endógeno e mascara transcritos alternativos. Como consequência, o produto gênico pode ser instável, interagir com parceiros não convencionais ou ter sua localização

subcelular incorreta (Park et al., 2017). No entanto, ao se trabalhar com a tecnologia CRISPR/dCas9, alguns aspectos merecem atenção:

- a) Como a tecnologia CRISPR/dCas9 não induz alterações definitivas na região genômica-alvo, as plantas geradas serão sempre transgênicas. Caso o intuito seja o desenvolvimento de um produto biotecnológico, isso pode acarretar em um processo longo e oneroso para liberação comercial.
- b) Por depender de processos celulares de controle transcricional, o efeito na modulação da expressão do gene-alvo pode ser variável entre as linhagens transgênicas obtidas. Apesar de herdável ao longo das gerações transgênicas, o efeito na modulação pode também ser variável entre planta-mãe e plantas-filhas ou ainda sofrer influência de fatores ambientais.
- c) A tecnologia exige um conhecimento mínimo da sequência da região promotora-alvo, o que pode ser um fator limitador quando se trabalha com espécies sem informação genômica detalhada.

Nesse contexto, os constantes aprimoramentos da técnica, principalmente visando aumentar sua especificidade e eficiência, como, por exemplo, com a utilização de outros ortólogos Cas, permitirão o aumento da abrangência do seu uso, sobretudo em pesquisas com plantas cultivadas, uma vez que até o momento foram realizadas pesquisas apenas com espécies modelo.

Referências

- 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.; KORNBURG, 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.; EI-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, p. 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 sulfonyleurea 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 Biotechnolog**, 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. **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-guided 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 transcriptional 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>