

*Empresa Brasileira de Pesquisa Agropecuária
Embrapa Agroenergia
Ministério da Agricultura, Pecuária e Abastecimento*

Tecnologia CRISPR na edição genômica de plantas

Biotecnologia aplicada à agricultura

*Hugo Bruno Correa Molinari
Letícia Rios Vieira
Nathalia Volpi e Silva
Guilherme Souza Prado
José Hernandes Lopes Filho*

Editores Técnicos

Embrapa
*Brasília, DF
2020*

Edição de genoma via *non-homologous end joining* (NHEJ) e ribonucleoproteínas (RNP)

Guilherme Souza Prado
Thaísa Tessutti Pinheiro
Josias Correa de Faria
Rosana Pereira Vianello

Introdução

A edição de genomas é uma ferramenta que permite a manipulação do material genético de forma a induzir mutações em regiões de interesse para que o organismo apresente um fenótipo desejável, podendo, em alguns casos, não ser considerado transgênico. Devido a isso sua aplicação em plantas de interesse agrônomico tem despertado grande interesse, principalmente por simplificar etapas de regulamentação.

Existem quatro principais classes de nucleases que são empregadas nas novas tecnologias de edição gênica, sendo elas: meganucleases, *zinc-finger nucleases* (ZFNs); *transcription activator-like effector nucleases* (TALENs); e *Clustered Regularly Interspaced Short Palindromic Repeats* (CRISPR)-*associated* (Cas) (Gaj et al., 2016). As meganucleases são endonucleases que clivam o DNA em regiões específicas, reconhecendo sequências maiores que 12 pb (pares de base). A família de meganucleases LAGLIDADG engloba I-Crel e I-SceI, que são as primeiras enzimas que foram usadas para edição gênica. Como apenas alguns resíduos de aminoácidos fazem contato com os nucleotídeos, essas enzimas foram manipuladas para clivar genes em locais específicos (Paques; Duchateau, 2007). Além disso, as ZFNs são enzimas artificiais que também foram uma das primeiras utilizadas com o intuito de induzir mutações dirigidas em plantas. Essas enzimas foram geradas a partir da fusão entre domínios do tipo *Zinc-Finger* e domínios da enzima de restrição *FokI*. Da mesma forma que as outras nucleases utilizadas na edição gênica, a ZFN insere quebras na dupla fita (DSBs-*double strand break*) em locais específicos no DNA que vão ser reparados e, por meio de falhas durante o mecanismo de reparo, podem surgir mutações (Carroll, 2011). O principal problema do uso desse sistema está na alta toxicidade dessa enzima e também pelo fato de ela gerar muitos efeitos *off-target* (Cornu et al., 2008; Ramirez et al., 2008), que podem prejudicar funções de genes que não deveriam ter seu funcionamento alterado (Zhang et al., 2015). Com a consolidação da edição

gênica como ferramenta biotecnológica e a busca por alternativas mais viáveis, a técnica TALEN foi desenvolvida a partir da fusão de efetores TAL (*transcription activator-like*), proteínas derivadas de patógenos que infectam plantas e têm a capacidade de se ligar ao DNA, ao domínio da endonuclease *FokI* (Moscou; Bogdanove, 2009). TALENs podem ser gerados a partir de uma biblioteca de plasmídeos que permite a montagem da proteína de interesse de forma combinatória. Apesar de essa técnica apresentar vantagens em relação ao ZFN, por utilizar proteínas menos tóxicas e mais fáceis de se manipularem (Mussolino et al., 2014), ainda é considerada mais trabalhosa em relação à CRISPR (Nemudryi et al., 2014).

Entre as técnicas que tornaram a edição de genoma mais acessível e revolucionária, o sistema CRISPR/Cas se destaca. Isso se deve pela simplicidade, baixo custo e alta eficiência da técnica em relação a outras ferramentas de edição, como ZFNs e TALENs. Devido a isso, o sistema CRISPR/Cas se tornou amplamente utilizado em laboratórios de biologia molecular de forma rotineira nos últimos anos, e atualmente é a principal ferramenta de edição genômica utilizada. O sistema CRISPR se baseia na indução de mutações no DNA por meio de cortes na fita dupla, de maneira extremamente específica e direcionada. Esses cortes são promovidos por endonucleases da família Cas e direcionados pelos chamados RNAs guias (gRNA). Eles serão reparados por uma das duas principais vias de reparo que a célula apresenta quando a dupla fita é danificada: recombinação homóloga (HR) ou *non-homologous end joining* (NHEJ) (Cardona; Morales, 2014). Maiores detalhes podem ser encontrados no Capítulo 1.

Focaremos aqui na utilização do sistema CRISPR/Cas para indução de mutações via NHEJ. O mecanismo NHEJ atua na maior parte do ciclo celular e não se baseia na homologia para reparar o DNA, podendo causar mutações do tipo inserções ou deleções (*indels*) ou ainda substituições. Desse modo, a expressão dos genes presentes no local onde houve reparo por NHEJ pode ser, em geral, totalmente interrompida, e as proteínas codificadas podem não ser funcionais (Cui et al., 2019). Com o objetivo de esclarecer o processo de edição do genoma por CRISPR, este capítulo visa abordar os detalhes do mecanismo NHEJ, bem como o uso de sistemas de ribonucleoproteínas (RNPs) para promover a edição, além de fornecer um procedimento prático para o delineamento de experimentos.

Mecanismos de reparo genético

O sistema CRISPR atua por meio de cortes específicos na dupla fita do DNA em regiões genômicas de forma direcionada utilizando um RNA guia (gRNA) que, em seguida, é reparado, podendo causar a modulação de genes ou regiões promotoras.

Desse modo, esclarecer os principais sistemas de reparo utilizados pela célula para reconstrução da dupla fita torna-se fundamental para o entendimento da técnica. A importância de a informação genética ser transmitida através das gerações de forma precisa é evidente, tanto pelo eficiente sistema de replicação quanto pela complexidade e investimento da célula em um sistema de reparo especializado na correção dos diversos tipos de desordens do DNA.

Com esse objetivo, para sanar as alterações ocorridas no DNA, as células apresentam mecanismo de resposta aos danos. Um deles é o sistema DDR (*DNA Damage Response*), que é eficiente no reparo do DNA e atua por diversas vias e enzimas, dependendo do tipo de lesão ocorrida. Nesse mecanismo, quando as células sofrem danos, o processo de replicação é suspenso até que o reparo seja efetuado (Silva; Ideker, 2019). Para isso, ocorre uma cascata de sinalização que gera modificações na cromatina e deixa a extremidade danificada acessível às proteínas de reparo do DNA (Riches et al., 2008). Com a extremidade do DNA danificada exposta, o complexo Mre11-Rad50-NBS1 (MRN) se liga à DSB, recrutando proteínas-membros da família ATM (Ataxia Telangiectasia Mutada), que regulam a resposta ao dano por meio de fosforilações (Lavin; Kozlov, 2007). Desse modo, são ativados pontos de verificação do ciclo celular e remodelação da cromatina; tudo para que o DNA seja reparado antes de o ciclo celular continuar (Silva; Ideker, 2019). Entretanto, nos animais multicelulares, se o dano no material genético for irremediável, é ativada a apoptose para que a lesão não seja perpetuada (Norbury; Zhivotovsky, 2004).

Para que a célula realize o reparo da lesão no material genético, cinco principais vias de reparo são ativadas em diferentes momentos do ciclo celular. Quando o dano no material genético acontece em apenas uma das fitas (SSB, *single strand break*), os mecanismos de reparo são do tipo reparo de mal-pareamento (MMR, *mismatch repair*), reparo por excisão de bases (BER, *Base Excision Repair*) reparo por excisão de nucleotídeo (NER, *Nucleotide Excision Repair*). Porém, quando o DNA é danificado em ambas as fitas (DSB), podem ocorrer as vias de reparo por recombinação homóloga (HR) ou por *non-homologous end joining* (NHEJ) (Cardona; Morales, 2014), que são a base molecular da tecnologia CRISPR.

Non-homologous end joining (NHEJ)

A via NHEJ não necessita de DNA-molde e é responsável por 75% dos reparos no material genético das células animais (Mao et al., 2008). Essa via de reparo leva cerca de 30 minutos para se completar, enquanto a HR é finalizada dentro de 7 horas ou mais (Mao et al., 2008). Na via NHEJ, quando a DSB é reparada, geralmente são formados *indels* que, ao modificarem a sequência de DNA original, podem causar

a inativação de genes (Her; Bunting, 2018). O reparo por NHEJ emprega, para isso, uma série de fatores essenciais que trabalham para detectar as DSBs, bem como alinhar e reparar as extremidades para que possam ser novamente ligadas (Her; Bunting, 2018).

Mecanismo bioquímico

Na presença da DSB, o mecanismo de resposta aos danos DDR é acionado. Desse modo, uma cascata de fosforilação é ativada, gerando modificações na cromatina, que deixam a extremidade danificada acessível às proteínas de reparo do DNA (Riches et al., 2008). Em plantas (Steinert et al., 2016) e mamíferos (Hartlerode; Scully, 2009), tanto HR quanto NHEJ são acionadas.

NHEJ, por sua vez, pode ser subdividida em duas vias: clássica (c-NHEJ), que é a principal via de reparo e de rearranjo cromossômico (Lieber, 2010), e a *backup* (b-NHEJ), também chamada de via alternativa (a-NHEJ) ou *microhomology-mediated end joining* (MMEJ) (Deriano; Roth, 2013). Essa segunda via é ativada caso ocorram problemas funcionais com a c-NHEJ, porém não apresenta um mecanismo de reparo tão preciso, o que geralmente leva a translocações cromossômicas e deleções excessivas (Deriano; Roth, 2013) (Figura 1).

A via c-NHEJ pode ser dividida em três principais etapas: o reconhecimento da DSB, o processamento do DNA e a ligação das fitas (Yang et al., 2016). O reconhecimento da DSB depende das proteínas Ku70-Ku80 (Ku), um heterodímero com duas subunidades de 70 kDa e 80 kDa, que se entrelaçam em um anel que envolve a extremidade final da DSB, formando o complexo Ku:DNA (Figura 2). Este complexo mantém as extremidades finais da DSB relativamente próximas, para que não haja translocações adicionais ou fusões (Downs; Jackson, 2004).

Assim que o complexo Ku:DNA é formado em cada extremidade da DSB, são recrutados fatores proteicos, como a subunidade catalítica da proteína quinase dependente de DNA (DNA-PKcs), proteína de reparo de raios-X transversal complementar 4 (XRCC4), fator tipo XLF, PAXX (parálogo de XRCC4 e XLF) e DNA-Ligase IV (Lig4), que mantêm próximas as extremidades do DNA, formando uma ponte entre elas para que ocorra o processamento e a ligação (Davis et al., 2014).

Se a DSB gera extremidades simples, estas podem ser diretamente ligadas após o reconhecimento da DSB. Entretanto, na maioria das DSBs existem extremidades incompatíveis por modificações químicas ou incompatibilidade dos terminos das fitas de DNA, que acabam impedindo que as fitas sejam ligadas diretamente. Desse modo, o DNA deve ser processado e, para isso, exonucleases degradam as extre-

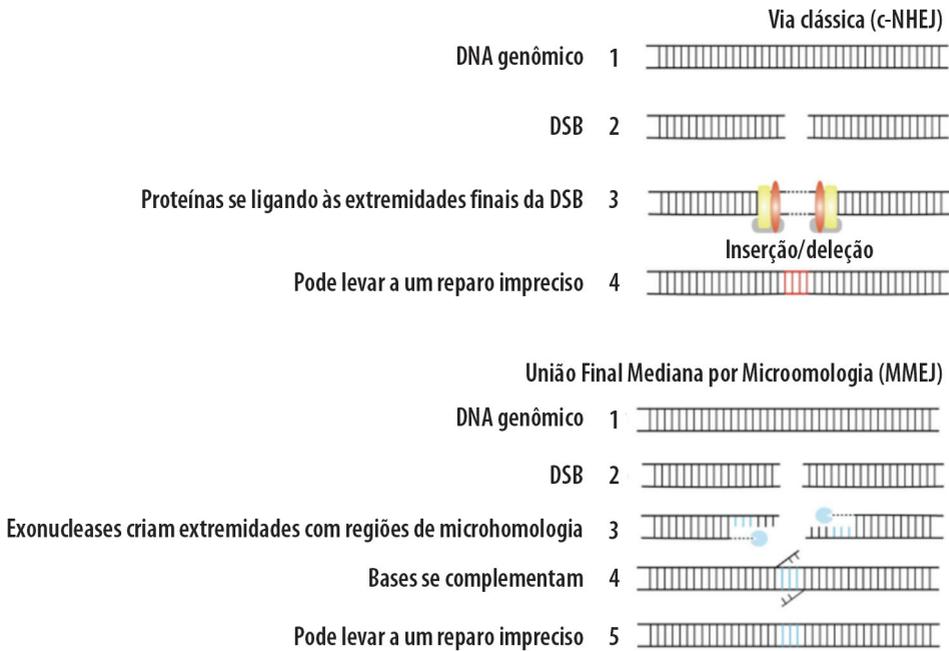


Figura 1. Vias de reparo por *non-homologous end joining* (NHEJ). (A) Via clássica *non-homologous end joining* (c-NHEJ), que pode levar a inserções ou deleções (*indels*). (B) União final mediada por micro-homologia (MMEJ), a qual sempre resulta em deleções.

Fonte: adaptado de Deriano e Roth (2013).

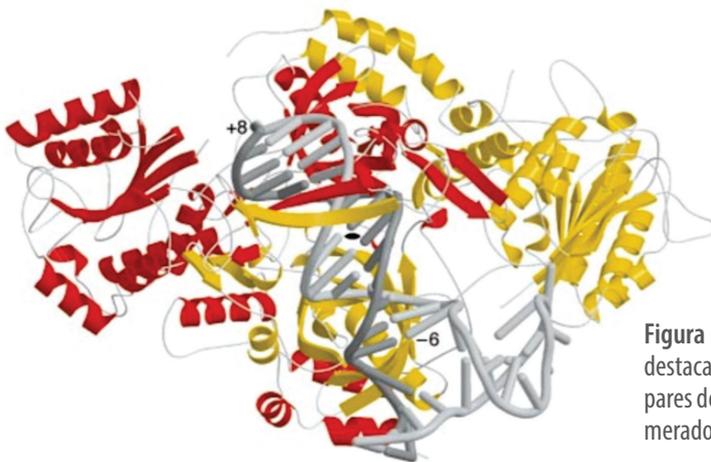


Figura 2. Complexo Ku:DNA. A proteína Ku70 é destacada em vermelho, e a Ku80 em amarelo. Os pares de bases terminais do duplex central são numerados +8 (extremidade do DNA quebrada) e -6.

Fonte: adaptado de Walker et al. (2001).

midades 5' ou 3', gerando regiões de micro-homologia. Esse processo, conhecido como recessão, permite que as duas extremidades se tornem compatíveis.

Durante a recessão das extremidades finais da DSB, as DNA-PKcs agem em complexo com a endonuclease Artemis que, após ser fosforilada pela DNA-PKcs, cliva ligações da fita simples e entre a fita dupla. Outros fatores envolvidos nesse processo são a polinucleotídeo quinase fosfatase (PNKP), aprataxina (APTX), APTX-polinucleotídeo quinase-fosfatase-like fator 1 (APLF ou PALF), werner (WRN), DNA polimerases e desoxinucleotídeo terminal transferase (TdT) (Lieber, 2010).

Após o processamento da DSB, as extremidades estão prontas para serem ligadas. Todo processamento e ligação do DNA é coordenado pelas DNA-PKcs, que fosforilam diversas proteínas (Jiang et al., 2015). Essas proteínas também podem ser autofosforiladas ou transfosforiladas por outros fatores, como ATM e ataxia relacionadas à telangiectasia (ATR) (Davis et al., 2014). Para que ocorra a ligação da dupla fita, são recrutadas ligases específicas e o complexo XRCC4-XLF-Lig4, que é o principal fator envolvido. Para a formação desse complexo, XRCC4 é ancorado ao complexo Ku:DNA-PKcs e, junto ao fator XLF, forma uma ponte com Lig4, promovendo a ligação da DSB (Andres et al., 2012).

Edição de Genomas via NHEJ – aplicações práticas

A aplicação do sistema CRISPR/Cas para edição de genomas em plantas tem sido usada com sucesso para a obtenção de cultivares melhoradas geneticamente. As características manipuladas com essa ferramenta abrangem desde o melhoramento nutricional, uso eficiente de nutrientes (Liang et al., 2014), resistência a doenças (Wang et al., 2014, Peng et al., 2017) e tolerância a estresses abióticos (Mushtaq et al., 2018) até o aumento da produtividade (Ma et al., 2016). Com isso, uma das frentes do melhoramento com CRISPR é o *knockout*, que elimina a função gênica por meio de *indels* derivados do reparo por NHEJ.

Visando à resistência de plantas de trigo (*Triticum aestivum*) ao fungo oídio (*Blumeria graminis* f. sp. *Tritici*), foi realizada a edição gênica via TALEN para introduzir mutações direcionadas em três genes homólogos que codificam as proteínas Mlo-Resistance Locus (MLO) (*TaMLO-A1*, *TaMLO-B1* e *TaMLO-D1*) e, também, a edição via CRISPR para *knockout* do gene *TaMLO-A1*. As proteínas MLO agem como reguladores negativos, reprimindo as vias de defesa da planta contra o oídio em tecidos não infectados (Piffanelli et al., 2002). Desse modo, quando os genes homólogos *MLO* foram editados via TALEN e CRISPR, observou-se o aumento da resistência das plantas de trigo ao fungo oídio em relação às plantas não editadas. Para a construção CRISPR, foi usado um vetor de expressão contendo a CDS da Cas9,

o promotor Ubiquitina 1 de milho e o sgRNA. O vetor foi introduzido no genoma de trigo via biobalística, obtendo-se eficiência de mutação nas plantas regeneradas igual a 5,6% (Wang et al., 2014).

Em outro trabalho, pesquisadores trabalharam com milho (*Zea mays*), o qual possui uma alta concentração de ácido fítico, um antinutriente que não é digerido por animais monogástricos e prejudica a absorção de nutrientes. Com o uso de CRISPR, foi realizado um *knockout* no gene *Zea mays Inositol Fosfato Cinase* (*ZmIPK*), o que gerou plantas editadas com baixa concentração de ácido fítico. Nesse estudo, a transformação genética via transfecção de protoplastos utilizando duas guias construídas para o direcionamento da Cas9 ao gene *ZmIPK* resultou em uma eficiência de mutação da ordem de 16,4% e 19,1% para cada guia utilizada (Liang et al., 2014).

Para soja (*Glycine max* (L.) Merrill), o primeiro trabalho de sucesso conduzido como prova de conceito com utilização do sistema CRISPR para realizar a mutagênese direcionada em raízes (*hairy roots*) demonstrou a eficácia da técnica por meio do desligamento de genes endógenos e do gene *bar* (Cai et al., 2015). Nesse estudo foram construídos sete vetores, um com a gRNAs para o transgene *bar* e outros seis com gRNAs direcionadas a dois genes endógenos de soja (*GmFEI2* e *GmSHR*), sendo três guias para cada gene (*GmFEI2-SP1*, *GmFEI2-SP2* e *GmFEI2-SP3* para o gene *GmFEI2* e *GmSHR-SP1*, *GmSHR-SP2* e *GmSHRSP3* para *GmSHR*). A expressão da Cas9 foi dirigida pelo promotor Ubiquitina de milho, as gRNAs foram dirigidas pelo promotor U6 de *Arabidopsis*, e a expressão do gene *GFP*, para a rápida visualização das *hairy roots* transgênicas, foi dirigida pelo promotor CAMV 35S. O método de transformação genética usado foi baseado em *Agrobacterium rhizogenes* e, como material vegetal, foram utilizadas plantas transgênicas homozigotas para o gene *bar* (para a edição do gene *bar*) e plantas selvagens, não transformadas geneticamente, para a edição dos genes *GmFEI2* e *GmSHR*. Para o transgene *bar*, foram identificadas 11 mutações nas 30 *hairy roots* transgênicas independentes analisadas, com frequência de 1,3% a 21,0%. Já para os genes endógenos, na guia *GmFEI2-SP1* foram identificadas mutações em 18 das 30 *hairy roots* transgênicas independentes, com frequência de *indels* entre 0,6% e 18,8%. Mutações no local da guia *GmFEI2-SP2* foram identificadas em 28 das 30 *hairy roots*, com frequência de *indels* variando de 1,0% a 95,0%. Para a guia *GmFEI2-SP3*, foram identificadas 3 mutações nas 30 *hairy roots* transgênicas independentes analisadas, com frequência de *indels* entre 5,4% a 27,0%. Já para o gene *GmSHR-SP2*, foram identificadas em 10 das 22 *hairy roots*, com frequência de *indels* de 8,7% a 30,0%. Para *GmSHR-SP1*, as mutações foram encontradas em 15 das 30 raízes analisadas, com frequência de *indels* variando entre 2,3% a 21,3%. Já no local *GmSHR-SP2* foram encontradas mutações em 10 das 22 raízes analisadas, com frequência de *indels* variando de 8,7% a 30,0% (Cai et al., 2015). Em *GmSHR-SP3* foram identificadas mutações em 10 das 28, com frequência variando de 2,8% a 28,7%.

Modulação e otimização de promotores

O uso de NHEJ com o objetivo de modular a expressão de genes também pode ser aplicado a elementos regulatórios, a exemplo das regiões promotoras. Essa estratégia faz com que a atividade transcricional, com base na modulação do acesso dos fatores de transcrição ao promotor, seja alterada e, como consequência, tem-se a mudança na expressão de genes (Cong et al., 2013). Quando sequências promotoras são modificadas por meio das técnicas de edição de genoma, dá-se o nome de engenharia promotora *in vivo* (IPE, do inglês *in vivo promoter engineering*) (Pandiarajan; Grover, 2018). Alguns trabalhos conduziram essa abordagem visando à geração de variantes de promotores com atividade transcricional alterada (Li et al., 2020).

Com o objetivo de obter plantas de arroz mais adaptadas a solos com alto teor de salinidade, a tecnologia CRISPR foi utilizada para gerar *indels* em regiões de unidades promotoras funcionais, os elementos *cis* regulatórios (CREs, do inglês *Cis Regulatory Elements*) do gene *OsRAV2*. Esse sistema foi estabelecido para confirmar que deleções nos CREs da sequência promotora do gene *OsRAV2* (nos elementos GT-1) estariam relacionadas à resistência ao estresse salino. Os genes *RAV* (*Related to ABI3/VP1*) codificam proteínas envolvidas na resposta ao estresse abiótico, sendo *OsRAV2* específico na resposta a ambientes com excesso de salinidade em arroz. Desse modo, foi construído um vetor contendo a Cas9 e um gRNA direcionado aos elementos GT-1 da região promotora do gene *OsRAV2*, e a transformação genética foi conduzida via *Agrobacterium tumefaciens*. Nesse estudo, 12 plantas foram editadas e todas apresentaram as mutações direcionadas desejadas, que promoveram a resistência das plantas a solos com alta concentração salina (Duan et al., 2016).

Em um trabalho que envolveu o cancro cítrico, causado por *Xanthomonas citri*, um patógeno responsável por perdas significativas no cultivo de *Citrus* spp., foi utilizado CRISPR em *Citrus sinensis* Osbeck, almejando a modulação do promotor do gene *Lateral Organ Boundaries 1* (*CsLOB1*), associado à suscetibilidade ao patógeno. O gene teve sua indução em resposta à infecção interrompida, e essa abordagem gerou plantas com maior resistência ao cancro cítrico em comparação com o tipo selvagem. Para a obtenção das plantas, foi realizada a construção de um vetor codificando a Cas9, além do promotor AtU6-1 dirigindo a expressão do gRNA direcionado à região promotora do gene *CsLOB1*. A transformação genética foi realizada por *A. tumefaciens*, e a eficiência de mutação foi entre 32.5% e 90.7% (Peng et al., 2017).

Ribonucleoproteínas (RNPs) nuclease-sgRNA

RNPs: sistemas DNA-free

A edição ou engenharia de genoma é, na maioria dos casos, conduzida utilizando sistemas de vetores plasmidiais carreando genes que, ao se integrarem no genoma do hospedeiro, codificam a expressão dos produtos necessários para o mecanismo de edição, sendo basicamente uma nuclease e o RNA-guia (gRNA), além de um gene marcador de seleção de transformantes, que confere resistência a um antibiótico ou um herbicida, quando o organismo-alvo é uma planta, e, eventualmente, genes-repórter (Xie; Yang, 2013).

Com isso, novas características são incorporadas ao hospedeiro tanto devido à mutagênese desencadeada no hospedeiro por meio da edição quanto devido aos genes exógenos que são introduzidos, o que configura o evento obtido como um OGM transgênico. Nesse caso, preocupações e problemas relacionados à regulamentação em biossegurança e à bioética podem se tornar obstáculos para que o produto biotecnológico desenvolvido se torne uma realidade mercadológica. Para contornar e/ou evitar barreiras que atrasam a chegada do organismo editado ao mercado, os pesquisadores normalmente recorrem a métodos que empregam a eliminação das sequências inseridas por meio de técnicas de hibridização e segregação mendeliana, de modo que o organismo editado passe a ser caracterizado apenas como GM, porém não como transgênico, uma vez que o material genético exógeno é perdido (Cyranoski, 2015).

Entretanto, essas técnicas podem se tornar inviáveis para a obtenção de plantas editadas não transgênicas que possuam processo de reprodução assexuada, a exemplo de uva (*Vitis* spp.), banana (*Musa* spp.) e batata (*Solanum tuberosum*), devido à ausência de recombinação que promoveria a eliminação dos transgenes na prole. Também podem ser pouco viáveis ou pouco vantajosas no caso de plantas perenes, como maracujá (*Passiflora* spp.), eucalipto (*Eucalyptus* spp.) e cereja (*Prunus* spp.), devido à própria falta de praticidade do processo de eliminação de transgenes, que se torna demasiadamente longo. Desse modo, plantas com ciclo de vida curto são mais passíveis de obtenção por meio desse método devido à maior praticidade (Woo et al., 2015).

Diante desse cenário, o rol de plantas-alvo se torna estreito, havendo a necessidade de desenvolver ou encontrar alternativas que contornem essa limitação. Sendo assim, uma das estratégias é o uso de sistemas DNA-free de edição gênica, dos quais

as ribonucleoproteínas (RNPs) são as mais conhecidas e cujo uso nos últimos anos segue expandindo para várias espécies (Kanchiswamy, 2016).

Em edição de genoma, as RNPs são complexos ribonucleoproteicos compostos por uma nuclease e um ou mais gRNAs. O gRNA é obtido por síntese *in vitro* ou síntese química, também podendo eventualmente ser otimizado para que sua inespecificidade seja reduzida, como por meio da truncagem para 17 resíduos de nucleotídeos, diferentemente dos 20 resíduos normalmente utilizados para Cas9, por exemplo. O gRNA também é estabilizado ao ser protegido contra a degradação por RNases devido à complexação com a nuclease (Kanchiswamy, 2016). A nuclease, por sua vez, pode ser obtida comercialmente a partir de fornecedores de empresas de biotecnologia ou expressa na forma recombinante em sistema heterólogo, comumente sistemas bacterianos baseados em *Escherichia coli*. Após a expressão, a proteína recombinante é purificada, preferencialmente por bioafinidade em coluna cromatográfica de níquel, com base em uma 6-His-tag fusionada a uma das extremidades da nuclease.

Conforme relatos da literatura científica, as RNPs são geralmente utilizadas para edição de genoma de células animais e, em comparação aos vetores plasmidiais, elas apresentam muito baixa citotoxicidade ao hospedeiro. Isso porque a citotoxicidade muitas vezes está associada ao próprio processo de transfecção plasmidial e a alguns reagentes utilizados para conduzir a transfecção.

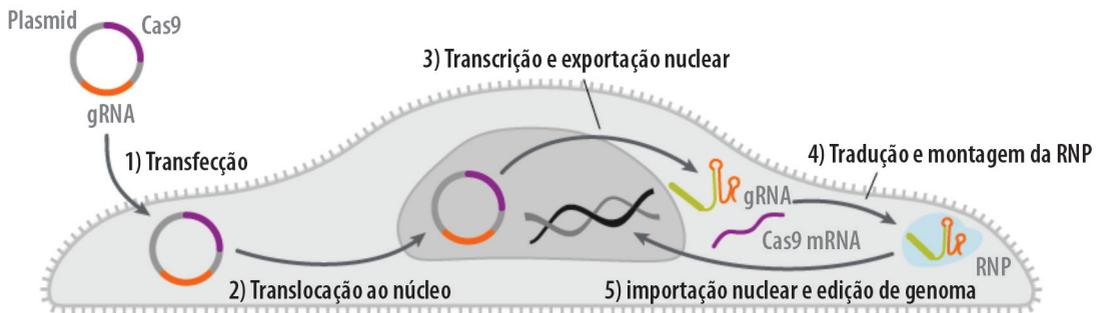
Algumas vantagens das RNPs em comparação com os plasmídeos são destacadas abaixo:

- a) O mecanismo plasmidial é mais complexo, uma vez que depende da maquinaria molecular de transcrição (para a nuclease e para o gRNA) e de tradução (para a nuclease), já que há a necessidade de integração dos genes de interesse no genoma-alvo, no núcleo, para que os transcritos sejam produzidos e, no caso da informação genética para a nuclease, exportado ao citoplasma para tradução gênica, com a nuclease retornando ao núcleo para realizar a edição. Observa-se, com isso, um trânsito contínuo e complexo entre diferentes compartimentos subcelulares, enquanto no uso de RNPs o mecanismo ocorre de forma direta, já que o complexo pré-formado se direciona imediatamente ao núcleo e executa sua função (Figura 3) (Amirkhanov; Stepanov, 2019; Wilbie et al., 2019).
- b) O uso de plasmídeos leva à integração gênica, a qual culmina na persistência da expressão dos agentes da edição, que permanecem funcionalmente ativos de forma prolongada, mantendo a atividade de edição, o que aumenta a probabilidade de ocorrerem efeitos *off-target*. Por outro lado, RNPs não

dependem de integração gênica para realizarem a edição, a qual ocorre imediatamente após o *delivery* do complexo no núcleo, com pico de mutagênese após cerca de 24 horas, sendo rapidamente degradadas e reduzindo, portanto, a ocorrência de efeitos *off-target* (Amirkhanov; Stepanov, 2019).

- c) Além da integração dos genes de interesse, o *backbone* dos plasmídeos também pode ser integrado no genoma do hospedeiro e persistir, mesmo após o uso de técnicas de hibridização, deixando “pegadas” (*footprints*) moleculares (Mao et al., 2018), o que pode se tornar outro problema na regulamentação da planta editada, por caracterizar um organismo transgênico.

Edição de genoma baseada e plasmídeo



Edição de genoma baseada e RNP

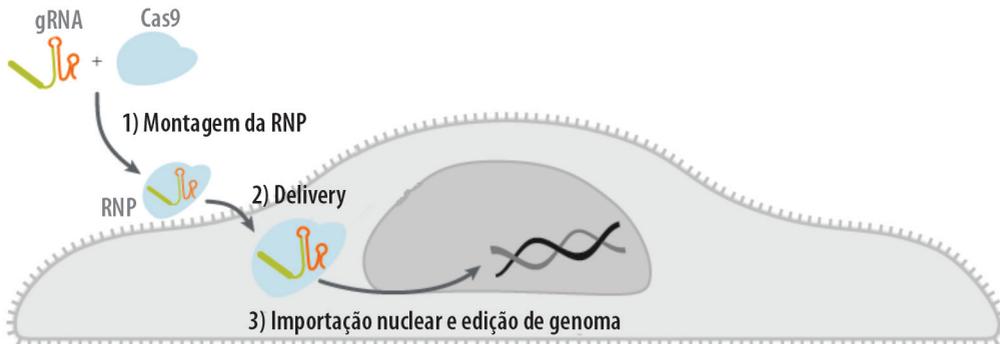


Figura 3. Plasmídeos versus RNPs na edição de genoma. (A) Vetor plasmidial carreando genes para codificação de uma proteína Cas9 e um gRNA, evidenciando etapas de um trânsito mais complexo entre diferentes compartimentos celulares até a edição do genoma. (B) RNP composta de uma Cas9 complexada a um gRNA, evidenciando uma etapa simples de trânsito direto ao núcleo para realizar a edição do genoma.

Fonte: adaptado de Spencer (2020).

RNPs não constituem a única estratégia DNA-free para edição de genoma. Outros sistemas também foram desenvolvidos e suas aplicações foram demonstradas não apenas para células animais, mas também para plantas de interesse econômico. Entre eles, os IVTs (*in vitro transcripts*, ou transcritos *in vitro*) e os vetores baseados em replicons virais são alternativas que também dispensam a integração genômica, sendo que cada um apresenta vantagens e desvantagens (Gil-Humanes et al., 2017; Liang et al., 2018). Em relação às várias vantagens oferecidas pelos vetores baseados em replicons virais, conferir o Capítulo 3.

Por meio do sistema de IVTs, são entregues moléculas de RNAm codificantes da nuclease e, embora os RNAm derivados de IVT apresentem a vantagem de não se integrarem ao genoma, as moléculas de RNAm estão mais sujeitas à degradação por RNases no meio intracelular do hospedeiro, resultando em menor estabilidade do sistema. Além disso, moléculas longas de RNA, como as codificantes da nuclease, são facilmente degradadas ao longo dos processos de manipulação em laboratório, dificultando o preparo. Já os vetores virais, frequentemente utilizados em sistemas baseados em replicon, exigem passos adicionais de clonagem e engenharia genética que são desnecessários no sistema RNPs (Gil-Humanes et al., 2017; Liang et al., 2018). Enquanto alguns vetores virais para células animais levam à integração do genoma viral no DNA do hospedeiro, em plantas isso não ocorre, e o replicon permanece ativo de forma episomal no citoplasma, dispensando as etapas de eliminação dos transgenes da planta editada.

Como exemplo de relatos da aplicação de RNPs em plantas, Woo et al. (2015) estabeleceram pela primeira vez a edição de genoma em sistemas vegetais via RNP, tendo como organismos hospedeiros: *Arabidopsis thaliana*, arroz (*Oryza sativa*), alface (*Lactuca sativa*) e tabaco-coiote (*Nicotiana attenuata*). Os autores realizaram a transfecção de protoplastos via PEG com RNPs pré-montados, obtendo eficiência de edição de até 46% (8,4% a 44% para *A. thaliana*, arroz e tabaco, e 46% para alface). Não foram detectados quaisquer efeitos *off-target*, e as edições se mantiveram estáveis na geração seguinte. As mutações geradas se demonstraram indistinguíveis das variações genéticas ocorridas de forma natural.

Svitashev et al. (2016) e Young et al. (2019), por sua vez, demonstraram a aplicabilidade das RNPs em milho (*Zea mays*), utilizando o bombardeamento de partículas para a transfecção do complexo em embriões imaturos, visando ao *knockout* dos genes *LIG1*, *Ms26* e *Ms45*, localizados em diferentes cromossomos. Mais especificamente, Young et al. (2019) estudaram os efeitos *off-target* dessa estratégia comparativamente ao uso de plasmídeos, demonstrando a ausência desses efeitos no genoma de milho. Liang et al. (2017) compararam RNPs com vetores plasmidiais em trigo (*Triticum aestivum*), verificando similaridade de atividade *on-target*, mas

efeito *off-target* cinco vezes menor quando foram utilizadas RNPs, corroborando o que foi mencionado sobre a correlação entre a atividade de edição prolongada na célula e o aumento dos casos de *off-target*. Por fim, Liang et al. (2018) compararam RNPs com IVTs para *T. aestivum*, demonstrando que ambos foram equivalentes em termos de eficácia, ao mesmo tempo eliminando a integração genômica e dando origem a plantas editadas DNA-free (não transgênicas). Outras espécies vegetais também foram transfectadas com RNPs, como, por exemplo: uva (*Vitis vinifera*) e maçã (*Malus domestica*) (Malnoy et al., 2016), petúnia (*Petunia x hybrida*) (Subburaj et al., 2016) e batata (*S. tuberosum*) (Andersson et al., 2018).

Nesse caso, os autores desses trabalhos utilizaram protoplastos como explantes para transfecção. Desses, apenas Andersson et al. (2018) regeneraram tecidos a partir dos protoplastos transfectados, obtendo brotações derivadas de calos induzidos. Para esse trabalho, a análise da eficiência de edição foi conduzida utilizando-se o próprio tecido foliar como fonte de amostra de DNA genômico, na qual concluíram que 9% dos eventos obtidos continham a mutação esperada em pelo menos um dos alelos do gene-alvo. Por outro lado, tanto Malnoy et al. (2016) quanto Subburaj et al. (2016) analisaram amostras de DNA genômico obtidas diretamente a partir dos protoplastos, sem passar por regeneração de plântulas. Com isso, Malnoy et al. (2016) verificaram eficiência de edição de 0,1% para uva e de 0,5% a 6,9% para maçã, enquanto Subburaj et al. (2016) obtiveram eficiência de 2,4% a 21% em petúnia.

Como é possível observar, as RNPs possuem aplicabilidade de grande alcance, com potencial para abranger outras abordagens úteis para o desenvolvimento de plantas editadas e economicamente importantes que sejam DNA-free. Desse modo, assim como plasmídeos, IVTs e vetores baseados em replicons virais, as RNPs são adequadas para gerar *knockout* via *indels* a partir do mecanismo de NHEJ, bem como podem ser adaptadas ao *knock-in* a partir do mecanismo de HDR, o qual será detalhado no próximo capítulo. Além disso, a regulação transcricional para edição e modulação de promotores via *indels*, visando à superexpressão ou ao *knock-down* de genes, bem como para a edição de base (*base editing*) e para a mais recente tecnologia de edição precisa de genoma, *prime editing*, é possível de ser executada em plantas utilizando RNPs, o que abre oportunidades para pesquisas que demonstrem pela primeira vez essa viabilidade e seu sucesso.

Entretanto, a limitação do uso de RNPs surge com abordagens que exigem a persistência do complexo nuclease-sgRNA atuando no núcleo. Isso porque as nucleases das RNPs são degradadas por proteinases intracelulares, perdendo sua atividade biológica tão logo sua tarefa de edição genômica seja cumprida. Portanto, estudos envolvendo a modulação e engenharia epigenética, que utilizam ativadores e inibidores via dCas9 em sistema CRISPRa/CRISPRi ou estudos voltados

ao imageamento com base na marcação genômica com sondas fluorescentes, por exemplo, se tornam inviáveis por meio das RNPs. Adicionalmente, a obtenção de RNPs exige a produção *in vitro* de transcritos de sgRNA, os quais, uma vez produzidos, possuem baixa durabilidade devido à sua alta sensibilidade (Liang et al., 2018). Além disso, pelo próprio fator mencionado a respeito da degradação, há uma janela curta disponível às RNPs para que a edição seja executada, o que costuma acarretar baixa eficiência de mutagênese por esse mecanismo. Por fim, ao se trabalhar com RNPs, não se faz uso de genes marcadores de seleção que auxiliariam na discriminação de explantes transformados, o que faz com que as RNPs sejam uma técnica mais trabalhosa, por demandar uma etapa adicional de caracterização molecular a fim de identificar as plantas editadas (Kanchiswamy, 2016).

Em suma, aprofundar e estender o uso das RNPs em espécies já testadas e em outras ainda não avaliadas, respectivamente, é de grande importância, dentro dos limites técnicos da abordagem, podendo impactar significativamente a agricultura mundial e, por consequência, a qualidade de vida humana.

Estratégias de *delivery* e internalização das RNPs

O *delivery* dos agentes moleculares da edição genômica é um passo fundamental para a obtenção de plantas editadas, e os parâmetros a ele associados estão diretamente ligados ao sucesso e à praticidade no desenvolvimento desses eventos. Com base nisso, uma ampla variedade de métodos de *delivery* para RNPs já é utilizada ou pode ser complementada com estratégias ainda não estabelecidas para plantas, embora já sejam amplamente aplicadas a células animais (Figura 4). Esses métodos podem ser classificados em:

- 1) **Transfecção simples:** abrange técnicas de introdução direta de RNPs – embora também seja utilizada para sistemas plasmidiais e outros mecanismos – na célula hospedeira por métodos físicos ou químicos, sem mediação de reagentes ou sistemas complexos e sofisticados para facilitar a transfecção. Exemplos:
 - a) Eletroporação: em se tratando de plantas, consiste na aplicação de pulsos de corrente elétrica em protoplastos, de modo a promover a abertura rápida de poros na membrana plasmática, permitindo a introdução das RNPs na célula.
 - b) PEG: o polietilenoglicol é um polímero mediador da internalização das RNPs, sendo frequentemente utilizado para transfecção de protoplastos como explantes.

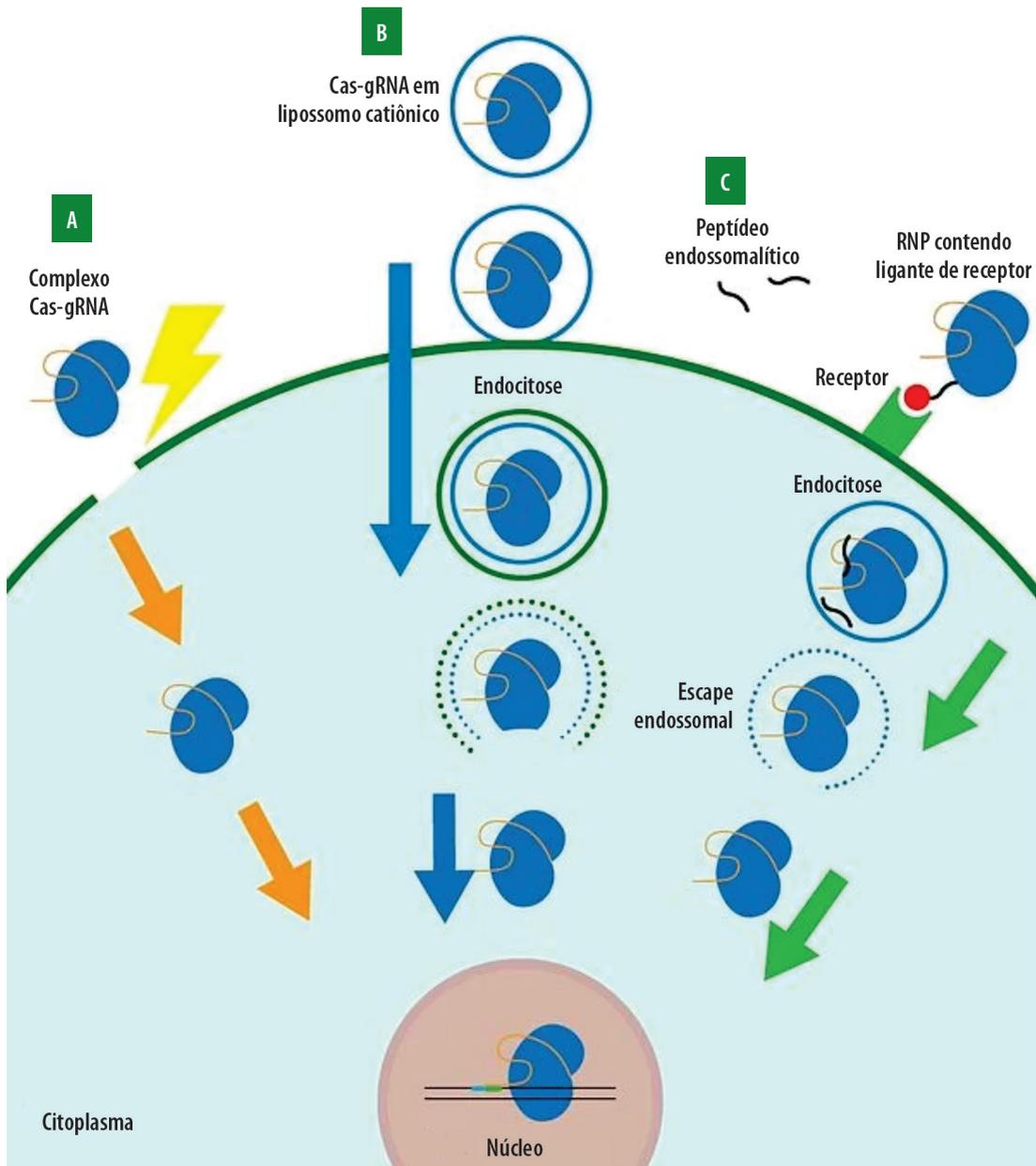


Figura 4. Métodos de *delivery* e internalização de RNPs. (A) Transfecção simples, exemplificada pela eletroporação. (B) Sistemas nanoestruturados, exemplificados pelo encapsulamento da RNP em lipossomas catiônicos. (C) Mecanismo ligante-receptor, exemplificado pela RNP fusionada a um ligante peptídico e endossomalítico. Todos os métodos culminam na edição do genoma nuclear, tendo ou não como passo intermediário a endocitose.

Fonte: adaptado de Hampstead (2020).

c) Bombardeamento de partículas (biobalística): método majoritariamente utilizado para plantas, consistindo na aceleração de micropartículas compostas, geralmente, de ouro ou tungstênio, a velocidades acima de 1.500 km/h, por meio do equipamento *gene gun*. Nesse procedimento, o complexo RNP é preparado e precipitado sobre as micropartículas, seguindo protocolos já bem-estabelecidos. Trata-se de uma técnica classicamente utilizada na transformação genética de plantas. Mais informações sobre a técnica podem ser verificadas no Capítulo 5 do livro *Manual de transformação genética de plantas* (Vianna et al., 2015).

2) **Sistemas nanoestruturados:** abrangem polímeros ou outras substâncias, produzidos na forma de partículas em escala nanométrica, que agem como nanocarreadores de RNPs (Zuris et al., 2015). Li et al. (2015) revisaram potenciais vetores não virais na forma de sistemas nanoestruturados ou nanomateriais para o *delivery* de CRISPR/Cas, os quais podem perfeitamente ser utilizados em um futuro próximo para edição em plantas via RNPs, como:

a) Lipossomos catiônicos: bicamadas lipídicas esféricas com cargas positivas, habilitando a ligação a nucleases negativamente carregadas, que, por sua vez, estão complexadas ao sgRNA. Podem ser compostos por DOTAP (1,2-dioleoil-3-trimetilamoniopropano), DOTMA (cloreto de N-[1-(2,3-dioleiloxi)propil]-N,N,N-trimetilamônio), DOSPA (trifluoroacetato de 2,3-dioleiloxi-N-[2(esperminacarboxamido)etil]-N,N-dimetil-1-propilamina) e DMRIE (brometo de 1,2-dimiristiloxipropil-3-dimetil-hidroxietilamônio).

b) Nanopartículas lipídicas sólidas (NLS): nanopartículas constituídas de lipídios com alto ponto de fusão, especialmente ácidos graxos (ex.: ácido esteárico e ácido palmítico), com um *core* sólido e cobertas por moléculas de surfactantes.

c) Nanopartículas poliméricas: polímeros em forma de partículas nanométricas, podendo ser compostos por PEI (polietilenimina), PLL (poli-L-lisina) e PDMAEMA (poli[2-(dimetilamino)etilmetacrilato]), por exemplo.

d) Nanopartículas de quitosana: nanopartículas compostas por um polissacarídeo derivado da desacetilação da quitina e com propriedades catiônicas.

e) Dendrímeros: polímeros altamente carregados positivamente, funcionando como “esponjas de prótons”, com alta afinidade para moléculas negativamente carregadas. A poliamidoamina (PAMAM) é um dos principais exemplos.

f) Nanopartículas de sílica (Torney et al., 2007).

g) Nanopartículas magnéticas (Zhao et al., 2017).

h) Nanotubos de carbono (Demirer et al., 2019).

Embora o uso de RNPs para edição de genoma em plantas seja de ocorrência esporádica, seus relatos na literatura majoritariamente envolvem a utilização de transfecção simples. Apenas recentemente, Liu et al. (2020) conduziram a transfecção de protoplastos de tabaco (*Nicotiana tabacum* cv. Bright Yellow-2) com base em sistemas nanoestruturados para lipofecção, utilizando dois reagentes lipídicos (Lipofectamina 3000 e RNAiMAX) para a montagem de lipossomas catiônicos, obtendo eficiência de *delivery* de 66% e 48% para esses reagentes, respectivamente. Para a Lipofectamina 3000, a eficiência de edição foi de 6%. Embora a lipofecção seja bem-estabelecida para edição de genoma em células animais, apenas nesse estudo foi estabelecida uma transição de nanomateriais para CRISPR/Cas aos sistemas vegetais.

- 3) **Mecanismo ligante-receptor:** abrangem a lógica de que a nuclease responsável pelo processo de edição pode ser fusionada a ligantes peptídicos ou proteicos de receptores celulares específicos (Rouet et al., 2018). Embora haja relatos do uso dessa abordagem para edição em células animais, sua aplicabilidade para CRISPR em plantas ainda não foi demonstrada, e tampouco os autores na literatura científica apresentam perspectivas nesse sentido.

Excluindo o bombardeamento de partículas, para todos os métodos e técnicas utilizados para *delivery*, é importante frisar a necessidade de se estabelecerem protocolos que facilitem o uso desses sistemas, principalmente para regeneração *in vitro* de plantas a partir de protoplastos, uma vez que são os explantes mais utilizados para conduzir esses procedimentos. Entretanto, sabendo que algumas espécies são recalcitrantes à regeneração a partir de protoplastos, o desenvolvimento de tais protocolos seria inviável, devendo-se optar por outras estratégias baseadas em tipos alternativos de explantes. De qualquer forma, para aquelas espécies para as quais esse procedimento é possível, surge esse benefício ao se trabalhar com RNPs.

Procedimento prático

Parte 1: *in silico*

O primeiro passo para realizar a edição genômica via CRISPR/Cas é escolher alvos genômicos da espécie desejada a ser editada. Essa escolha servirá de base

para identificar em seguida, dentro dos respectivos genes, sítios com sequências complementares ao gRNA utilizado para guiar a nuclease, desde que atendam aos requisitos da respectiva sequência PAM exigida pela enzima para a clivagem. Além disso, torna-se essencial realizar a predição de potenciais sítios *off-target*, de modo a sugerir gRNAs candidatos mais seguros quanto à especificidade genômica. Por fim, escolher particularmente uma estratégia adequada de *delivery*, a depender do tipo de explante necessário para a transfecção, bem como um vetor de expressão adequado com configuração apropriada de seus elementos genéticos, são a etapa final e fundamental – embora variem conforme os objetivos e a espécie de cada estudo – de um conjunto de procedimentos que deve ser conduzida de forma complementar à análise *in silico*, conforme será detalhado a seguir.

Identificação e seleção de alvos genômicos

Visando ao *knockout* em sequências codantes (CDS) ou à regulação transcricional de promotores (gerando variantes de promotores), ambos via NHEJ, é necessário que o gene-alvo esteja caracterizado quanto às sequências nucleotídicas que o compõem, incluindo seus elementos regulatórios. Nesse caso, é possível obter as sequências nucleotídicas no banco de dados anotados do National Center for Biotechnology Information (NCBI)¹, utilizando termos que identificam a espécie e/ou a sequência, bem como números de acesso, quando possível. Outra possibilidade é obter a sequência completa do gene a partir de sequências menores (*query*), por meio de similaridade com sequências depositadas (*subject*) no banco de dados. Para isso, essa informação pode ser obtida com utilização das diferentes modalidades do *Basic Local Alignment Search Tool* (BLAST)² para diferentes tipos de sequência.

Entretanto, o recurso BLAST pode ser acessado indiretamente e de forma personalizada para muitas das principais espécies vegetais de importância econômica, como, por exemplo, no banco de dados Phytozome v12.1³ (Figura 5).

Com a identificação da sequência-alvo (Figura 6), a composição nucleotídica pode ser detalhada na aba “Sequence”. A subseção “Transcript Sequence” dispõe os componentes do transcrito maduro diferenciados por cores entre 5’-UTR, CDS e 3’-UTR, o que pode ser útil para o delineamento de gRNAs para ambas as abordagens (CDS e elementos regulatórios). Da mesma forma que no NCBI, a sequência pode ser obtida no formato FASTA e armazenada para análise de gRNAs.

¹ Disponível em: <https://www.ncbi.nlm.nih.gov/nucleotide>

² Disponível em: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>

³ Disponível em: <https://phytozome.jgi.doe.gov/pz/portal.html>

Figura 5. Phytozome v12.1. Mecanismo de busca por seqüências ao longo do genoma de diferentes espécies vegetais possíveis. Uma barra de busca para a inserção da seqüência nucleotídica da espécie selecionada, à direita, é mostrada, possibilitando a busca mediada por BLAST, a partir da função “GO”.

Figura 6. Busca por seqüências de transcrito maduro no Phytozome v12.1. Etapa de diferenciação dos elementos genéticos que compõem a seqüência nucleotídica do gene após a transcrição, a qual servirá como modelo para similaridade na análise de gRNAs candidatos. Em verde, a seqüência relativa à região 5'-UTR (útil para a seleção de gRNAs visando à modulação transcricional de promotores). Em azul, as seqüências relativas dos éxons vizinhos (úteis para a seleção de gRNAs visando ao *knockout* do gene), separados por tons claros e escuros.

Paralelamente, é recomendado que, no caso de CDS como alvo, a sequência da cadeia polipeptídica também seja acessada e armazenada em qualquer ferramenta de edição de texto, para outro tipo de análise, como será mostrado a seguir. Nesse caso, no Phytozome é possível acessar a respectiva sequência na subseção “Peptide Sequence”, dentro da mesma aba “Sequence”, enquanto no NCBI as respectivas sequências são mostradas na página da tabela de características de cada acesso, na seção “CDS”.

Desenho de gRNA

Possuindo o sítio-alvo delimitado *in silico* e contendo seus elementos genéticos, deve-se, dependendo da estratégia, desenhar um ou mais gRNAs complementares a um elemento regulatório ou à CDS, levando em conta a sequência-alvo em nível genômico. Com o objetivo de promover o *knockout* de um gene, deve-se eliminar a funcionalidade do produto da expressão por meio da truncagem da cadeia polipeptídica codificada. Portanto, deve-se partir de duas premissas fundamentais:

- a) O *knockout* deve ocorrer com base na mutagênese em regiões mais distais a 5'; isto é: quanto mais *upstream* da CDS ocorrer a clivagem, menor a probabilidade de que os domínios proteicos nativos e funcionais permaneçam sendo expressos com base na sequência remanescente *in frame*.
- b) O pareamento do gRNA com o genoma deve ocorrer em regiões restritas a éxons específicos e únicos, isto é: é crucial que a sequência PAM e o sítio de clivagem estejam localizados no mesmo éxon. Portanto, localizar essas sequências em íntrons ou em junções éxon-íntron pode tornar a sequência candidata irrelevante para uma edição genômica eficaz e com resultados concretos.

Considerando esses fatores para iniciar a segunda etapa, uma ferramenta bastante prática, intuitiva e ampla para a busca e predição de gRNAs candidatos é a CRISPOR⁴ (Concordet; Haeussler, 2018), que será detalhada a seguir, instruindo o leitor a selecionar as melhores guias para os propósitos supracitados:

- 1) Ao acessar o CRISPOR (Figura 7), deve-se inserir a sequência-alvo na barra de sequência nucleotídica em “Step 1”. Caso prefira, é possível atribuir um nome à sequência, nessa mesma seção.

Obs.: a sequência a ser buscada pode ser genômica ou de cDNA. Em ambos os casos, deve-se atentar à segunda premissa exposta acima, descartando gRNAs candidatos que pareiem com íntrons ou que abranjam mais de 1 éxon, respectivamente.

⁴ Disponível em: <http://crispor.tefor.net/>

- 2) Em seguida, selecionar, na seção “Step 2”, entre os 536 genomas disponíveis, o genoma do hospedeiro-alvo que se pretende editar. Caso o genoma da espécie não esteja disponível, é possível desabilitar a seleção de um genoma específico, selecionando a opção “No genome”. Nesse caso, o *output* terá menor riqueza de dados, sem detalhes de especificidade.
- 3) Por fim, na seção “Step 3”, selecionar o PAM correspondente à sequência PAM da nuclease com a qual se irá trabalhar na bancada e cuja abreviação é mostrada ao lado de cada sequência genérica. A seleção do PAM já inclui o comprimento padrão das sequências de gRNA a serem retornadas como *output*, o que varia conforme a nuclease utilizada na edição. Após isso, clicar em “Submit” para iniciar a busca.



CRISPOR ([paper](#)) is a program that helps design, evaluate and clone guide sequences for the CRISPR/Cas9 system. [CRISPOR Manual](#)
 Apr 2020: [bugfixes](#), [ShCas9 enzymes](#), [21bp guides for Cpf1](#) [Full list of changes](#)

Step 1

Planning a lentiviral gene knockout screen? Use [CRISPOR Batch](#)

Sequence name (optional):

Enter a single genomic sequence, < 2000 bp, typically an exon

[Clear Box](#) - [Reset to default](#)

```

ATGGTAAAGTCCACCAACATTCCAGCCACACCAAGGCCGTGTCTCATCATCGAGCAACGGTTTC
ATAGGCCAAGTTGTCACTGAGGCAAGCCTCTCACCAGCACCCCACTTACTGCTCTCAAGGCCACCA
CCCTTGTCCCTTCCAAGGATGCCATTGTTAAACCTTCCAAGAAAAGGGTCCATGATCATTGATGGT
GTGATAAACCAACAAGACTTTGTGGAAAAGATTTGAAAGAGCATGAATAGACATTGTGATTTCTGCT
ATCGGGCTAAGAGCTTGCCTGGATCACTTACTGGTGGSHGOCATGAAATCCGTTAAGCTATTAGG
AGGTTTTGGCTTCAGAGTTGGACATGATGTGGACAGAGCTGATCCTGTGGAGCAGGTCTAAACAATG
TACAAGGAAAGCAATTGGTTAGGCCGTGGTTGAGCAATCTGGGGTCCATACACCAATATTTGTGTC
AATTCATTGCTTCATGGCCCTTACTATGACAATTGCCCCCATCTCAGCTTCTCCACCCCTGGATCAA
TTGCCAATATATGGTCATGGTAATGTCAAAAGCTTACTTTGTTGATGGCATGATATTGGAAAAGTTCACA
ATGAGGTTATTGATGATGTAAGAAATCAACAATAATGTTCAATTTGACCATCTAAGAATCTGCTAC
AGCATCAATGAATCTGCTTCTTGTGGGAAATGAAATTTGGCCGCAATCCCAAGAGTGAACATATCA
GAAGATGATCTTCAAGCAGCAGCTGCAGAAAATGCAATCCCAAGATTTGTGGCATCATTCACTCAT

```

Text case is preserved, e.g. you can mark ATGs with lowercase.
 Instead of a sequence, you can paste a chromosome range, e.g. chr1:11,130,540-11,130,751

Step 2

Select a genome

We have 536 genomes, but not yours? Search [NCBI assembly](#) and send a GCF_/GCA_ID to [CRISPOR support](#).

Step 3

Select a Protospacer Adjacent Motif (PAM)

[SUBMIT](#)

Figura 7. Busca por gRNAs candidatos no CRISPOR. Janela de *input* da sequência-alvo para desenho dos gRNAs candidatos, detalhando as barras de inserção da sequência (Step 1), de seleção das opções de genomas-alvo (Step 2) e de seleção da sequência PAM (Step 3).

- 4) Logo após, novas informações serão mostradas em uma nova página (Figura 8). Primeiramente, a sequência-alvo será disponibilizada dentro de uma caixa de cor cinza, com resíduos identificados de 10 em 10 nucleotídeos. Sob os resíduos, aparecerão as sequências iniciais de gRNAs sobrepostos ao alvo. A depender da especificidade de cada gRNA, estes se apresentarão com três diferentes cores: verde (sequências com alta especificidade), amarelo (média especificidade) ou vermelho (baixa especificidade). Recomenda-se fortemente que apenas sequências com cor verde sejam selecionadas. As sequências podem estar alinhadas à sequência *Forward* (iniciando com nucleotídeos e

terminando com traços) ou *Reverse* (iniciando com traços e terminando com nucleotídeos) em relação à sequência-alvo do *input*.

Obs. 1: quando se trabalha com nucleases que clivam a sequência-alvo em ambas as fitas, é indiferente selecionar candidatos *Forward* ou *Reverse*.

Obs. 2: é possível clicar em cada sequência da caixa cinza para ser direcionado às informações respectivas na tabela que contém todos os gRNAs.

- 5) A tabela de informações, apresentada na seção *Predicted guide sequences for PAMs*, contém os principais dados sobre todos os gRNAs candidatos encontrados conforme os parâmetros estabelecidos. Essa tabela pode ser exportada e salva como arquivo com extensão .xls, ao se clicar nas diferentes opções logo acima dela (*Download as Excel tables*). Quanto aos parâmetros da tabela, os gRNAs listados são dispostos em ordem decrescente de especificidade (*on-target*, *MIT Specificity Score*, que varia de 0 a 100), sendo esta inversamente proporcional à probabilidade de ocorrência de efeitos *off-target* usando a respectiva guia. Na primeira coluna da tabela (*Position/Strand*), é informada a posição e a fita da sequência-alvo que contém complementariedade ao gRNA. Na segunda coluna (*Guide sequence + PAM + restriction enzymes*), são demonstradas as sequências de cada guia, contendo o PAM, bem como os sítios de restrição presentes e passíveis de clivar cada guia. Nessa coluna também é indicada, quando for o caso, uma observação de que a guia pode não ser adequada por conter sinais de terminação na forma de resíduos sequenciais de timina (T_n). Também é possível restringir guias iniciadas com resíduos nucleotídicos específicos (essa restrição geralmente é utilizada ao se trabalhar com promotores U6 e U3 para dirigir a expressão do sgRNA, embora, mesmo ao se selecionarem tais promotores para dirigir a expressão, esses resíduos possam ser adicionados artificialmente à extremidade 5' de qualquer sgRNA para torná-los compatíveis ao respectivo promotor). As colunas *CFD* e *Specificity Score* são relativas aos *scores* de especificidade e de eficiência de clivagem, porém de menor utilidade para a escolha dos gRNAs. A coluna *Outcome* representa parâmetros relacionados ao resultado da clivagem, como *out-of-frame* e *frameshifts*, ocasionados por deleções e *indels*, respectivamente. Nesse caso, os valores também variam de 0 a 100, com o número representando a porcentagem teórica de eventos/clones a possuírem esses resultados de edição. A coluna *Off-targets for 0-1-2-3-4 mismatches* aponta o número de casos de sequências não alvo no genoma hospedeiro que contém de 0 a 4 *mismatches* (mal-pareamentos), nessa ordem, servindo de parâmetro crucial para definir as melhores guias que representem menor risco de gerar efeitos *off-target*. Quanto menor o número de casos, especialmente se houver

poucos ou nenhum caso com o menor número de *mismatches*, mais específica será a guia. É preferível que as guias selecionadas não possuam quaisquer casos com menos de 3 *mismatches*, especialmente na região *seed* do gRNA. A última coluna detalha a posição genômica de cada um dos casos *off-target*, detalhando o pareamento ao se passar o cursor do mouse sobre cada caso e marcando com asteriscos os *mismatches* entre os resíduos, o que facilita constatar se a ocorrência dos *mismatches* está dentro ou fora da *seed*.

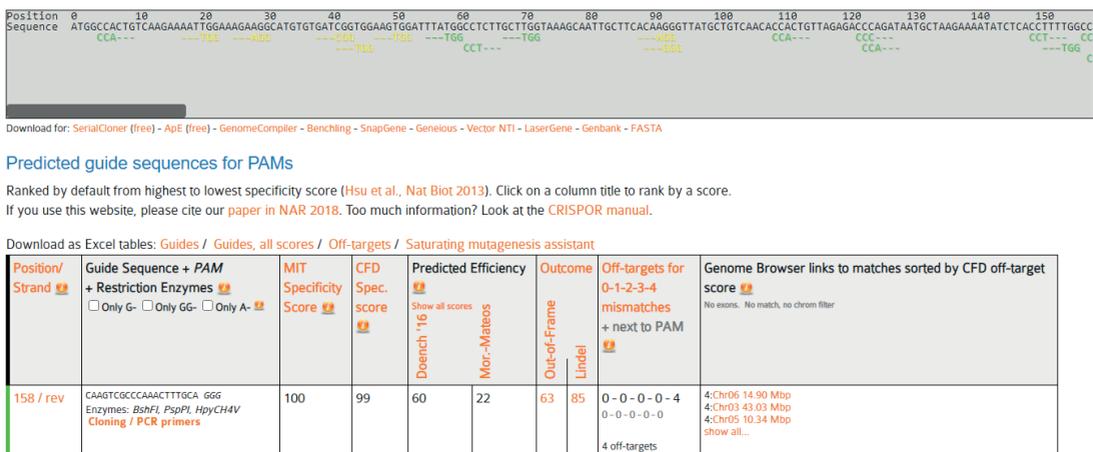


Figura 8. Resultado da predição de gRNAs candidatos no CRISPOR. Janela de *output* das seqüências de gRNAs candidatos, evidenciando a sobreposição de cada um na seqüência-alvo fornecida, sua composição nucleotídica, bem como localização na respectiva fita e parâmetros de *on-target* e *off-target*, detalhando cada um dos casos de potenciais *off-targets*.

Além do CRISPOR, para algumas espécies vegetais – especialmente as de grande importância econômica, como soja, arroz e milho –, é possível desenhar gRNAs com outras ferramentas computacionais, a exemplo da CRISPR-P v2.0⁵ (Lei et al., 2014). Nesse caso, basta acessar a página principal do programa e clicar na aba *Submit*. Na nova janela (Figura 9), deve-se selecionar a espécie-alvo, a seqüência PAM, o promotor U6 ou U3 e o comprimento da seqüência desejada de gRNA a ser retornada (podendo variar de 15 a 22 resíduos conforme *default* do programa), também inserindo a seqüência nucleotídica a ser editada. A seqüência do *scaffold* de sgRNA já se encontra pré-estabelecida. Após isso, deve-se clicar em *Submit* para a busca de gRNAs.

Após a submissão, uma nova página (Figura 10) será aberta com o detalhamento das guias candidatas para seleção. Nessa página, os gRNAs estarão representados

⁵ Disponível em: <http://crispr.hzau.edu.cn/cgi-bin/CRISPR2/SCORE>

CRISPR design

PAM

snoRNA promoter U6 U3

RNA Scaffold

Guide Sequence Length

Target Genome Data from Bolbase

Locus Tag eg: Bol000102, (input % for some tips)

OR

Position eg: Scaffold001737.614.785

OR

Sequence

Figura 9. Busca por gRNAs candidatos no CRISPR-P v2.0. Janela de *input* da sequência-alvo para desenho dos gRNAs candidatos, na aba *Submit* da página, detalhando as barras de seleção da sequência PAM, do comprimento da guia, do genoma-alvo (espécie) e de inserção da sequência-alvo nucleotídica.

graficamente em alinhamento à sequência no painel superior, com as sequências *Forward* e *Reverse* em cores distintas. No painel inferior, são elencadas as sequências candidatas, com vários parâmetros informados à esquerda (%GC, *on-target score*, composição nucleotídica, presença ou não em região codante). À direita, o número de casos de *off-targets* é informado, com o detalhamento de cada um deles quanto à composição nucleotídica (com *mismatches* destacados em vermelho), número de *mismatches*, localização cromossômica e presença ou não em região codante (cada caso é classificado como *intergenic* ou CDS e, neste último caso, a ID do gene correspondente é apontada na coluna ao lado). Também são informados os sítios de restrição presentes na sequência de cada gRNA predito, os quais podem ser úteis na etapa de subclonagem das guias no vetor de expressão, por exemplo.

Adicionalmente, é possível desenhar mais de um gRNA para o mesmo gene-alvo, o que assegura maior probabilidade de ocorrência do *knockout* do referido gene, além de se poderem excisar fragmentos de grande comprimento devido à produção de pelo menos dois sítios de DSB. Indo mais além, gRNAs múltiplos podem ser desenhados e utilizados, geralmente clonados no mesmo vetor de expressão, e não necessariamente direcionados ao mesmo gene, mas a genes-alvo distintos (Cong et al., 2013), em uma estratégia denominada “engenharia de genoma multiplex”. A engenharia multiplex permite, portanto, modular a expressão de vários genes

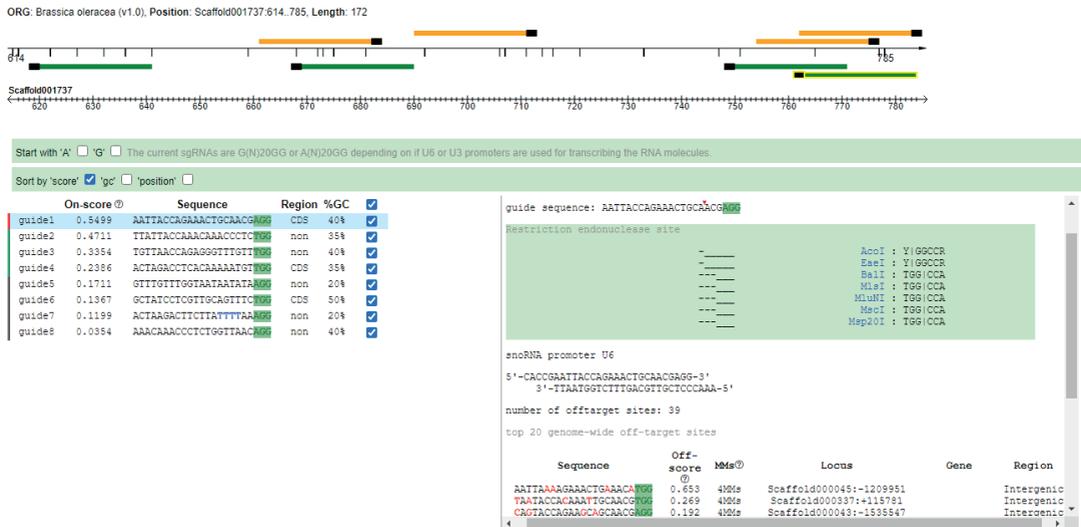


Figura 10. Resultado da predição de gRNAs candidatos no CRISPR-P v2.0. Janela de *output* das seqüências de gRNAs candidatos, evidenciando a sobreposição de cada um na seqüência-alvo fornecida, sua composição nucleotídica, bem como sítios de restrição presentes e parâmetros de *on-target* e *off-target*, detalhando cada um dos casos de potenciais *off-targets*.

associados a caracteres bastante discrepantes ou que estejam relacionados entre si convergindo para um caractere em comum.

Para conduzir a abordagem multiplex, é comum que os gRNAs desenhados sejam inseridos no vetor de expressão de forma que fiquem dispostos em cassetes policistrônicos. Para isso, além da configuração clássica de diferentes gRNAs dispostos em cassetes de expressão distintos (ou seja, de forma monocistrônica), com expressão dirigida pelos respectivos promotores, as unidades de gRNA podem ser intercaladas por seqüências de RNA transportador (RNAt), em uma estratégia que se baseia no processamento endógeno de moléculas de RNAt por RNases para excisão de transcritos em condições *in vivo*, promovendo a liberação dos gRNAs a elas fusionados e permitindo, assim, que, uma vez livres no núcleo, possam, paralelamente entre si, se associar à nuclease para guiá-la a sítios-alvo diversos no DNA genômico, também aumentando a eficiência da edição do genoma (Xie et al., 2015; Wolabu et al., 2020). Entretanto, o sistema de processamento baseado em RNAt-gRNA se fundamenta no uso de promotores para RNA Polimerase III para dirigir a expressão desses genes sintéticos (He et al., 2017). O uso desses promotores, entretanto, a exemplo das variantes de U6 e U3, possui algumas limitações (Zhang et al., 2017): são constitutivamente ativos e, portanto, ausentes de especificidade celular ou tecidual; são pouco caracterizados para alguns organismos; e não são passíveis de utilização na transcrição *in vitro* devido à falta de versões comerciais de RNA Polimerase III.

De modo a contornar esses obstáculos, uma forma alternativa de conduzir a engenharia multiplex é utilizar promotores para RNA Polimerase II. Esses promotores, a princípio, também possuem algumas limitações no uso para engenharia de genoma (Zhang et al., 2017), como: os transcritos por eles dirigidos na expressão passam por processamento pós-transcricional (capeamento, *splicing* e poliadenilação), cujas modificações no transcrito podem inativar funcionalmente o gRNA; o RNA maduro é transportado ao citosol, ficando confinado em um compartimento celular distinto de onde deve atuar juntamente com a nuclease Cas, o que constitui uma barreira física que bloqueia a atividade de edição. Entretanto, quando esses promotores são associados a uma segunda estratégia de processamento, eles podem se tornar suficientes para promover a edição. Essa estratégia, denominada "RGR" (*Ribozyme-gRNA-Ribozyme*), tem como princípio o uso de gRNAs flanqueados por ribozimas em ambas as extremidades 5' e 3' das guias (Lee et al., 2016), sendo geralmente as ribozimas *hammerhead* (HH) e do *Hepatitis D Virus* (HDV).

Com base nisso, o princípio desse mecanismo é similar ao princípio da estratégia RNAt-gRNA, embora se fundamente na atividade autocatalítica das ribozimas a partir do momento em que são transcritas, o que promove igualmente a liberação dos gRNAs no núcleo celular, contornando os processamentos que um RNA maduro receberia, como mencionado acima. Desse modo, torna-se possível produzir gRNAs a partir de quaisquer promotores em quaisquer organismos, tornando as escolhas de promotores para conduzir a edição de genoma virtualmente ilimitadas (He et al., 2017). Embora essa estratégia também possa ser utilizada para um único gRNA com expressão dirigida por promotores para RNA Polimerase II, por meio de cassetes monocistrônicos similares aos mencionados sobre promotores para RNA Polimerase III, ela é bastante conveniente para a abordagem multiplex, como mencionado anteriormente, em uma configuração policistrônica.

Identificação de *off-targets* potenciais

Embora ambas as ferramentas apresentadas no tópico anterior sejam capazes de prever potenciais sítios *off-target* para cada gRNA desenhado, há ferramentas especializadas nessa tarefa e com maior acurácia. Entre elas, a ferramenta Cas-OFFinder⁶ (Bae et al., 2014) é bastante completa, intuitiva, com layout simples e mais utilizada, sendo, portanto, recomendada pelos autores para esse procedimento, o qual é detalhado a seguir:

- 1) Acessando o Cas-OFFinder (Figura 11), deve ser realizado o *input* simultâneo das sequências de gRNA selecionadas na etapa anterior. Nessa janela,

⁶ Disponível em: <http://www.rgenome.net/cas-offinder/>

a ferramenta é subdividida em 3 seções: (a) *PAM Type*: primeiramente, deve ser selecionada a sequência PAM correspondente à nuclease com a qual se irá trabalhar; (b) *Target Genome*: aqui deve ser selecionado o grupo de organismos dentro do qual a espécie que contém o genoma-alvo está localizada, sendo os grupos *Vertebrates*, *Insects*, *Plants*, *Others* e *Non-reference Vertebrates*, e depois deve ser selecionada a espécie com seu genoma de referência; (c) *Query Sequences*: nesta seção, as sequências de resíduos dos gRNAs são inseridas uma abaixo da outra (em linhas separadas), sem a sequência PAM, podendo se estabelecer um limite máximo de *mismatches*, dando maior objetividade à análise. Ao final, clicar em *Submit* para iniciar a busca.

- 2) Em seguida, novas informações serão mostradas em nova página (Figura 12). Nela são apresentadas duas seções: (a) *Summary*: apresenta, para cada sequência de gRNA, uma linha com o número de casos *on-target* e outra linha com o número de casos *off-target*, de acordo com o número de *mismatches*, de forma similar ao que é demonstrado no CRISPOR; (b) *Details*: tanto para os casos de *on-target* quanto de *off-target*, é apresentado o pareamento entre o gRNA e o alvo genômico, bem como indicado em qual cromossomo o alvo está localizado, além de sua posição de início em termos de resíduo nucleotídico e da orientação da sequência na fita (sense +, ou antissense -). Com essa informação, é possível voltar ao Phytozome e localizar a sequência do sítio *off-target*, caracterizando-a quanto à expressão. Nesse caso, é recomendado verificar primeiro se os sítios *off-target* se encontram em sequências intergênicas ou transcritas. Estando em sequências transcritas, é importante determinar a região da sequência, avaliando sua importância no desempenho do produto final, de modo que, uma vez realizado o *knockout*, o resultado não gere efeitos prejudiciais ao organismo ou que interfiram na análise. Por exemplo: clivagens em regiões 5'-UTR e éxons costumam ser críticas; portanto, um gRNA desenhado com um *off-target* potencial nessas clivagens não deve ser priorizado, ao contrário de outros localizados em regiões intrônicas (exceto em regiões de consenso que sejam preponderantes para o *splicing*).

Simulação de *knockout* e predição de homologia

Uma vez verificados e filtrados apenas gRNAs com boa especificidade e sem atividade *off-target* crítica, quando houver sítios *off-target* potenciais, é importante simular o *knockout* resultante da edição, para que possa ser feita a predição de homologia com domínios proteicos funcionais a partir da cadeia polipeptídica remanescente, uma vez que se espera que ela seja truncada e sua atividade biológica seja abolida. É importante ressaltar que esse tipo de análise deve ser feita apenas

Cas-OFFinder

A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases.

PAM Type

CRISPR/Cas-derived RNA-guided Endonucleases (RGENs)

- SpCas9 from *Streptococcus pyogenes*: 5'-NGG-3'
- SpCas9 from *Streptococcus pyogenes*: 5'-NRG-3' (R = A or G)
- StCas9 from *Streptococcus thermophilus*: 5'-NNAGAAW-3' (W = A or T)
- NmCas9 from *Neisseria meningitidis*: 5'-NNNNGMTT-3' (M = A or C)
- SaCas9 from *Staphylococcus aureus*: 5'-NNGRR-3' (R = A or G)
- CJCas9 from *Campylobacter jejuni*: 5'-NNNRYAC-3' (V = G or C or A, R = A or G, Y = C or T)
- CJCas9 from *Campylobacter jejuni*: 5'-NNNNRYAC-3' (R=A or G, Y= C or T)
- AsCpf1 from *Acidaminococcus* or LbCpf1 from *Lachnospiraceae*: 5'-TTTN-3'
- AsCpf1 from *Acidaminococcus* or LbCpf1 from *Lachnospiraceae*: 5'-TTTV-3' (V = G or C or A)
- SpCas9 from *Streptococcus pasteurianus*: 5'-NNGTGA-3'
- FnCpf1 from *Francisella*: 5'-TTN-3'
- SaCas9 from *Staphylococcus aureus*: 5'-NNNRT-3' (R=A or G)
- FnCpf1 from *Francisella*: 5'-KYTV-3'
- VRER SpCas9 from *Streptococcus pyogenes*: 5'-NGCG-3'
- VQR SpCas9 from *Streptococcus pyogenes*: 5'-NGA-3'
- XCas9 3.7 (TLIKDIV SpCas9) from *Streptococcus pyogenes*: 5'-NGT-3'
- XCas9 3.7 (TLIKDIV SpCas9) from *Streptococcus pyogenes*: 5'-NG-3'
- BHCas12b from *Bacillus hisashii*: 5'-TTN-3'
- BHCas12b v4 from *Bacillus hisashii*: 5'-ATTN-3'
- BHCas12b v4 from *Bacillus hisashii*: 5'-DTTN-3'
- Spy-macCas9 from *Streptococcus pyogenes* and *Streptococcus macacae*: 5'-NAAN-3'
- Nme2Cas9 from *Neisseria meningitidis*: 5'-NNNNCC-3'
- RR AsCpf1 from *Acidaminococcus*: 5'-TYCV-3'
- RVR AsCpf1 from *Acidaminococcus*: 5'-TATV-3'
- CcCas9 from *Clostridium cellulolyticum*: 5'-NNNNGNA-3'
- MAD7 nuclease: 5'-YTTV-3'

Query Sequences

Query sequences (5' to 3'), one sequence per line.

Please write crRNA sequences **without PAM sequences** (e.g. without NGG for SpCas9). The length of each query sequence should be between **15 and 25 nt**, and **all be the same length!**

```
CAGCAACTCCAGGGGCCGC
AAAGGAAACCATTGTGTAA
```

Mismatch Number (eq or less than)

DNA Bulge Size (eq or less than)

RNA Bulge Size (eq or less than)

Please note that large number of bulge size will significantly increase the calculation time!

Mixed bases are allowed.

The count of query sequence must be less than 1000.

<DNA bulge>

<RNA bulge>

Target Genome

Organism Type

Plants

Genomes

- Arabidopsis thaliana* (TAIR10) - Thale cress
- Oryza sativa* (OSv4) - Rice
- Solanum lycopersicum* (SL2.4) - Tomato
- Zea mays* (AGPv3) - Corn
- Chlamydomonas reinhardtii* (Chlr4)
- Solanum tuberosum* (PGSC v4.03) - Potato
- Glycine max* (v1.0) - Soybean
- Vitis vinifera* (IGGP_12X/Ensplant26) - European grapevine
- Manihot esculenta* (JGI 4.1) - Cassava
- Malus domestica* (JGI 1.0) - Apple
- Hordeum vulgare* (Ensembl Plants 28) - Barley
- Nicotiana benthamiana* (v1.0.1)
- Fragaria vesca* (1.0) - Wild strawberry
- Citrus sinensis* (1.0) - Sweet orange
- Theobroma cacao* (CIRAD 1.0) - Cacao
- Theobroma cacao* (CGD 1.1) - Cacao

Figura 11. Busca por potenciais sítios *off-target* no Cas-OFFinder. Janela de *input* dos gRNAs selecionados na etapa anterior, possibilitando a mudança de diferentes parâmetros em cada seção. *PAM Type*: seleção da sequência PAM associada à nuclease em experimentação; *Target Genome*: seleção da espécie que contém o genoma-alvo a ser analisado para as potenciais regiões não alvo, com as espécies segmentadas em grupos de organismos que também precisam ser selecionados; *Query Sequences*: inserção das sequências de gRNAs simultaneamente, atendendo às exigências dispostas nesta seção, também sendo possível estabelecer limite de número de *mismatches*.

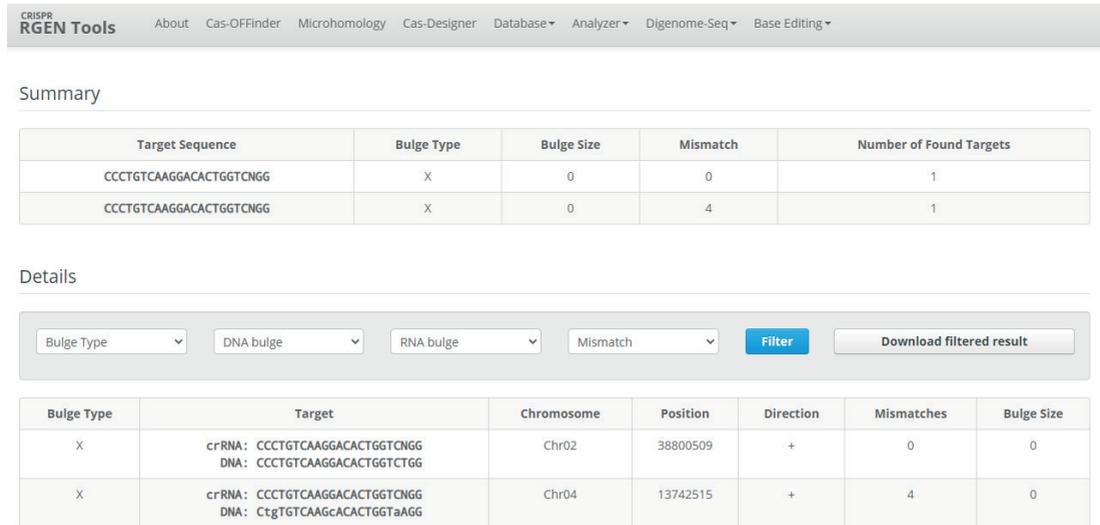


Figura 12. Resultado da predição de potenciais sítios *off-target* no Cas-OFFinder. Janela de *output* dos gRNAs selecionados na etapa anterior, com a análise dos casos de sítios potenciais de *off-targets* encontrados no genoma de referência. A seção *Summary* resume, para o mesmo gRNA, casos *on-target* em uma linha (mostrado por 0 *mismatch*, sendo a própria sequência-alvo no genoma) e *off-target* em outra linha (mostrado por 1 caso com 4 *mismatches*, neste exemplo). Na seção *Details*, ambos os casos são detalhados quanto ao pareamento entre o gRNA e a sequência genômica similar, porém informando a localização cromossômica, nas colunas *Chromosome* e *Position*.

quando o objetivo almejado é o *knockout* por edição, via *indels*, das sequências codantes de proteínas-alvo. Para isso, o procedimento adequado deve ser realizado da seguinte maneira:

- 1) Deve ser conduzida uma tradução *in silico* a partir do resíduo nucleotídico onde teoricamente ocorrerá a clivagem pela nuclease, de modo que seja gerada uma cadeia polipeptídica truncada para posterior verificação de homologia. É possível conduzi-la na ferramenta online ExPASy Translate Tool⁷. A partir do ponto de clivagem, selecionar a sequência *upstream* e inseri-la na barra de sequência nucleotídica (Figura 13). Recomenda-se, como formato de *output*, manter a opção *Compact* selecionada, bem como os resultados mostrados em ambas as fitas, habilitando as duas opções, sem necessidade de selecionar o código genético específico. Por fim, deve-se clicar em *Translate!*. É importante lembrar que nem todo *indel* ocasionará a formação de um *stop codon* mais precoce ao longo da sequência-alvo ou, de modo mais amplo, uma *frameshift* que altere a cadeia polipeptídica a 3' do ponto de clivagem, modificando sua estrutura tridimensional secundária e/ou terciária e, com isso, inviabilizando

⁷ Disponível em: <https://web.expasy.org/translate/>

a atividade biológica nativa. Isso porque é possível que ocorram inserções ou deleções de resíduos nucleotídicos em números múltiplos de 3, o que mantém o restante da sequência *in frame*, não necessariamente provocando este efeito deletério pela presença ou ausência de alguns poucos resíduos de aminoácido. Porém, devido à imprevisibilidade do tipo de *indel* a ocorrer em cada evento, deve-se trabalhar assumindo a ocorrência dos *indels* não múltiplos de 3, possibilitando a simulação.

Translate is a tool which allows the translation of a nucleotide (DNA/RNA) sequence to a protein sequence.

DNA or RNA sequence

```
ATGGCCACTGTC AAGAAAATTGGAAAGAAGGCATGTGTGATCGGTGGAAGTGGATTTATGGCCCTCTGCTTGGTAAAGC
AATTGCTTCA CAAGGGTTATGCTGTCAACCACTGTTAGAGACCCAGATAATGCTAAGAAAATATCTCACCTTTTGGC
CCTGCAAAAGTTGGGCGACTTGAACATATTTGGAGCAGATCTAACAGGTGAAAAGGGATTTTGTGCCCTATAGCAGGC
TGTGAACTTGTCTTCA GCTTGTACACCTGTGAACCTTTCCTCTGAAAGATCCTGAGAAATGACACGATCAAGCCTGC AA
TCTCAGGAGTGTGAAATGTGTTGAAAGCATGTGCACGAGCAA
```

Output format

- Verbose: Met, Stop, spaces between residues
- Compact: M, -, no spaces
- Includes nucleotide sequence
- Includes nucleotide sequence, no spaces

DNA strands

- forward reverse

Genetic codes - See NCBI's genetic codes

Standard

reset TRANSLATE!

Figura 13. Tradução in silico de sequências truncadas no ExPASy Translate Tool. Janela de *input* da sequência nucleotídica simulada quanto à truncagem, evidenciando diferentes formatos possíveis de *output* e a possibilidade de representação de 3 *frames* (apenas *Forward* ou *Reverse*) ou 6 *frames* (*Forward* e *Reverse*).

- 2) Com o resultado da análise, uma nova página é aberta (Figura 14), com os resultados da tradução in silico, em todas as 6 possíveis ORFs (*Open Reading Frames*). Como a sequência do *input* geralmente é inserida a partir do primeiro resíduo do *start codon* (ATG), todo resultado desse modo será encontrado na *Frame 1* (5'-3'), com a sequência truncada marcada continuamente em vermelho.

Results of translation

- Open reading frames are highlighted in red
- Select your initiator on one of the following frames to retrieve your amino acid sequence

[Download all the translated frames](#)

5'3' Frame 1

MAIVKIKKACVIGSGFMASCLVQQLLHKGYAVNTTVRDPDPAKIKSHLLALQSLGDLNIFGADLTGERDFDAPDIAGCELVQQLATPVNFASEDPENDTIKPAI
SGVLNVLRACARA

5'3' Frame 2

WPLSRKLERRHV-SVEVDLWPLAW-SNCFTRVMLSTPILLETQIMLRKYLTFWPCVKVWAT-TYLEQI-QVKGILMPL-QAVNLSFSLHLH-TLLKILRMRSSLQS
QEC-MC-KHVHEQ

5'3' Frame 3

GHCQENWKEGMDRQWKVIYGLLKGAIASQGLCCQHHH-C-RPR-C-ENISPFPGPAKFRLEHIWRSRNR-KGF-CPYSRL-TCLSACYTCELCF-RS-E-HDQACNL
RSVECVESMCTS

3'5' Frame 1

LLVHMLSTHSTLLRLQA-SCHSQDLQKSSQV-QAERQVHSLH-GHQNPFHLLDLLQICSSRPNFAGPKGEIFS-HYLGL-QWC-QHNPCETALPSKRP-IHFHR
SHMPSFQFS-QWP

3'5' Frame 2

CSCTCFQHIQHS-DCRLDRVILRIFRSKVHRCCLKDKFTACYRGIKIPFTC-ICSKYVQVAQTLQGQKRVYFLSIIWVSNQVDSITLVKQLLYQARGHKSTSTD
HTCLLSNFLDSGH

3'5' Frame 3

ARAHAFNIFNTPEIAGLIVSFGSGSEAKFTGVAS-KTSSQPAIGASKSLSPVRSAPNMFKSPKLCRAKR-DIFLALSGSLTVVLA-PL-SNCFTRQEAINPLPPI
THAFFPIFLTVA

Figura 14. Resultado da tradução *in silico* no ExPASy Translate Tool. Janela de *output* da sequência traduzida, em 6 diferentes ORFs (*Open Reading Frames*), com a sequência da cadeia polipeptídica, neste caso, presente na *Frame 1 5'-3'*.

- Tendo a sequência polipeptídica truncada, é possível copiá-la e inseri-la na ferramenta on-line para predição de homologia com domínios proteicos funcionais, InterPro⁸. Nessa página (Figura 15), realiza-se o *input* da sequência no formato FASTA, conduzindo a busca e mantendo os parâmetros pré-estabelecidos. Logo após, uma página com o *output* da análise (Figura 16) será aberta, informando sobre a presença ou ausência de predição de famílias proteicas referentes à sequência polipeptídica remanescente. Não havendo predição, é provável que a sequência resultante não forme domínios funcionais que desempenhem atividade biológica residual, o que implica na deleção da função da proteína associada ao gene submetido ao *knockout*.

Uma vez confirmada ausência de homologia, o gRNA atende a todos os requisitos e propriedades de um candidato útil e com maior probabilidade de exercer sua função sem que efeitos indesejados ocorram, podendo ser validado *in vitro* e preparado para subclonagem em vetores de expressão (abordagem: uso de plasmídeos), ou então usado isoladamente (abordagem: IVT) ou complexado a uma nuclease (abordagem: RNP).

⁸ Disponível em: <http://www.ebi.ac.uk/interpro/>

InterPro Classification of protein families

Home Search Browse Results Release notes Download Help About

Search by sequence Search by text Search by Domain Architecture

Sequence, in FASTA format

> Protein 1
 MATVKKIKGKACVIGGSGFMASCLVQLLHKGYAVNTTVRDPNNAKTIHLLALQSLGDLNIFGADLTGERDFDAPAIAGCELVFQLATPVNFASE

Valid Sequence.

Choose file Example protein sequence

Figura 15. Busca por homologia proteica no InterPro. Janela de *input* da sequência polipeptídica truncada para predição de homologia com domínios proteicos funcionais por associação a famílias proteicas.



Figura 16. Resultado da predição de homologia com domínios funcionais no InterPro. Janela de *output* da sequência polipeptídica truncada, evidenciando ausência de homologia com famílias proteicas, indicando que a sequência remanescente não desempenha atividade biológica conhecida.

Parte 2: in vitro

O segundo passo para realizar a edição genômica via CRISPR/Cas consiste na validação *in vitro* das moléculas isoladas e mediadoras do processo de edição, o que pode ser conduzido com utilização de kits comerciais. Devido ao fato de que os elementos genéticos que compõem um vetor de expressão (nas abordagens de

uso de plasmídeos) e de que os métodos de transfecção variam bastante de acordo com a espécie-alvo e o explante utilizado, não será detalhado aqui o delineamento dessas etapas, uma vez que isso deve ser estabelecido de forma particular em cada projeto de pesquisa. Entretanto, a validação *in vitro* dos guias desenhados e das nucleases associadas é um processo mais universal para os diferentes organismos que se pretende editar.

Eficiência de clivagem e validação do complexo nuclease-sgRNA

É fundamental que o sgRNA seja capaz de guiar a nuclease ao sítio-alvo. Mais ainda, o processo somente será viável se, além dessa capacidade, a nuclease for eficiente em promover a clivagem da sequência no sítio de clivagem esperado. Dessa forma, é possível verificar a atividade de ambos por meio de um teste de eficiência de clivagem *in vitro*, no qual a sequência-alvo é amplificada por Reação em Cadeia da Polimerase (PCR, do inglês *Polymerase Chain Reaction*) e o produto da PCR é clivado pelo complexo. A formação de fragmentos de digestão a partir do produto de PCR indica que tanto o sgRNA quanto a nuclease executam sua função normalmente, estando aptos a proceder à edição genômica *in vivo*. Caso contrário, a ausência de produtos de clivagem indica ausência de eficiência da função de pelo menos um desses componentes do complexo. Para isso, é necessário obter a RNP correspondente ao complexo que será formado *in vivo*. Nesse caso, a obtenção da nuclease e do sgRNA pode ocorrer conforme descrito no tópico RNPs: sistemas DNA-free, deste capítulo. Para trabalhos que utilizem o sistema de RNP para transfecção, após a validação o complexo já pode ser diretamente utilizado para proceder à edição. No caso de sistemas plasmidiais ou de transcritos, procede-se às sequências nucleotídicas correspondentes das moléculas validadas.

Um procedimento simples e sucinto (GenScript, 2018) para essa validação, baseado em kit de ensaio *in vitro*, pode ser seguido conforme descrito abaixo:

- 1) Realizar a amplificação da região-alvo por PCR (Nota: é possível desenhar pares de *primers* tanto para a região-alvo quanto para potenciais *off-targets* no CRISPOR, a partir da tabela de informações do *output* da busca). O produto de PCR será utilizado como substrato para a digestão com a nuclease da RNP. É recomendado que o produto contenha pelo menos 200 pb flanqueando a sequência com a qual o gRNA se pareia, tanto a 5' quanto a 3', sendo preferencial que o produto de PCR tenha cerca de 1 kb para melhor visualização da digestão.

- 2) Preparar 16 μL de RNP em um tubo nuclease-free, conforme o seguinte sistema de reação, adaptado de GenScript (2018):

Componente	Volume/Quantidade de matéria
sgRNA	3,6 pmol
Nuclease Cas (1 $\mu\text{g}/\mu\text{L}$)	1,5 pmol
Tampão de reação (Cas) (10X)	2 MI
H ₂ O nuclease-free	13,6 μL
Volume final	16 μL

- 3) Conduzir o *assembly* da RNP incubando o tubo do sistema de reação a 37 °C por 10 minutos.
- 4) Adicionar 450 ng do produto de PCR ao mix de RNP montado no passo 2, completando o volume com água nuclease-free para 20 μL e agitando levemente para promover a mistura.
- 5) Incubar a mistura produto-RNP por, no mínimo, 30 minutos a 37 °C.
- 6) Verificar a ocorrência ou não da digestão por meio de visualização baseada em eletroforese em gel de agarose.

Parte 3: in vivo

Por último, após ter sido realizada a transfecção e regeneração de eventos candidatos a plantas editadas, é necessário caracterizar as edições ocorridas in vivo por meio de técnicas diversas, cada uma apresentando suas vantagens e desvantagens. Aqui não será detalhado um passo a passo de cada uma, mas o leitor será direcionado com literatura a respeito das mesmas, bem como com informações gerais pertinentes, para buscar sobre como conduzi-las conforme sua preferência, adequação ao desenho de seu estudo e disponibilidade de materiais, reagentes e equipamentos em seu laboratório.

Identificação de plantas editadas

Para que os eventos candidatos sejam caracterizados quanto à alteração da sequência genômica, é de grande interesse filtrar, quando possível, as plantas que apresentem fenótipo, reduzindo o volume de avaliações, visando a tornar a caracterização mais prática, simples e menos trabalhosa.

Ao se trabalhar com sistemas plasmidiais e integração genômica, é possível, com base no vetor de expressão utilizado e seus elementos genéticos, selecionar explantes que se regenerem em meio seletivo com resistência garantida pelo marcador com o qual se trabalha, ou que sejam visualmente observáveis quanto à característica que o gene repórter inserido confere. Entretanto, ao se trabalhar com IVTs ou RNPs, por exemplo, a seleção com base nessas características se torna inviável, já que geralmente não se aplica. Dessa forma, caso não seja possível conduzir uma triagem (*screening*) com base em agente seletivo, ou no próprio fenótipo resultante da edição genômica, a caracterização deve se proceder para a seleção de mutantes. Para essa seleção, os seguintes métodos são alguns dos que podem ser utilizados na genotipagem:

- Sequenciamento: bastante útil para detalhar o tipo e a composição do *indel* gerado após a clivagem e reparo, informando sobre cada uma das fitas a sequência exata dos resíduos nucleotídicos, podendo ser utilizados tanto o método de Sanger quanto abordagens de NGS (*Next-Generation Sequencing*). De forma geral, o sequenciamento costuma ser uma abordagem central ou complementar a outros métodos menos informativos em nível de sequência genômica, sendo usado, quando aplicado como método complementar, para caracterizar apenas os eventos previamente genotipados.
- PCR-RE (*PCR-restriction enzyme*): baseada em uma etapa de amplificação das regiões a serem avaliadas, que geralmente são as sequências-alvo e regiões de possível *off-target*. Em seguida, é realizada uma digestão com enzimas de restrição com sítio de reconhecimento englobando o sítio de clivagem pela nuclease da edição. Dessa forma, a digestão do produto de PCR revela que o sítio de restrição permanece intacto e, portanto, não houve edição; a ausência de digestão revela que o sítio de restrição foi eliminado devido ao *indel* resultante da edição de genoma. Embora frequentemente utilizada, a limitação dessa técnica reside na necessidade da existência de tais sítios de restrição, o que nem sempre ocorre, dada a reduzida diversidade de gRNAs candidatos que sejam apropriados. Para mais informações, ler o trabalho de Xie e Yang (2013).
- Ensaio T7E1/Surveyor™: baseado na capacidade de reconhecimento e clivagem de moléculas de DNA heteroduplex (derivadas dos *mismatches* entre as fitas) pelas enzimas T7 endonuclease e Surveyor. Essas enzimas reconhecem moléculas de DNA cujas fitas se encontram mal-pareadas. Por essa razão, são úteis na detecção de heteroduplexes, que correspondem a indivíduos heterozigotos para a edição. Trata-se de técnicas simples e baratas, quando comparadas ao sequenciamento, mas que não revelam a sequência nucleotí-

dica das regiões avaliadas. Também não possuem reprodutibilidade robusta e estão suscetíveis a gerar resultados falso-positivos quando há ocorrência de polimorfismos ou mutações randômicas nas sequências avaliadas, embora sua utilização seja relatada comumente na literatura. Além disso, apenas indivíduos heterozigotos para as edições podem ser detectados. Para mais informações, ler o trabalho de Vouillot et al. (2015).

- qPCR (PCR quantitativa): são inúmeras as variações da qPCR para genotipagem dos *indels* de eventos, tanto em nível genômico quanto em nível transcriptômico (via RT-qPCR), cujos *primers* desenhados se sobrepõem ao sítio-alvo a ser editado, de modo que o fundamento dessa técnica reside no fato de que a ocorrência ou não da mutagênese é indicada pelo insucesso ou sucesso, respectivamente, da amplificação. Isso porque a edição tende a anular o anelamento do *primer* na região editada, interferindo na geração de amplicon. Entretanto, *indels* menores tendem a interferir em menor grau na amplificação. Para isso, a técnica de HRM (*High Resolution Melting*) pode ser uma alternativa para identificar *indels* menores, pois consiste na alteração do padrão de fluorescência na análise da curva de *melting*, que é gerado durante a formação do duplex de DNA. Com isso, com base no perfil da curva, é possível distinguir nitidamente um indivíduo selvagem (*wild-type*, WT) de um mutante. De forma geral, a qPCR/RT-qPCR é uma técnica acurada, porém com custo relativo um pouco maior que algumas outras técnicas de genotipagem. Desse modo, recomendamos que o leitor compreenda os detalhes das principais vertentes metodológicas a partir dos trabalhos de Yu et al. (2014), Falabella et al. (2017) e Li et al. (2018).
- Western blot: técnica que consiste na marcação de uma proteína-alvo utilizando um anticorpo conjugado a uma molécula que produz um sinal visível após reação enzimática, sendo o anticorpo direcionado para reconhecer uma região específica da proteína-alvo. A proteína é detectada sobre uma membrana após a transferência a partir de um gel de SDS-PAGE (eletroforese em gel de poliacrilamida com dodecilsulfato de sódio). Havendo presença da proteína em sua forma completa, a região reconhecida pelo anticorpo permanece presente, resultando em sinal colorimétrico devido à reação catalisada; no caso de ausência da forma completa, nenhum sinal é teoricamente detectado. Trata-se de um método menos utilizado para a caracterização, já que além de ser mais trabalhoso e exigir maior tempo de trabalho, é possível realizar a caracterização em nível genômico ou transcriptômico, sem necessidade de avaliação em nível proteômico. Além disso, há a possibilidade de gerar falso-positivos, uma vez que pode ocorrer reação cruzada de ligação do anticorpo com proteínas estruturalmente similares à proteína-alvo, ainda que

esta já esteja truncada devido à edição. Normalmente, é conduzida a detecção da proteína Cas, verificando eventos que expressem a nuclease e, portanto, estão a princípio passando por um processo de edição (Endo et al., 2019).

- Gel de poliacrilamida (PAGE): o gel de poliacrilamida forma uma malha com propriedades que geram alta capacidade de resolução de moléculas de DNA, o que permite detectar diferenças de até cerca de 2 pb entre as fitas e entre diferentes amostras. Desse modo, é uma alternativa bastante acurada para detectar todas as faixas de comprimento dos *indels*, apesar de ser um pouco trabalhosa e não muito rápida. Para mais informações, conferir o trabalho de Zhu et al. (2014).
- Gel de agarose (AGE) 4%-6%: embora géis de agarose, em suas concentrações usuais de trabalho (1% a 3%) não sejam úteis para distinguir *indels* por CRISPR/Cas e, portanto, para genotipar eventos, Bhattacharya e Van Meir (2019) recentemente desenvolveram um método simples para genotipagem utilizando géis mais concentrados, com 4% a 6% de agarose. Nessa abordagem, *indels* de pelo menos 3 pb podem ser detectados visualmente com base na propriedade de mobilidade de moléculas heteroduplex ao longo da malha do gel. Inicialmente, heterozigotos podem ser diferenciados de homozigotos após serem submetidos a 34 ciclos de PCR. Em seguida, os homozigotos mutantes podem ser diferenciados dos homozigotos selvagens após um ciclo adicional de desnaturação/renaturação por meio da hibridização com um controle selvagem, o que leva à formação de homoduplexes no caso de homozigotos selvagens e de heteroduplexes no caso de homozigotos mutantes. A técnica é bastante vantajosa quanto à praticidade e simplicidade frente a outras como qPCR, PAGE, Western blot e ensaios T7E1/Surveyor, mas até o momento foi utilizada apenas para animais, sendo promissora para testar sua aplicação em plantas.

Referências

ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Biologia molecular da célula**. 5. ed. Porto Alegre: Artmed, 2010. 1396 p.

AMIRKHANOV, R. N.; STEPANOV, G. A. Systems of delivery of CRISPR/Cas9 ribonucleoprotein complexes for genome editing. **Russian Journal of Bioorganic Chemistry**, v. 45, p. 431–437, Nov. 2019.

ANDERSSON, M.; TURESSON, H.; OLSSON, N.; FALT, A-S.; OHLSSON, P.; GONZALEZ, M. N.; SAMUELSSON, M.; HOFVANDER, P. Genome editing in potato via CRISPR-Cas9 ribonucleoprotein delivery. **Physiologia Plantarum**, v. 164, n. 4, p. 378–384, Dec. 2018.

- ANDRES, S. N.; VERGNES, A.; RISTIC, D.; WYMAN, C.; MODESTI, M.; JUNOP, M. A human XRCC4-XLF complex bridges DNA. **Nucleic Acids Research**, v. 40, n. 4, p. 1868–1878, Feb. 2012.
- BAE, S.; PARK, J.; KIM, J. S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. **Bioinformatics**, v. 30, n. 10, p. 1473-1475, May 2014.
- BHATTACHARYA, D.; VAN MEIR, E. G. A simple genotyping method to detect small CRISPR-Cas9 induced indels by agarose gel electrophoresis. **Scientific Reports**, v. 9, article 4437, 2019.
- CAI, Y.; CHEN, L.; LIU, X.; SUN, S.; WU, C.; JIANG, B.; HAN, T.; HOU, W. CRISPR/Cas9- mediated genome editing in soybean hairy roots. **PLoS One**, v. 10, n. 8, e0136064, Aug. 2015.
- CARDONA, Y. T.; MORALES, M. A. M. Main repair mechanisms for damages in the DNA molecule. **Biosalud**, v. 13, n. 2, p. 95-110, jul./dic. 2014.
- CARROLL, D. Genome engineering with zinc-finger nucleases. **Genetics**, v. 188, n. 4, p. 773-782, Aug. 2011.
- CICCIA, A.; ELLEDGE, S. J. The DNA damage response: making it safe to play with knives. **Molecular Cell**, v. 40, n. 2, p. 179-204, Oct. 2010.
- CONCORDET, J.; HAEUSSLER, M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. **Nucleic Acids Research**, v. 46, n. 1, p. 242-245, July 2018.
- 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**, v. 339, n. 6121, p. 819-823, Feb. 2013.
- COOK, E. H.; SCHERER, S. W. Copy-number variations associated with neuropsychiatric conditions. **Nature**, v. 455, n. 7215, p. 919-923, Oct. 2008.
- CORNU, T. I.; THIBODEAU-BEGANNY, S.; GUHL, E.; ALWIN, S.; EICHTINGER, M.; JOUNG, J. K.; CATHOMEN, T. DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. **Molecular Therapy**, v. 16, n. 2, p. 352-358, Feb. 2008.
- CUI, Y.; DONG, H.; MA, Y.; ZHANG, D. Strategies for applying nonhomologous end joining-mediated genome editing in Prokaryotes. **ACS Synthetic Biology**, v. 8, n. 10, p. 2194–2202, 2019.
- CYRANOSKI, D. **CRISPR tweak may help gene-edited crops bypass biosafety regulation**. 2015. Disponível em: <https://www.nature.com/news/crispr-tweak-may-help-gene-edited-crops-bypass-biosafety-regulation-1.18590>. Acesso em: 02 jun. 2020.
- DAVIS, A. J.; CHEN, B. P.; CHEN, D. J. DNA-PK: a dynamic enzyme in a versatile DSB repair pathway. **DNA Repair (Amst)**, v. 17, p. 21–29, May 2014.
- DEMIRER, G. S.; ZHANG, H.; GOH, N. S.; GONZÁLEZ-GRANDÍO, E.; LANDRY, M. P. Carbon nanotube-mediated DNA delivery without transgene integration in intact plants. **Nature Protocols**, v. 14, p. 2954-2971, Oct. 2019.
- DERIANO, L.; ROTH, D. B. Modernizing the non-homologous end-joining repertoire: alternative and classical NHEJ share the stage. **Annual Review of Genetics**, v. 47, p. 433-455, 2013.
- DOWNS, J. A.; JACKSON, S. P. A means to a DNA end: the many roles of Ku. **Nature Reviews Molecular and Cell Biology**, v. 5, p. 367-378, May 2004.

DUAN, Y. B.; LI, J.; QIN, R. Y.; XU, R. F.; LI, H.; YANG, Y. C.; MA, H.; LI, L.; WEI, P. C.; YANG, J. B. Identification of a regulatory element responsible for salt induction of rice OsRAV2 through *ex situ* and *in situ* promoter analysis. **Plant Molecular Biology**, v. 90, n. 1/2, p. 49-62, Jan. 2016.

ENDO, M.; MIKAMI, M.; ENDO, A.; KAYA, H.; ITOH, T.; NISHIMASU, H.; NUREKI, O.; TOKI, S. Genome editing in plants by engineered CRISPR–Cas9 recognizing NG PAM. **Nature Plants**, v. 5, p. 14–17, Jan. 2019.

FALABELLA, M.; SUN, L.; BARR, J.; PENA, A. Z.; KERSHAW, E. E.; GINGRAS, S.; GONCHAROVA, E. A.; KAUFMAN, B. A. Single-step qPCR and dPCR detection of diverse CRISPR-Cas9 gene editing events *in vivo*. **G3: Genes, Genomes, Genetics**, v. 7, n. 10, p. 3533-3542, Oct. 2017.

GAJ, T.; SIRK, S. J.; SHUI, S. L.; LIU, J. Genome-editing technologies: principles and applications. **Cold Spring Harb Perspectives in Biology**, v. 8, n. 12, p. a023754, Dec. 2016.

GENSCRIPT. **CRISPR: Ribonucleoprotein (RNP) user manual**. New Jersey, 2018. p. 7.

GIL-HUMANES, J.; WANG, Y.; LIANG, Z.; SHAN, Q.; OZUNA, C. V.; SÁNCHEZ-LEÓN, S.; BALTES, N. J.; STARKER, C.; BARRO, F.; GAO, C.; VOYTAS, D. F. High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. **The Plant Journal**, v. 89, n. 6, p. 1251-1262, Mar. 2017.

GLASS, Z.; LEE, M.; LI, Y.; XU, Q. Engineering the delivery system for CRISPR-based genome editing. **Trends in Biotechnology**, v. 36, n. 2, p. 173-185, Feb. 2018.

GRIFFITHS, A. J. F.; MILLER, J. H.; SUZUKI, D. T.; LEWONTIN, R. C.; GELBART, W. M. **An introduction to genetic analysis**. 7th KREBS, J. E; GOLDSTEIN, E. S.; KILPATRICK, S. T; LEWIN, B. **Lewin's Genes X**. 10

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**, v. 7, n. 9, p. 1494-1496, Sept. 2014.

LI, L.; HE, Z.; WEI, X.; GAO, G.; WEI, Y. Challenges in CRISPR/Cas9 delivery: potential roles of nonviral vectors. **Human Gene Therapy**, v. 26, n. 7, p. 452-462, July 2015.

LI, Q.; SAPKOTA, M.; VAN DER KNAAP, E. Perspectives of CRISPR/Cas-mediated *cis*-engineering in horticulture: unlocking the neglected potential for crop improvement. **Horticulture Research**, v. 7, n. 36, Mar. 2020.

LI, S.; LIU, S.; LIU, Y.; LU, H.; TAN, Y.; HUANG, J.; WEI, P.; SHU, Q. Y. HRM-facilitated rapid identification and genotyping of mutations induced by CRISPR/Cas9 mutagenesis in rice. **Crop Breeding and Applied Biotechnology**, v. 18, n. 2, p. 184-191, Apr./June 2018.

LIANG, Z.; CHEN, K.; LI, T.; ZHANG, Y.; WANG, Y.; ZHAO, Q.; LIU, J.; ZHANG, H.; LIU, C.; RAN, Y.; GAO, C. Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. **Nature Communications**, v. 8, article 14261, Jan. 2017.

LIANG, Z.; CHEN, K.; ZHANG, Y.; LIU, J.; YIN, K.; QIU, J. L.; GAO, C. Genome editing of bread wheat using biolistic delivery of CRISPR/Cas9 *in vitro* transcripts or ribonucleoproteins. **Nature Protocols**, v. 13, p. 413-430, Mar. 2018.

LIANG, Z.; ZHANG, K.; CHEN, K.; GAO, C. Targeted mutagenesis in *Zea mays* using TALENs and the CRISPR/Cas system. **Journal of Genetics and Genomics**, v. 41, n. 2, p. 63–68, Feb. 2014.

LIEBER, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. **Annual Review of Biochemistry**, v. 79, p. 181–211, 2010.

- LIU, W.; RUDIS, M. R.; CHEPLICK, M. H.; MILLWOOD, R. J.; YANG, J. P.; ONDZIGHI-ASSOUME, C. A.; MONTGOMERY, G. A.; BURRIS, K. P.; MAZAREI, M.; CHESNUT, J. D.; STEWART, C. N. Lipofection-mediated genome editing using DNA-free delivery of the Cas9/gRNA ribonucleoprotein into plant cells. **Plant Cell Reports**, v. 39, p. 245-257, 2020.
- MA, X.; ZHU Q.; CHEN Y.; LIU Y-G. CRISPR/Cas9 platforms for genome editing in plants: developments and applications. **Molecular Plant**, v. 9, n. 7, p. 961-974, July 2016.
- MALNOY, M.; VIOLA, R.; JUNG, M. H.; KOO, O. J.; KIM, S.; KIM, J. S.; VELASCO, R.; KANCHISWAMY, C. N. DNA-free genetically edited grapevine and apple protoplast using CRISPR/Cas9 ribonucleoproteins. **Frontiers in Plant Science**, v. 7, n. 1904, 2016.
- MALZAHN, A.; LOWDER, L.; QI, Y. Plant genome editing with TALEN and CRISPR. **Cell Bioscience**, v. 7, n. 21, Apr. 2017.
- MAO, C. Z.; ZHENG, L.; ZHOU, Y. M.; WU, H. Y.; XIA, J. B.; LIANG, C. Q.; GUO, X. F.; PENG, W. T.; ZHAO, H.; CAI, W. B.; KIM, S. K.; PARK, K. S.; CAI, D. Q.; QI, X. F. CRISPR/Cas9-mediated efficient and precise targeted integration of donor DNA harboring double cleavage sites in *Xenopus tropicalis*. **The FASEB Journal**, v. Online, Jun. 2018.
- MAO, Z.; BOZZELLA, M.; SELUANOV, A.; GORBUNOVA, V. Comparison of nonhomologous end joining and homologous recombination in human cells. **DNA Repair (Amst)**, v. 7, n. 10, p. 1765–1771, Oct. 2008.
- MONTGOMERY, S. B.; GOODE, D. L.; KVIKSTAD, E. et al. The origin, evolution, and functional impact of short insertion-deletion variants identified in 179 human genomes. **Genome Research**, v. 23, n. 5, p. 749–761, May 2013.
- MOSCOU, M. J.; BOGDANOVA, A. J. A simple cipher governs DNA recognition by TAL effectors. **Science**, v. 326, n. 5959, p. 1501, Dec. 2009.
- MUSHTAQ, M.; BHAT J. A.; MIR, Z. A.; SAKINA, A.; ALI, S.; SINGH, A. K.; TYAGI, A.; SALGOTRA, R. K.; DAR, A. A.; BHAT, R. CRISPR/Cas approach: a new way of looking at plant-abiotic interactions. **Journal of Plant Physiology**, v. 224–225, p. 156-162, May/June 2018.
- MUSSOLINO, C.; ALZUBI, J.; FINE, E. J.; MORBITZER, R.; CRADICK, T. J.; LAHAYE, T.; BAO, G.; CATHOMEN, T. TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. **Nucleic Acids Research**, v. 42, n. 10, p. 6762–6773, June 2014.
- NEMUDRYI, A. A.; VALETIDINOVA, K. R.; MEDVEDEV, S. P.; ZAKIAN, S. M. TALEN and CRISPR/Cas genome editing systems: tools of discovery. **Acta Naturae**, v. 6, n. 3, p. 19-40, July 2014.
- NUSSBAUM, R. L.; MCINNES, R. R.; WILLARD, H. F. **Genética médica**. 7. ed. Rio de Janeiro: Guanabara Koogan, 2008. 640 p.
- OTTO, S. P. The evolutionary consequences of polyploidy. **Cell**, v. 131, n. 3, p. 452-462, Nov. 2007.
- PANDIARAJAN, R.; GROVER, A. *In vivo* promoter engineering in plants: are we ready? **Plant Science**, v. 277, p. 132-138, Dec. 2018.
- PÂQUES, F.; DUCHATEAU, P. Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. **Current Gene Therapy**, v. 7, n. 1, p. 49–66, Feb. 2007.

PENG, A.; CHEN, S.; LEI, T.; XU, L.; HE, Y.; WU, L.; YAO, L.; ZOU, X. Engineering canker-resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene *CsLOB1* promoter in citrus. **Plant Biotechnology Journal**, v. 15, n. 12, p. 1509–1519, Dec. 2017.

RAMIREZ, C. L.; FOLEY, J. E.; WRIGHT, D. A.; MULLER-LERCH, F.; RAHMAN, S. H.; CORNU, T. I.; WINFREY, R. J.; SANDER, J. D.; FU, F.; TOWNSEND, J. A.; CATHOMEN, T.; VOYTAS, D. F.; JOUNG, J. K. Unexpected failure rates for modular assembly of engineered zinc fingers. **Nature Methods**, v. 5, n. 5, p. 374–375, May 2008.

RICHES, L. C.; LYNCH, A. M.; GOODERHAM, N. J. Early events in the mammalian response to DNA double-strand breaks. **Mutagenesis**, v. 23, n. 5, p. 331–339, Sept. 2008.

ROUET, R.; THUMA, B. A.; ROY, M. D. et al. Receptor-mediated delivery of CRISPR-Cas9 endonuclease for cell-type-specific gene editing. **Journal of the American Chemical Society**, v. 140, n. 21, p. 6596–6603, 2018.

SEHN, J. K. Insertions and deletions (Indels). In: KULKARNI, S.; PFEIFER, J. (Ed.). **Clinical genomics**. London: Academic Press, 2015. p. 129–150.

SILVA, E.; IDEKER, T. Transcriptional responses to DNA damage. **DNA Repair (Amst)**, v. 79, p. 40–49, July 2019.

SPENCER, N. Y. **Are you still using CRISPR plasmids? Consider ribonucleoprotein (RNP) instead!** IDT. Disponível em: [https://www.idtdna.com/pages/education/decoded/article/are-you-still-using-crispr-plasmids-consider-ribonucleoprotein-\(rnp\)-instead!](https://www.idtdna.com/pages/education/decoded/article/are-you-still-using-crispr-plasmids-consider-ribonucleoprotein-(rnp)-instead!). Acesso em: 02 jun. 2020.

STEINERT, J.; SCHIML, S.; PUCHTA, H. Homology-based double-strand break-induced genome engineering in plants. **Plant Cell Reports**, v. 35, n. 7, p. 1429–1438, July 2016.

STRAUSS, B. S. Mechanisms of mutation. In: WEISS, R. E.; REFETTOFF, S. (Ed.). **Genetic diagnosis of endocrine disorders**. 2nd ed. Amsterdam: Academic Press, 2015. p. 3–18.

SUBBURAJ, S.; CHUNG, S. J.; LEE, C.; RYU, S.-M.; KIM, D. H.; KIM, J.-S.; BAE, S.; LEE, G.-J. Site-directed mutagenesis in *Petunia × hybrida* protoplast system using direct delivery of purified recombinant Cas9 ribonucleoproteins. **Plant Cell Reports**, v. 35, n. 7, p. 1535–1544, July 2016.

SVITASHEV, S.; SCHWARTZ, C.; LENDERTS, B.; YOUNG, J. K.; CIGAN, A. M. Genome editing in maize directed by CRISPR-Cas9 ribonucleoprotein complexes. **Nature Communications**, v. 7, article 13274, Nov. 2016.

TORNEY, F.; TREWYN, B. G.; LIN, V. S.; WANG, K. Mesoporous silica nanoparticles deliver DNA and chemicals into plants. **Nature Nanotechnology**, v. 2, p. 295–300, Apr. 2007.

TORRES, E. M.; WILLIAMS, B. R.; AMON, A. Aneuploidy: cells losing their balance. **Genetics**, v. 179, n. 2, p. 737–746, June 2008.

VIANNA, G. R.; ARAGÃO, F. J. L.; RECH, E. L. O sistema biobalístico. In: BRASILEIRO, A. C. M.; CARNEIRO, V. T. C. (Ed.). **Manual de transformação genética de plantas**. 2. ed. Brasília, DF: Embrapa, 2015. p. 89–104.

VOUILLOT, L.; THÉLIE, A.; POLLET, N. Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. **G3: Genes, Genomes, Genetics**, v. 5, n. 3, p. 407–415, Mar. 2015.

WALKER, J. R.; CORPINA, R. A.; GOLDBERG, J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. **Nature**, v. 412, n. 6847, p. 607–614, Aug. 2001.

WANG, Y.; CHENG, X.; SHAN, Q.; ZHANG, Y.; LIU, J.; GAO, C.; QIU, J. L. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. **Nature Biotechnology**, v. 32, n. 9, p. 947–951, Sept. 2014.

WILBIE, D.; WALTHER, J.; MASTROBATTISTA, E. Delivery aspects of CRISPR/Cas for in vivo genome editing. **Accounts of Chemical Research**, v. 52, n. 6, p. 1555-1564, Jun. 2019.

WOLABU, T. W.; CONG, L.; PARK, J.-J.; BAO, Q.; CHEN, M.; SUN, J.; XU, B.; GE, Y.; CHAI, M.; LIU, Z.; WANG, Z.-Y. Development of a highly efficient multiplex genome editing system in outcrossing tetraploid alfalfa (*Medicago sativa*). **Frontiers in Plant Science**, v. 11, 1063, July 2020.

WOLT, J. D.; WANG, K.; YANG, B. The regulatory status of genome-edited crops. **Plant Biotechnology Journal**, v. 14, n. 2, p. 510–518, Feb. 2016.

WOO, J. W.; KIM, J.; KWON, S. I.; CORVALÁN, C.; CHO, S. W.; KIM, H.; KIM, S.-G.; KIM, S.-T.; CHOE, S.; KIM, J.-S. DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. **Nature Biotechnology**, v. 33, p. 1162–1164, Oct. 2015.

XIE, K.; MINKENBERG, B.; YANG, Y. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. **Proceedings of the National Academy of Sciences of the United States of America**, v. 112, n. 11, p. 3570-3575, Mar. 2015.

XIE, K.; YANG, Y. RNA-guided genome editing in plants using a CRISPR/Cas system. **Molecular Plant**, v. 6, n. 6, p. 1975–1983, Nov. 2013.

YANG, K.; GUO, R.; XU, D. Non-homologous end joining: advances and frontiers. **Acta Biochimica et Biophysica Sinica**, v. 48, n. 7, p. 632-640, July 2016.

YOUNG, J.; ZASTROW-HAYES, G.; DESCHAMPS, S. et al. CRISPR-Cas9 editing in maize: systematic evaluation of off-target activity and its relevance in crop improvement. **Scientific Reports**, v. 9, article 6729, Apr. 2019.

YU, C.; ZHANG, Y.; YAO, S.; WEI, Y. A PCR-based protocol for detecting indel mutations induced by TALENs and CRISPR/Cas9 in zebrafish. **PLoS ONE**, v. 9, n. 6, e98282, June 2014.

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, e2148, Feb. 2017.

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, e264, 2015.

ZHAO, X.; MENG, Z.; WANG, Y.; CHEN, W.; SUN, C.; CUI, B.; CUI, J.; YU, M.; ZENG, Z.; GUO, S.; LUO, D.; CHENG, J. Q.; ZHANG, R.; CUI, H. Pollen magnetofection for genetic modification with magnetic nanoparticles as gene carriers. **Nature Plants**, v. 3, p. 956-964, Nov. 2017.

ZHU, X.; XU, Y.; YU, S.; LU, L.; DING, M.; CHENG, J.; SONG, G.; GAO, X.; YAO, L.; FAN, D.; MENG, S.; ZHANG, X.; HU, S.; TIAN, Y. An efficient genotyping method for genome-modified animals and human cells generated with CRISPR/Cas9 system. **Scientific Reports**, v. 4, n. 6420, Sept. 2014.

ZURIS, J. A.; THOMPSON, D. B.; SHU, Y.; GUILINGER, J. P.; BESSEN, J. L.; HU, J. H.; MAEDER, M. L.; JOUNG, J. K.; CHEN, Z. Y.; LIU, D. R. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. **Nature Biotechnology**, v. 33, n. 1, p. 73-80, Jan. 2015.