

Predição *in silico* de efeitos *off-target* da aplicação de CRISPR em ambiente não controlado

Aline Martins Hoepers^{(1)*} , Alex Sandro Poltronieri⁽²⁾

(1) Acadêmico do curso de Agronomia do Centro de Ciências Agrárias da Universidade Federal de Santa Catarina. Rod. Admar Gonzaga,1346, Bairro Itacorubi, Caixa Postal 476, CEP 8840-900, Florianópolis-SC, Brasil.

(2) Professor, Depto. de Fitotecnia, Centro de Ciências Agrárias da Universidade Federal de Santa Catarina. Rod. Admar Gonzaga,1346, Bairro Itacorubi, Caixa Postal 476, CEP 8840-900, Florianópolis-SC, Brasil.

*Autor Correspondente - E-mail: alinehoepers@gmail.com

Resumo

CRISPR/Cas9, uma potente ferramenta de engenharia genética amplamente adotada na agricultura, é capaz de introduzir novas características em plantas em larga escala e sem métodos convencionais de melhoramento. Apesar da sua eficiência, surgiram preocupações relativamente às consequências não intencionais em ambientes não controlados. O objetivo desse trabalho foi avaliar a atividade fora do alvo em organismos que poderiam ser expostos à edição do genoma em ambientes não controlados. Três cenários baseados em cenários agrícolas na agricultura, nomeadamente controle de pragas e doenças, foram desenvolvidos. Usando dados de acesso público, foram previstos efeitos *off-target* em múltiplas espécies comumente encontradas no agroecossistema, incluindo humanos. A análise de enriquecimento metabólico previu efeitos no desenvolvimento dos sistemas nervoso e respiratório. As descobertas enfatizam a importância de ter precaução ao considerar o uso dessa edição genética em ambientes não controlados. Alterações genômicas não intencionais podem ocorrer em organismos não intencionais, sublinhando a importância de compreender os riscos potenciais e de ampliar a discussão sobre medidas de biossegurança para proteger a saúde humana e o ambiente.

Palavras-chave: edição genética; biossegurança; avaliação de risco; organismos não-alvo

Abstract

CRISPR/Cas9, a potent genetic engineering tool widely adopted in agriculture, is capable of introducing new characteristics into plants on a large scale and without conventional

breeding methods. Despite its remarkable efficiency, concerns have arisen regarding unintended consequences in uncontrolled environments. This work aimed to assess potential activity in organisms that could be exposed to genome editing in uncontrolled environments. Three scenarios based on outdoor uses in agriculture were developed namely pest and disease control. Using publicly available software, off-target effects were predicted in multiple species commonly found in the agroecosystem, including humans. Metabolic enrichment analysis (gene IDs) predicted effects on the development of nervous and respiratory systems. The findings emphasize the importance of exercising caution when considering the use of this genome editing in uncontrolled environments. Unintended genomic alterations may occur in unintended organisms, underscoring the significance of understanding potential hazards and implementing safety measures to protect human health and the environment.

Keywords: gene-editing; biosafety; risk assessment; non-target organisms

Introdução

A edição genética facilita a manipulação de grandes genomas, como os de plantas, fungos e animais, e, inclusive, a criar mutações nulas (por exemplo, *knockout* de genes). Com o avanço da técnica, também é possível mudar bases em locais específicos em uma molécula de DNA ou RNA (TAO; BAUER; CHIARLE, 2023). Esses avanços da tecnologia genética contribuem para o desenvolvimento de variedades agrícolas mais produtivas e agroecossistemas sustentáveis (SHI et al., 2017; ZHANG; WEN; GUO, 2014; WANG et al., 2015). A edição genética engloba diversas técnicas direcionadas para mudanças em polímeros de DNA ou RNA alvo (SEGAL; MECKLER, 2013), como métodos de mutagênese de oligonucleotídeos (ODM), TALENs, nucleases de dedo de zinco (ZFNs) e sistemas CRISPR/Cas.

O sistema CRISPR/Cas é uma técnica de edição genética de fácil manuseio (quando comparado a técnicas convencionais de melhoramento) e que permite a síntese de gRNA (RNAs guia) direcionáveis para um ácido nucleico alvo já previamente conhecido e sequenciado. Na engenharia genética o sistema funciona através de uma proteína chamada Cas9 que se liga a uma molécula de RNA, esta proteína é direcionada para um local específico no DNA ou RNA por meio de uma molécula chamada crRNA (SHARMA et al., 2022). Esta combinação forma um complexo que procura por sequências de DNA que

correspondam ao RNA. Quando o sítio de reconhecimento (PAM) NGG é encontrado a enzima Cas9 opera clivando o DNA (CUI et al., 2018).

Quando um polímero de DNA é quebrado, as células ativam vias de reparo do DNA, que pode levar à inserção ou deleção de nucleotídeos, como a Junção Final Não Homóloga (NHEJ) ou as vias de Reparo Dirigido por Homologia (HDR) (BORTESI; FISCHER, 2015; HSU et al., 2013). A ligação específica entre um guia de RNA e a sequência alvo de DNA ou RNA aumenta a taxa de quebra do polímero de ácido nucleico no local desejado. Embora a ligação entre o guia e o substrato seja importante, não é tão simples como um modelo de "chave e fechadura". Essa interação pode tolerar pequenas diferenças entre ácidos nucleicos, que variam de acordo com as condições celulares, como a concentração de íons e a temperatura. Essa faixa de tolerância pode levar a Cas a cortar o DNA em locais não desejados, denominadas regiões *off-target* (ALKAN et al., 2018).

Ferramentas de tecnologia de edição genética, como a edição genética, estão sendo desenvolvidas com o objetivo de editar organismos no campo (HEINEMANN; WALKER, 2019). A baixa eficiência dessas abordagens fora de um laboratório é limitante, mas pode melhorar (DEMIRER et al., 2021 VATS et al., 2022). O uso proposto em ambientes não controlados poderia ser viabilizado por mudanças nas regulamentações, como a criação de novos quadros regulatórios, os quais excluem algumas aplicações da edição genética das regulamentações existentes sobre tecnologia de edição genética (HEINEMANN, 2019; HEINEMANN et al., 2023). Para que a eficiência do uso em ambientes não controlados atenda a objetivos comerciais, um desafio para a avaliação de riscos surge (LI et al., 2021; MUELLER et al., 2021). Isso se deve à necessidade de considerar o ambiente receptor por completo, incluindo o alvo e todos os organismos do ecossistema durante essa avaliação.

Quando ferramentas biotecnológicas direcionadas a locais são projetadas - quer a proteína dependa ou não de um guia - os locais alvo previstos são, primeiramente, informados por pesquisas computacionais (*in silico*) de sequências de nucleotídeos em moléculas de DNA ou RNA. Nenhum método de pesquisa e design *in silico* atingiu os padrões científicos para substituir os testes diretos em animais vivos, com o objetivo de avaliar efeitos *off-target* (HANNING et al., 2013). Em um estudo sobre CRISPR/Cas9 foram detectados, em peixes-zebra (*Danio rerio*), inserções e deleções de até vários quilobases, tanto nos locais desejados quanto em locais não desejados, com variações significativas entre os resultados individuais de edição genética que não foram previstas por ferramentas de computação disponíveis (HÖIJER et al., 2022). Apesar disso, as pesquisas computacionais

são úteis para identificar locais de interesse e priorizá-los para avaliação de riscos. Embora possam gerar mais falsos positivos do que negativos, é aceitável considerando a biossegurança como prioridade. A avaliação de riscos depende de um processo rigoroso de identificação sendo um processo de descrever o que deve ser protegido (um objetivo de proteção) e formular hipóteses sobre como um risco pode afetá-lo. Esse trabalho tem como objetivo apresentar uma análise *in silico* e uma discussão dos efeitos *off-target* em organismos não alvo através dos quais a tecnologia de aplicação ao ar livre de CRISPR/Cas9 poderia causar efeitos adversos.

Material e métodos

1. Descrição dos cenários

Uma análise *in silico* foi realizada referente às potenciais atividades *off-targets* em espécies expostas, a fim de prever os impactos ambientais da tecnologia de edição genética ser utilizada para o controle de pragas e doenças. Três cenários, baseados em casos realistas, foram descritos com o objetivo de identificar hipóteses de riscos aos objetivos de proteção selecionados (determinados através das espécies escolhidas para análise). Em cada cenário, foi selecionada uma praga agrícola relevante e um gene que apresenta alto percentual de mortalidade na espécie, além de ser consideradas as estratégias atuais de aplicação de produtos agrícolas, que poderia ser replicado com uma solução CRISPR/Cas9 (Figura 1). Os genes letais foram identificados com base na literatura sobre aqueles utilizados e mapeados em tecnologias similares às propostas neste artigo, como técnicas de dsRNA e RNAi.

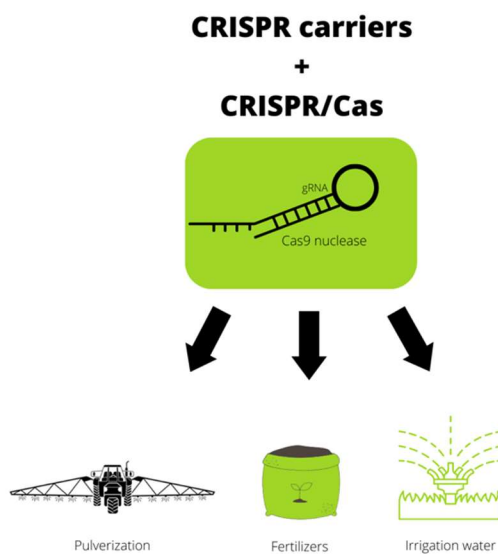


Figura 1. Resumo gráfico dos métodos de aplicação propostos para uma tecnologia ao ar livre de CRISPR/Cas. Neste estudo, foram consideradas as aplicações de moléculas CRISPR/Cas através de água de irrigação (irrigação por aspersão e/ou pivô central) –

Cenário #1, tecnologia de fumigação (preencher uma área com pesticidas gasosos) – Cenário #2, e estratégias de fertilização (aplicações diretas no solo).

No primeiro cenário, a lagarta do milho (*Diabrotica virgifera virgifera*) foi selecionada como alvo por ser uma praga principal do milho (*Zea mays*). Suas larvas se alimentam das raízes, prejudicando seu desenvolvimento (MEINKE et al., 2009). Essa praga é considerada uma das principais nos Estados Unidos (SAPPINGTON et al., 2006) e na Europa (MILLER et al., 2005), especialmente devido à resistência que desenvolveu devido ao uso excessivo de inseticidas (MEINKE; SOUZA; SIEGFRIED, 2021). No segundo cenário, o besouro vermelho da farinha (*Tribolium castaneum*) foi escolhido como alvo. *T. castaneum* é uma praga global de alimentos armazenados (CAMPBELL et al., 2022), infestando uma variedade de farinhas de cereais (AGARWAL; AGASHE, 2020). No terceiro cenário, o fungo necrotrófico *Sclerotinia sclerotium* foi o alvo. Este fungo é um patógeno agrícola importante que pode afetar mais de 400 culturas, incluindo *Glycine max*, *Phaseolus vulgaris*, *Gossypium hirsutum* e *Brassica napus* (BOLAND; HALL, 1994). Em condições favoráveis, pode causar perdas de rendimento acima de 50% (ALKOORANEE et al., 2017).

A aplicação ao ar livre de CRISPR/Cas resulta em exposições diretas e indiretas a organismos alvo e não alvo por meio de contato, inalação ou ingestão (Figura 2). A liberação de reagentes, como aerossóis ou géis, cria exposições de contato em plantas ou pragas alvo. As exposições indiretas decorrem da ingestão de organismos que foram expostos e contêm reagentes em forma ativa em suas células. Por exemplo, os polinizadores poderiam ingerir pólen e os insetos predadores poderiam ingerir pragas que internalizaram a pulverização. Para invertebrados herbívoros, animais de fazenda e organismos do solo, a via de exposição são folhas, raízes e grãos (BACHMAN et al., 2016).

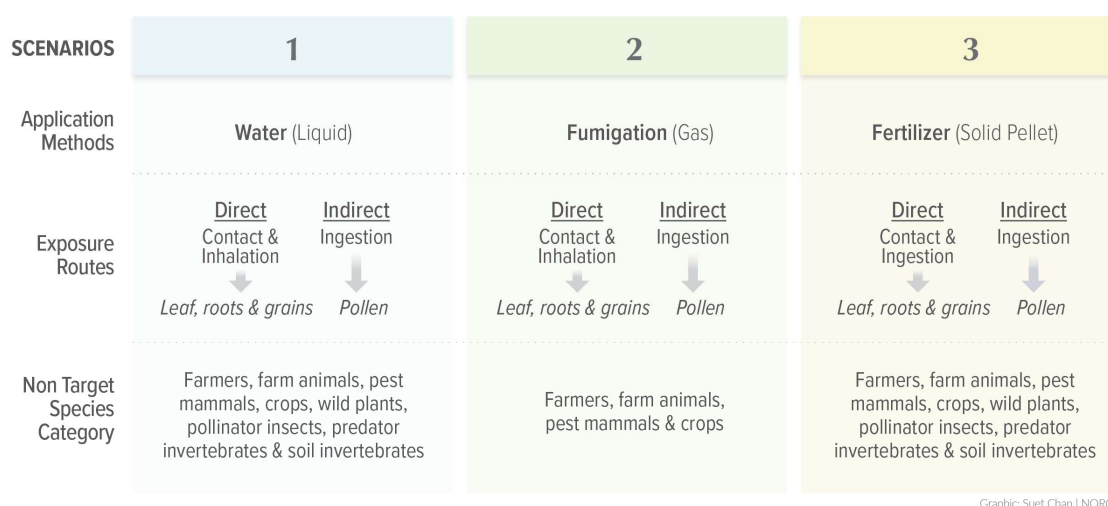





Figura 2. Representação esquemática das rotas de exposição direta e indireta para as espécies não-alvo selecionadas para este estudo. Para cada cenário (linha superior), são ilustrados um método de aplicação (segunda linha), uma via de exposição (terceira linha) e uma exposição não visada (linha inferior).. A ingestão pode ser direta (por exemplo, beber néctar contaminado) ou indireta (por exemplo, comer pólen que internalizou os reagentes). As rotas de exposição indireta incluem a ligação através de redes alimentares. (Fonte: Suet Chen|NORCE)

2. Seleção de genes alvo e design de gRNA

O gene ortólogo *Snf7 DvSnf7* foi escolhido para o cenário #1 como alvo para *Diabrotica virgifera virgifera* (Gene ID: 114337301) por ter sido descrito como um alvo potencial de RNAi para o controle de *D. v. virgifera* (BOLOGNESI et al., 2012). O dsRNA de 240 pb descrito por Bolognesi et al. (2012) foi identificado no Centro Nacional de Informações sobre Biotecnologia (NCBI) usando a Ferramenta Básica de Pesquisa de Alinhamento Local (BLAST), contra o genoma da *D. v. virgifera* (taxid:50390). Para o cenário #2, o gene *RpII140* (Gene ID: 663520) de *Tribolium castaneum* foi selecionado como alvo. Este gene faz parte da subunidade da RNA polimerase II dirigida por DNA que catalisa a transcrição do DNA em RNA (KNORR et al., 2018). Quando usado como alvo para RNAi, o silenciamento mata 88,89% dos *T. castaneum* expostos. Para o cenário #3, foi escolhido o gene *SS1G_01703 - ABHYDROLASE-3* (Gene ID: 5493374) do gene *Sclerotinia sclerotiorum*, previamente identificado como um potencial alvo de RNAi (WYTINCK et al., 2022). A descrição de cada cenário desse estudo está apresentada na Figura 3.

CASE	Description	Target Species	Target Gene (Genebank ID)	Knock-Out Metabolism
1	Western corn rootworm infestation in maize fields	<i>Diabrotica virgifera virgifera</i>  <small>© Russel IFM (source: russellipm.com)</small>	Charged multivesicular body protein 4b (Gene ID: 114337301)	Essential protein involved in intracellular trafficking
2	Red flour beetle infestation in stored cereals	<i>Tribolium castaneum</i>  <small>© Geosty Goergen (source: CABI CPC)</small>	Rpl140 (Gene ID: 658374)	Part of the DNA-directed RNA polymerase II subunit that catalyzes the transcription of DNA into RNA
3	White mold infection in soybean, common beans, and cotton fields	<i>Sclerotinia sclerotium</i> 	ABHYDROLASE-3 gene (Gene ID: 593374)	Predicted aflatoxin biosynthesis gene and pathogenicity factor

Graphic: Suet Chan | NORCE

Figura 3. Descrição de cada cenário investigado neste estudo. (Suet Chan |NORCE)

Os gRNAs foram projetados através da ferramenta online CRISPOR (CONCORDET; HAEUSSLER, 2018) e foram selecionados com base no alto conteúdo de GC% (entre 40 e 70%) e altas pontuações *out-of-frame* (eficiência de nocaute completo no alvo e pontuação de especificidade). O número de incompatibilidades em locais *off-target* no genoma alvo também foi considerado para evitar locais de ligação em gêneros e espécies relacionados. Os gRNAs com menores números de incompatibilidades foram selecionados para fornecer maior especificidade das sequências de gRNA.

3. Predição *in silico* de sequências *off-target* CRISPR

Para identificar possíveis atividades *off-targets*, as três sequências de gRNA foram analisadas pelo Cas-OFFinder no CRISPR RGEN Tools (PARK; BAE; KIM, 2015). A ferramenta CasOFFinder realiza pesquisas de DNA com base em genomas de referência de espécies selecionadas. Aqui, a escolha de espécies foi baseada nos organismos de cada cenário. Esta plataforma lê os dados da sequência de entrada, que nesse estudo foram sequências de gRNA projetadas para cada cenário, e coleta informações sobre locais específicos contendo sequências *protospacer adjacent motifs* (PAM). Essas sequências são então entregues a um comparador, que conta o número de incompatibilidades e seleciona potenciais locais *off-targets* que possuem menos números de incompatibilidades do que os parâmetros fornecidos (BAE et al., 2014). Neste caso, os parâmetros permitiram até quatro incompatibilidades de bases, além da região (PAM) idêntica, neste caso, a região PAM com padrão 5'-NGG-3' de SpCas9 de *Streptococcus pyogenes* (SMITH et al., 2020). A plataforma

CasOFFinder fornece uma saída de possíveis sequências de DNA *off-target* com as bases e número de incompatibilidade, juntamente com o número do cromossomo, posição, direção e número de bases incompatíveis. O resultado do Cas-OFFinder foi analisado na ferramenta NCBI Genome Data Viewer (GDV), para identificar sequências de DNA correspondentes (RANGWALA et al., 2021). As informações fornecidas ao GDV incluíram cromossomo, localização e organismo analisado. Além disso, foi avaliada se a sequência *off-target* analisada corresponde a um gene e, em caso afirmativo, qual gene (RANGWALA et al., 2021). Os alvos que não correspondiam a nenhum gene foram rotulados como “regiões intergênicas”.

4. Enriquecimento metabólico de sequências *off-target*

A ferramenta de Pesquisa para Recuperação de Genes/Proteínas Interagentes (STRING) foi usada para montar uma rede metabólica com os genes *off-target*. STRING é uma ferramenta online que identifica conexões entre proteínas, avaliadas e comparadas com os mapas de vias KEGG (Enciclopédia de Genes e Genomas de Kyoto) (SZKLARCZYK et al., 2023). O KEGG reconstrói sistemas biológicos computacionalmente capturando informações de dados publicados na literatura e acumulados no banco de dados KEGG PATHWAY (KANEHISA et al., 2021). A saída do STRING é uma pontuação de confiança para cada associação de proteínas, variando de zero a um, indicando a probabilidade de ser verdadeira com base nas evidências disponíveis nos bancos de dados KEGG. Essas pontuações de confiança são calculadas para cada tipo de evidência e integradas para gerar uma pontuação final para uma determinada via metabólica (SZKLARCZYK et al., 2023).

O conjunto de entrada incluiu o gene e a lista de ocorrências intergênicas para cada espécie em cada cenário. No total, foram realizadas 12 análises em nove organismos utilizando STRING. Os acessos genéticos foram enriquecidos pelo algoritmo da plataforma para conectar redes metabólicas e vias bioquímicas para a análise das seguintes categorias: Biological Process (Gene Ontology), Molecular Function (Gene Ontology), Cellular Component (Gene Ontology), Reference publications (PubMed), Local network cluster (STRING), KEGG Pathways, Reactome Pathways, WikiPathways, Tissue expression (TISSUES), Human Phenotype (Monarch), Annotated Keywords (UniProt), Protein Domains (Pfam), Protein Domains and Features (InterPro), Protein Domains (SMART). Para algumas espécies, a categoria de associações gene-doença (DISEASES) também foi analisada. Vias com False Discovery Rate (FDR) < 0,05 foram considerados significativamente enriquecidos.

Resultados e discussão

1. Atividade prevista de gRNA em organismos alvo

Foram escolhidas sequências de gRNA para cada cenário com base na sua posição dentro de um exon do gene correspondente, teor de GC% e pontuação *out-of-frame* (Tabela 1). Para os gRNAs selecionados nos cenários #1 e #3, nenhum local de ligação *off-target* foi previsto nos genomas dos organismos alvos. Ao permitir até quatro *incompatibilidades*, previu-se que o gRNA selecionado com melhor classificação no cenário # 2 teria três locais de ligação *off-target* no genoma alvo do *Tribolium castaneum*. Esses locais *off-target* estavam em regiões intergênicas (NC_007424.3), intrônicas (NC_007422.5) e exônicas (NC_007421.3).

Tabela 1. Parâmetros de design de gRNA para organismos alvo nos três cenários. Os parâmetros relevantes do gRNA incluem: a posição no gene alvo (localização do gene e sentido da fita), em qual exon o gRNA está localizado, conteúdo CG (porcentagem de guanina e citosina na sequência de gRNA), pontuação fora do quadro (0-100; previsão da probabilidade de um gene nocaute). As últimas 5 colunas são acertos *off-target* previstos, categorizados pelo número de incompatibilidades de nucleotídeos em relação à sequência de gRNA.

Cenário	Alvo e ID do gene	Sequência de gRNA (5' a 3')	Posição (pb)/fita	Éxon	Conteúdo do GC (% sem PAM)	Pontuação <i>out-of-frame</i>	Incompatibilidades				
							0	1	2	3	4
#1	Dvv ID do gene: 114337301	TGAATAGCC TCCCCAGGA CT AGG	130/rev	1	55%	67	0	0	0	0	0
#2	Tc ID do gene: 663520	GTCCGAATA TCCCGAATC A TGG	86/rev	2	50%	55	0	0	0	0	3
#3	Ss ID do gene: 5493374	GGCGCATAA CCCTCGGTTG A TGG	10/rev	1	60%	65	0	0	0	0	0

2. Identificação de objetivos de proteção e exposições relevantes

Para definir e identificar os objetivos de proteção, cada cenário envolveu um organismo alvo e um ambiente específico. Para ser considerado, o objetivo de proteção tinha que ser um organismo não alvo (ONA) com papel importante no ecossistema agrícola. Além disso, deveria existir a possibilidade de exposição a agentes da engenharia genética, seja diretamente ou através das cadeias alimentares primárias. A lista de espécies analisadas

contém seres humanos (por exemplo, os agricultores), animais, espécies de interesse e os organismos benéficos (polinizadores, predadores, invertebrados do solo) (Tabela 2). A seleção de espécies foi limitada pelo desequilíbrio da representação destas nas bases de dados disponíveis publicamente. Por exemplo, *Eisenia andrei*, *Coleomegilla maculata*, *Poecilus calcitans* (organismos do solo), *Coccinella septempunctata* (predador invertebrado), *Parthenium hysterophorus* e *Brachiaria reptans* (plantas selvagens que também podem ser ervas daninhas) estão sub-representadas. Para superar essa limitação, foram selecionadas espécies relacionadas disponíveis com função ecológica igual ou semelhante para cada cenário (ou seja, *Metarhizium anisoplia* ; *Brassica sp.*, *Sorghum bicolor*, *Panicum virgatum*, *Setaria viridis*).

Tabela 2. Função ecológica e espécies não-alvo analisadas (indicadas por pontos) em cada cenário. Os organismos não incluídos num cenário são indicados por traços.

Categoria ONA/Função ecológica	Organismos não-alvo	Cenário #1 Alvo DvSnf7	Cenário #2 Alvo RpII140	Cenário #3 Alvo ABHYDRO LASE3
Agricultor	<i>Homo sapiens</i> (humano)	•	•	•
Animais de fazenda (ruminantes e onívoros)	<i>Bos taurus</i> (gado)	•	•	•
	<i>Gallus gallus</i> (frango)	•	•	•
Mamíferos pragas (roedores)	<i>Mus musculus</i> (rato)	•	•	•
Culturas (organismos fotossintéticos e cicladores de nutrientes)	<i>Zea mays</i> (milho)	•	•	-
	<i>Glycine max</i> (soja)	•	-	•
	<i>Phaseolus vulgaris</i> (feijão)	-	-	•
	<i>Gossypium hirsutum</i> (algodão)	-	-	•
Plantas selvagens (organismos fotossintéticos e cicladores de nutrientes)	<i>Brassica sp.</i>	•	-	•
	<i>Sorgo bicolor</i>	•	-	•
	<i>Panicum virgatum</i>	•	-	•
	<i>Setaria viridis</i>	•	-	•
Insetos polinizadores	<i>Apis mellifera</i> (abelha)	•	-	•
	<i>Bombus terrestris</i> (mamangava-de-cauda-amarela-clara)	•	-	•

Invertebrados predadores	<i>Camponotus floridanus</i> (formiga carpinteira)	•	-	•
	<i>Lasius Niger</i> (formiga preta)	•	-	•
Invertebrados do solo (cicladores de nutrientes)	<i>Metarhizium anisoplia</i> (fungo muscardina verde)	•	-	•
	<i>Caenorhabditis elegans</i> (lombriga)	•	-	•

3. Locais de atividade previstos de gRNA em organismos não-alvo

A atividade *off-target* pode ser evitada ou minimizada usando gRNAs que não possuam potencial de hibridização em nenhum lugar dos genomas do ONA, ou garantindo que os ONAs não sejam expostos aos gRNAs. Apesar da distância evolutiva entre organismos alvo (insetos e fungos) e organismos não-alvo (humanos e animais), estes últimos detinham os locais de ligação *off-target* mais previstos para gRNA. Nenhum *off-target* foi identificado nos genomas dos predadores e dos invertebrados do solo analisados, mas esse resultado deve ser interpretado com precaução, pois resultados falsos negativos podem surgir de vários fatores. Em primeiro lugar, pode ser devido a um pequeno número de genomas de algumas espécies de ONA em bases de dados disponíveis publicamente. Em segundo lugar, a diversidade intra-espécies conhecida dos ONAs pode variar. O genoma do *Homo sapiens* foi revisado pela última vez no NCBI em 2022, mas o genoma do *Caenorhabditis elegans* foi revisado pela última vez em 2013. Em terceiro lugar, os modelos *in silico* não conseguem prever as variáveis bioquímicas que podem influenciar a hibridização. Portanto, uma análise *in silico* como esta pode fornecer uma falsa garantia de alguns sítios de ligação em espécies alvo de proteção sub-representadas nas bases de dados. Dos 18 ONAs selecionados, foram encontrados potencial hibridização *off-target* em 12 espécies, e estas variaram de um *off-target* (por exemplo, cenário #1 *Apis mellifera*) a 36 (cenário #2, *Homo sapiens*), dependendo do organismo alvo (Tabela 3).

Tabela 3. Número de locais de ligação *off-target* não intencionais previstos em ONAs para os 3 cenários.

Organismo não alvo (nome da espécie)	Cenário # 1 (<i>Diabrotica virgifera virgifera</i>)	Cenário # 2 (<i>Tribolium castaneum</i>)	Cenário # 3 (<i>Sclerotinia sclerotium</i>)
<i>Homo sapiens</i> (humanos)	24 <i>off-target</i> (8 regiões intergênicas)	36 <i>off-target</i> (15 regiões intergênicas)	5 <i>off-target</i> (2 regiões intergênicas)
<i>Bos taurus</i> (gado)	19 <i>off-target</i> (11 regiões intergênicas)	3 <i>off-target</i> (1 região intergênica)	-

<i>Gallus gallus</i> (frango)	9 off-target (5 regiões intergênicas)	-	-
<i>Mus musculus</i> (rato)	29 off-target (17 regiões intergênicas)	4 off-target (1 região intergênica).	4 off-target (1 região intergênica)
<i>Zea mays</i> (milho)	4 off-target (2 regiões intergênicas)	1 off-target (1 região intergênica)	-
<i>Glycine max</i> (soja)	2 off-target (1 região intergênica)	-	1 off-target (1 região intergênica)
<i>Gossypium hirsutum</i> (algodão)	-	-	1 off-target (1 região intergênica)
<i>Sorgo bicolor</i> (painço)	3 off-target (2 regiões intergênicas)	-	2 off-target
<i>Panicum virgatum</i>	2 off-target	-	2 off-target (2 regiões intergênicas)
<i>Setaria viridis</i>	2 off-target (1 região intergênica)	-	-
<i>Apis mellifera</i> (abelha)	-	-	1 off-target
<i>Caenorhabditis elegans</i>	1 off-target (1 região intergênica)	-	-

Locais de ligação potenciais foram identificados em quase metade (oito de 16) dos organismos não-alvo no cenário #1 (Tabela 4). Nestas oito espécies, o número de potenciais locais de ligação variou de um, na soja e no sorgo, a 16, nos humanos. O número de incompatibilidades variou de dois (bovinos) a quatro. Os humanos foram os mais propensos a serem afetados pela exposição à água de irrigação CRISPR/Cas com base no número de locais potenciais de hibridização.

Tabela 4. Potenciais locais de hibridização *off-target* do gene *DvSnf7* em ONAs (Cenário #1)

Organismo	<i>incompatibilidad es</i>	Nome do Gene	Gene ID
<i>Homo sapiens</i>	3	MTHFD1L - methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1 like	25902
<i>Homo sapiens</i>	4	EGF like, fibronectin type III and laminin G domains	133584
<i>Homo sapiens</i>	4	LOC105376737	105376737
<i>Homo sapiens</i>	4	neuron navigator 1	89796
<i>Homo sapiens</i>	4	cAMP responsive element binding protein 5	9586
<i>Homo sapiens</i>	4	FLII actin remodeling protein	2314
<i>Homo sapiens</i>	4	ADAMTS like 1	92949
<i>Homo sapiens</i>	4	zinc finger protein 462	58499
<i>Homo sapiens</i>	4	WHRN - whirlin	25861
<i>Homo sapiens</i>	4	NHS like 2	340527

<i>Homo sapiens</i>	4	Neurexin 3	9369
<i>Homo sapiens</i>	4	KHDRBS2 - KH RNA binding domain containing, signal transduction associated 2	202559
<i>Homo sapiens</i>	4	FYN proto-oncogene, Src family tyrosine kinase	2534
<i>Homo sapiens</i>	4	DENN domain containing 2B	6764
<i>Homo sapiens</i>	4	LOC105370324	105370324
<i>Homo sapiens</i>	4	RAB43, member RAS oncogene family	339122
<i>Bos taurus</i>	2	zinc finger protein, FOG family member 1	528338
<i>Bos taurus</i>	4	ALDH4A1 - aldehyde dehydrogenase 4 family member A1	100126042
<i>Bos taurus</i>	4	adrenoceptor alpha 1A	282134
<i>Bos taurus</i>	4	ryanodine receptor 3	539899
<i>Bos taurus</i>	4	serine/threonine kinase like domain containing 1	614899
<i>Bos taurus</i>	4	rabphilin 3A like (without C2 domains)	617294
<i>Bos taurus</i>	4	galactosidase beta 1	507188
<i>Bos taurus</i>	4	cytoplasmic polyadenylation element binding protein 3	537016
<i>Gallus gallus</i>	3	GTPase IMAP family member 7-like 2	101747807
<i>Gallus gallus</i>	4	cAMP responsive element binding protein 5	428435
<i>Gallus gallus</i>	4	RasGEF domain family member 1A	423631
<i>Gallus gallus</i>	4	alpha kinase 1	100858173
<i>Mus musculus</i>	3	Arhgap15 Rho GTPase activating protein 15	76117
<i>Mus musculus</i>	3	Mall, T cell differentiation protein-like	228576
<i>Mus musculus</i>	3	myocardin	214384
<i>Mus musculus</i>	3	TSPO associated protein 1	207777
<i>Mus musculus</i>	4	solute carrier family 10 (sodium/bile acid cotransporter family), member 7	76775
<i>Mus musculus</i>	4	quiescin Q6 sulfhydryl oxidase 1	104009
<i>Mus musculus</i>	4	DENN domain containing 1A	57706
<i>Mus musculus</i>	4	predicted gene, 30319	102632174
<i>Mus musculus</i>	4	utrophin	22288
<i>Mus musculus</i>	4	death-associated protein kinase 3	13144
<i>Mus musculus</i>	4	cardiac mesoderm enhancer-associated non-coding RNA	328968
<i>Zea mays</i>	4	LOC100381711 putative WAK receptor-like protein kinase family protein	100381711
<i>Zea mays</i>	4	LOC100283564 transferase/ transferase, transferring glycosyl groups	100283564
<i>Glycine max</i>	4	uncharacterized LOC100801289	100801289
<i>Panicum virgatum</i>	4	LOC120681520	120681520
<i>Panicum virgatum</i>	4	LOC120687334	120687334
<i>Sorghum bicolor</i>	4	loc8082872	8082872

Cinco organismos não-alvo foram analisados no cenário #2. Potenciais locais de ligação foram encontrados em quatro deles (Tabela 5), variando de três no genoma do *Mus*

musculus a 21 no *Homo sapiens*. Os potenciais locais de ligação no genoma do *Homo sapiens* incluíam proteínas com atividade antimicrobiana, como as codificadas pelo gene da defensina beta. Outros possíveis locais de ligação nos genomas humanos incluem o gene delta do sarcoglicano, que é produzido no músculo esquelético e cardíaco. Para *Mus musculus*, os potenciais locais de ligação off-target estavam em genes envolvidos no funcionamento do aparelho de Golgi, na camada ventricular e no desenvolvimento do sistema nervoso.

Tabela 5. Potenciais locais de hibridização *off-target* do gene RpII140 em ONAs (Cenário #2)

Organismo	incompatibilidades	Nome do gene	Gene ID
<i>Homo sapiens</i>	3	SUFU negative regulator of hedgehog signaling	51684
<i>Homo sapiens</i>	4	defensin beta 107B	503614
<i>Homo sapiens</i>	4	sarcoglycan delta	6444
<i>Homo sapiens</i>	4	solute carrier family 44 member 5	204962
<i>Homo sapiens</i>	4	LOC100419798 gon-4 like pseudogene	100419798
<i>Homo sapiens</i>	4	flavin containing dimethylaniline monooxygenase pseudogene	9,116123
<i>Homo sapiens</i>	4	complement C3b/C4b receptor 1	1378
<i>Homo sapiens</i>	4	uncharacterized LOC122455341	122455341
<i>Homo sapiens</i>	4	endogenous retrovirus group K member 6 Env polyprotein	124901580
<i>Homo sapiens</i>	4	LOC105369787 endogenous retrovirus group K member 21 Env polyprotein	105369787
<i>Homo sapiens</i>	4	MIR548XHG (MIR548X host gene)	101927797
<i>Homo sapiens</i>	4	G protein subunit alpha q	2776
<i>Homo sapiens</i>	4	meiotic double-stranded break formation protein 4	101928601
<i>Homo sapiens</i>	4	uncharacterized LOC105376655	105376655
<i>Homo sapiens</i>	4	asparaginase and isoaspartyl peptidase 1	80150
<i>Homo sapiens</i>	4	LOC124902738 endogenous retrovirus group K member 25 Env polyprotein	124902738
<i>Homo sapiens</i>	4	LOC124902766 endogenous retrovirus group K member 7 Env polyprotein-like	124902766
<i>Homo sapiens</i>	4	uncharacterized LOC124904606	124904606
<i>Homo sapiens</i>	4	zinc finger protein 420	147923
<i>Homo sapiens</i>	4	LOC124906262 endogenous retrovirus group K member 5 Gag polyprotein	124906262
<i>Homo sapiens</i>	4	ectonucleotide pyrophosphatase/phosphodiesterase pseudogene 4	7100421816
<i>Bos taurus</i>	4	chromosome 5 C12orf75 homolog	100270756
<i>Bos taurus</i>	4	mediator complex subunit 13L	517336

<i>Mus musculus</i>	3	polypeptide N-acetylgalactosaminyltransferase 9	231605
<i>Mus musculus</i>	4	Mir9-3hg	101694
<i>Mus musculus</i>	4	erythrocyte membrane protein band 4.1 like 3	13823

Locais de ligação potenciais foram encontrados em quatro dos dezessete genomas de ONA analisados no cenário #3 (Tabela 6). Eles variaram entre um encontrado no genoma de *Apis mellifera* e três no *Homo sapiens*. Um gene potencialmente afetado foi a metaloproteinase de matriz-2 de *Apis mellifera*, envolvido no processamento de proteínas extracelulares.

Tabela 6. Potenciais locais de hibridização *off-target* do gene ABHYDROLASE-3 em ONAs (Cenário #3)

Organismo	incompatibilidades	Nome do gene	Gene ID
<i>Homo sapiens</i>	4	zinc finger protein 92	168374
<i>Homo sapiens</i>	4	GDP-mannose 4,6-dehydratase	2762
<i>Homo sapiens</i>	4	Rho/Rac guanine nucleotide exchange factor 18	23370
<i>Mus musculus</i>	4	predicted gene 5364	384931
<i>Mus musculus</i>	4	synaptic vesicle glycoprotein 2c	75209
<i>Mus musculus</i>	4	tet methylcytosine dioxygenase 2	214133
<i>Sorghum bicolor</i>	4	LOC8082218 - protein GAMETE EXPRESSED 3	8082218
<i>Sorghum bicolor</i>	3	LOC8070878 - E4 SUMO-protein ligase PIAL2	8070878
<i>Apis mellifera</i>	4	LOC409383 matrix metalloproteinase-2	409383

4. Distúrbios metabólicos previstos em ONAs expostos

Uma análise da via de enriquecimento metabólico foi utilizada para prever possíveis efeitos das hibridizações de gRNA em ONAs. Essa análise baseia-se na suposição de que a hibridização de gRNA nas sequências de DNA complementares identificadas acima resultaria na perda da função anotada dos genes através de uma reação de edição *off-target*. Para construir hipóteses de potenciais efeitos adversos da interrupção desses genes, foram usados os bancos de dados KEGG e GenBank para recuperar informações sobre metabolismo, processos biológicos e respostas fisiológicas após exposição ao CRISPR/Cas. A partir disso, as principais vias metabólicas estatisticamente significativas foram identificadas. Vários locais de hibridização de genes (21) das seções acima não foram encontrados pela pesquisa no banco de dados de strings porque são compostos principalmente de ncRNA, Long non-coding RNA ou pseudogenes e não foram anotados.

4.1 Humano (*Homo sapiens*)

Usando as previsões de KEGG, a ligação potencial do gRNA direcionado ao DvSnf7 a locais identificados no genoma do *Homo sapiens* poderia resultar na atividade de edição

que alterou a função dos genes que afetam 43 vias metabólicas no cenário # 2. Os cenários #2 e #3 apresentam resultados onde os efeitos adversos incluem surdez, câncer do sistema nervoso central, síndrome do carcinoma basocelular nevóide e meduloblastoma. Quando montada a rede metabólica para os genes afetados de *Homo Sapiens* para o cenário 2 (Figura 3), três interações entre oito genes distintos foram encontradas.

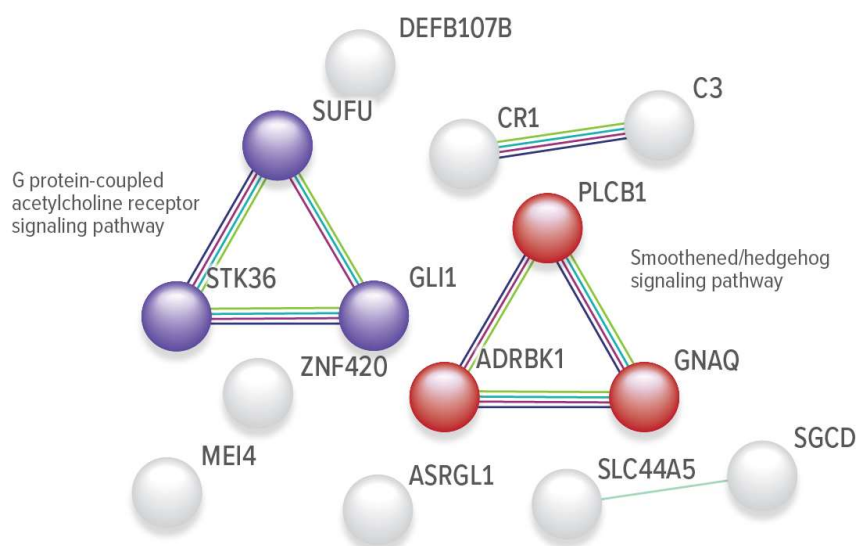


Figura 3. Rede de interação de proteínas enriquecidas por banco de dados para o cenário #2. A análise de enriquecimento da via KEGG em sítios de ligação do gRNA ilustrados no cenário #2 identifica efeitos potenciais nas vias envolvidas na sinalização hedgehog, na sinapse glutamatérgica, no câncer, no sistema endócrino e na reabsorção de cálcio, entre outros (Tabela 7).

Tabela 7. ID das vias metabólicas; descrição; contagem genética observada; e taxa de falsa descoberta para o *Homo sapiens* no cenário #2.

ID	Metabolismo	Descrição do metabolismo	Contagem genética	Taxa de falsa descoberta
hsa04340	Hedgehog signaling pathway		3	0.0023
GO:0007213	G protein-coupled acetylcholine receptor signaling pathway		3	0.0057
hsa04724	Glutamatergic synapse		3	0.01
hsa05142	Chagas disease		3	0.01
hsa04062	Chemokine signaling pathway		3	0.0311
hsa05143	African trypanosomiasis		2	0.0311
hsa05200	Pathways in cancer		4	0.0311
hsa04961	Endocrine and other factor-regulated calcium reabsorption		2	0.0381
GO:0007224	Smoothened signaling pathway		3	0.0408

hsa04071	Sphingolipid signaling pathway	2	0.041
hsa04540	Gap junction	2	0.041
hsa04713	Circadian entrainment	2	0.041
hsa04720	Long-term potentiation	2	0.041
hsa04725	Cholinergic synapse	2	0.041
hsa04726	Serotonergic synapse	2	0.041
hsa04730	Long-term depression	2	0.041
hsa04750	Inflammatory mediator regulation of TRP channels	2	0.041
hsa04911	Insulin secretion	2	0.041
hsa04912	GnRH signaling pathway	2	0.041
hsa04916	Melanogenesis	2	0.041
hsa04918	Thyroid hormone synthesis	2	0.041
hsa04922	Glucagon signaling pathway	2	0.041
hsa04924	Renin secretion	2	0.041
hsa04925	Aldosterone synthesis and secretion	2	0.041
hsa04927	Cortisol synthesis and secretion	2	0.041
hsa04928	Parathyroid hormone synthesis, secretion and action	2	0.041
hsa04929	GnRH secretion	2	0.041
hsa04935	Growth hormone synthesis, secretion and action	2	0.041
hsa04970	Salivary secretion	2	0.041
hsa04971	Gastric acid secretion	2	0.041
hsa04972	Pancreatic secretion	2	0.041
hsa05146	Amoebiasis	2	0.041
hsa05217	Basal cell carcinoma	2	0.041
hsa04611	Platelet activation	2	0.0416
hsa04728	Dopaminergic synapse	2	0.0442
hsa04270	Vascular smooth muscle contraction	2	0.0448
hsa04371	Apelin signaling pathway	2	0.0448
hsa04915	Estrogen signaling pathway	2	0.0448
hsa05017	Spinocerebellar ataxia	2	0.0448
hsa04723	Retrograde endocannabinoid signaling	2	0.0487
hsa04261	Adrenergic signaling in cardiomyocytes	2	0.0488
hsa04921	Oxytocin signaling pathway	2	0.0488
hsa04934	Cushing syndrome	2	0.0436

4.2 Gado doméstico (*Bos taurus*) e frango (*Gallus gallus*)

O enriquecimento metabólico para o cenário #1 *Bos taurus* mostrou efeitos potenciais em 13 vias metabólicas. Possíveis alterações ocorrem no metabolismo da arginina e da prolina, alanina, aspartato e glutamato (Tabela 8).

Tabela 8. ID da via metabólica; descrição do processo; contagem genética observada; e taxa de descoberta falsa para *Bos taurus* e *Gallus gallus* no cenário #1.

ID metabolismo	Descrição do metabolismo	contagem genética observada	Taxa de falsa descoberta
<i>Bos taurus</i>			

bta00330	Arginine and proline metabolism	5	4.52E-08
GO:0006536	Glutamate metabolic process	5	5.56E-08
GO:0006560	Proline metabolic process	4	6.32E-07
GO:0009065	Glutamine family amino acid catabolic process	4	2.68E-06
GO:0010133	Proline catabolic process to glutamate	3	8.44E-06
GO:0010724	Regulation of definitive erythrocyte differentiation	2	0.0023
GO:0071733	Transcriptional activation by promoter-enhancer looping	2	0.0035
bta01100	Metabolic pathways	6	0.0232
GO:0030220	Platelet formation	2	0.0254
GO:0030219	Megakaryocyte differentiation	2	0.0297
GO:0030851	Granulocyte differentiation	2	0.0297
GO:0035162	Embryonic hemopoiesis	2	0.0297
bta00250	Alanine, aspartate and glutamate metabolism	2	0.0332
Gallus gallus			
gga04261	Adrenergic signaling in cardiomyocytes	3	0.0026
gga04137	Mitophagy - animal	2	0.0273

Quando montada a rede metabólica para os genes afetados de *Bos Taurus* para o cenário 1 (Figura 4), quatro interações entre 14 genes distintos foram encontradas.

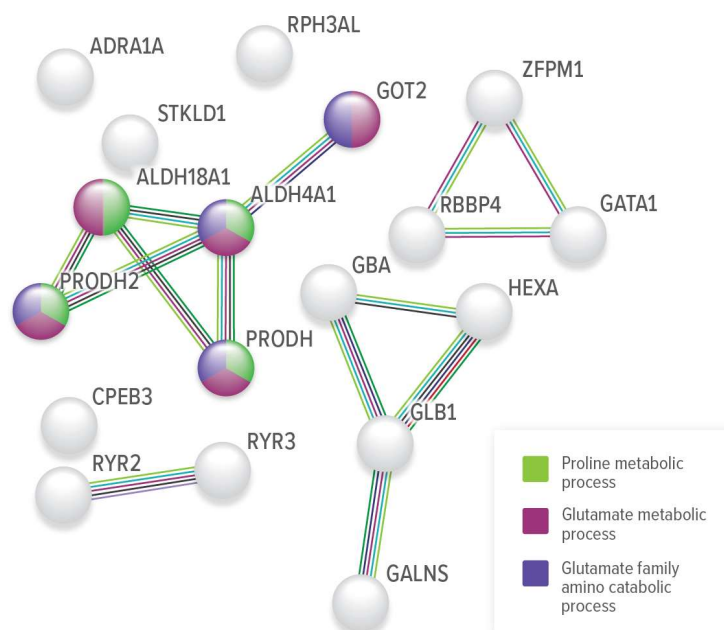


Figura 4. Rede de interação de proteínas enriquecidas por banco de dados para o cenário #1 do *Bos taurus*.

O enriquecimento metabólico baseado em genes potencialmente alterados pela edição em *Gallus gallus* identificou efeitos na contração do músculo cardíaco. A via da mitofagia também foi potencialmente afetada. Esta via está ligada à autofagia mitocondrial,

um processo fisiológico que contribui para a integridade de uma rede mitocondrial funcional e para a prevenção da morte celular programada através da remoção de mitocôndrias danificadas.

4.3 Rato (*Mus musculus*)

O enriquecimento metabólico para o cenário #2 do *Mus musculus* identificou quatro processos biológicos que poderiam afetar a homeostasia, a resposta imune, a inflamação, a embriogênese e o desenvolvimento do tecido neuronal (Tabela 9).

Tabela 9. ID da via metabólica; descrição do processo; contagem genética observada; e taxa de descoberta falsa para *Mus musculus* no cenário #2

ID metabolismo	Descrição do metabolismo	Contagem genética observada	taxa de descoberta falsa
GO:0071205	Protein localization to juxtaparanode region of axon	3	1.41E-05
mmu04514	Cell adhesion molecules	3	0.0041
GO:0002175	Protein localization to paranode region of axon	2	0.0078
GO:0030913	Paranodal junction assembly	2	0.01

4.4 Abelha (*Apis mellifera*)

O enriquecimento metabólico para o cenário #3 de *Apis mellifera* revelou 39 processos biológicos afetados (Tabela 10). As vias estão associadas ao desenvolvimento do sistema nervoso, do sistema respiratório, do coração, do coração larval e do zigoto para um animal multicelular.

Tabela 10. ID da via metabólica; descrição do processo; contagem genética observada; e taxa de falsa descoberta para *Apis mellifera* no cenário #3.

ID metabolismo	Descrição do metabolismo	contagem genética observada	taxa de descoberta falsa
GO:0030198	Extracellular matrix organization	5	4.07E-06
GO:0030574	Collagen catabolic process	3	8.59E-05
GO:0007275	Multicellular organism development	9	0.00065
GO:0003344	Pericardium morphogenesis	2	0.002
GO:0016043	Cellular component organization	9	0.002
GO:0016340	Calcium-dependent cell-matrix adhesion	2	0.002
GO:0090129	Positive regulation of synapse maturation	2	0.0022
GO:0007155	Cell adhesion	4	0.0024
GO:0006930	Substrate-dependent cell migration, cell extension	2	0.0031
GO:0007419	Ventral cord development	3	0.0031
GO:0007508	Larval heart development	2	0.0031
GO:0007229	Integrin-mediated signaling pathway	2	0.0035
GO:0060560	Developmental growth involved in morphogenesis	3	0.0058
GO:0007399	Nervous system development	6	0.0062

GO:0048731	System development	7	0.0065
GO:0030182	Neuron differentiation	5	0.0088
GO:0042330	Taxis	4	0.0088
GO:0051492	Regulation of stress fiber assembly	2	0.0088
GO:0021551	Central nervous system morphogenesis	2	0.0096
GO:0034446	Substrate adhesion-dependent cell spreading	2	0.0096
GO:0007157	Heterophilic cell-cell adhesion via plasma membrane cell2 adhesion molecules		0.0153
GO:0016339	Calcium-dependent cell-cell adhesion via plasma2 membrane cell adhesion molecules		0.0153
GO:0048808	Male genitalia morphogenesis	2	0.0153
GO:0007015	Actin filament organization	3	0.0163
GO:0030336	Negative regulation of cell migration	2	0.0163
GO:0071711	Basement membrane organization	2	0.0173
GO:0033627	Cell adhesion mediated by integrin	2	0.018
GO:0035162	Embryonic hemopoiesis	2	0.018
GO:0010669	Epithelial structure maintenance	2	0.0187
GO:0007494	Midgut development	2	0.0215
GO:0007444	Imaginal disc development	4	0.0217
GO:0048666	Neuron development	4	0.0292
GO:0009987	Cellular process	11	0.0321
GO:0045214	Sarcomere organization	2	0.0324
GO:0060541	Respiratory system development	3	0.034
GO:0016203	Muscle attachment	2	0.036
GO:0048513	Animal organ development	5	0.0438
GO:0007411	Axon guidance	3	0.0486
GO:0007608	Sensory perception of smell	2	0.0486

4.5 Culturas agrícolas e plantas silvestres (*Zea mays* e *Sorghum bicolor*)

O enriquecimento metabólico para o cenário #1 de *Zea mays* sugeriu que oito processos biológicos poderiam ser afetados, incluindo vias necessárias para a biossíntese lipídica. A biossíntese lipídica atua em funções celulares como reconhecimento, adesão celular, proliferação e diferenciação. As modificações previstas para o cenário #3 do *Sorghum bicolor* incluem vias para a proteólise da ubiquitina, envolvidas no processo celular, e vias associadas ao processo de desenvolvimento envolvido na reprodução e no desenvolvimento embrionário (Tabela 11).

Tabela 11. ID das vias metabólicas; descrição do processo; contagem genética observada; e taxa de falsa descoberta para *Zea mays* e *Sorghum bicolor*.

ID metabolismo	Descrição do metabolismo	contagem genética observada	taxa de falsa descoberta
----------------	--------------------------	-----------------------------	--------------------------

zma00603	Glycosphingolipid biosynthesis - globo and isoglobo6 series		2.19E-18
zma00604	Glycosphingolipid biosynthesis - ganglio series	5	8.08E-16
zma00531	Glycosaminoglycan degradation	5	2.37E-14
zma00511	Other glycan degradation	5	2.94E-14
zma00513	Various types of N-glycan biosynthesis	5	6.72E-13
zma00520	Amino sugar and nucleotide sugar metabolism	5	3.79E-10
zma01100	Metabolic pathways	6	6.17E-06
GO:0005975	Carbohydrate metabolic process	5	0.0051
<i>Sorghum bicolor</i>			
GO:0016925	Protein sumoylation	4	5.08E-08
sbi04120	Ubiquitin mediated proteolysis	3	0.00027
sbi03013	RNA transport	3	0.00027
GO:0003006	Developmental process involved in reproduction	4	0.0141
GO:0009793	Embryo development ending in seed dormancy	3	0.0141
GO:0009737	Response to abscisic acid	3	0.0178

A questão central deste trabalho foi a identificação dos perigos que podem surgir do uso de técnicas de edição genética fora das instalações laboratoriais, não sendo restritas a uma única espécie. Aqui, argumenta-se que, se as ferramentas *in silico* são eficazes na previsão das sequências de DNA alvo onde ocorrerão as reações de edição, então elas também são úteis para prever efeitos *off-target* quando organismos, seja de maneira intencional ou não, são expostos. Foi avaliada a hipótese de que a aplicação do produto formulado CRISPR/Cas9 em ambiente ao ar livre pode impactar outras espécies em três cenários. O uso dessas ferramentas em ambientes não controlados tem o potencial de elevar a edição do genoma à classe de contaminantes ambientais emergentes.

Os ONA selecionados foram identificados como objetivos de proteção suscetíveis de serem expostos. A partir do potencial previsto para hibridização entre um gRNA e sequências complementares de nucleotídeos no genoma de um ONA, muitos genes *off-target* poderiam ser afetados (Tabelas 3-6). Historicamente, os ONAs não têm sido uma fonte importante de risco porque a engenharia genética é realizada no organismo pretendido em laboratório, que minimiza o potencial de exposição dos ONAs aos procedimentos de modificação genética. Nem a relação do alvo com o ONA, nem o reino biológico do ONA eram preditivos de atividade *off-target*. Portanto, todas as espécies que despertam preocupação devem ser examinadas numa avaliação de risco. A atividade *off-target* pode resultar em efeitos biológicos em todas as plantas e animais não-alvo examinados, como os efeitos nas respostas imunológicas, na biossíntese de moléculas essenciais e no sistema

nervoso central. Estes resultados, consistentes com a literatura existente sobre técnicas de edição, fornecem uma justificativa para a necessidade de avaliar ONAs vulneráveis em qualquer uso proposto de *spray* ou técnicas tópicas no ambiente. A desregulamentação dos procedimentos, conforme proposto em estruturas de risco emergentes (HEINEMANN, 2023; COMISSAO EUROPEIA, 2023), tornaria possível utilizá-los em ambientes que não controlam exposições ou liberação de organismos modificados após uma exposição não intencional.

O uso aplicado externamente de reagentes que facilitam a penetração de células para a entrega de reagentes de edição genética ou silenciamento de genes está ganhando atenção devido ao seu potencial para melhorar o manejo de pragas, reduzir a toxicidade das exposições associadas a pesticidas convencionais e fornecer benefícios aos consumidores (HEINEMANN, 2019). O relatório descrito pelo Grupo de Especialistas Técnicos em Biologia Sintética (AHTEG, 2019) identificou agentes tópicos de penetração para uso na edição do genoma ou no silenciamento de genes (RNAi), entre várias novas aplicações da biologia sintética. Devido aos rápidos avanços no campo, os futuros organismos da biologia sintética podem ficar fora da definição existente de “organismo vivo modificado”, conforme descrito no Protocolo. Isto levanta preocupações à avaliação e monitoramento de tais organismos e destaca a necessidade de definições de avaliação de risco de OGM fluidas e em evolução (LI et al., 2021; HEINEMANN et al., 2023).

As previsões *in silico* apresentadas neste estudo fornecem uma visão geral dos efeitos potenciais em organismos expostos durante o uso de reagentes aplicados topicamente para causar edição do genoma ou silenciamento de genes. No entanto, estas listas não são exaustivas por alguns motivos. Primeiramente, as análises baseiam-se apenas em organismos com genomas sequenciados e anotados em bancos de dados disponíveis publicamente. Apenas alguns dos muitos objetivos de proteção que habitariam os locais onde essas técnicas seriam utilizadas estavam representados nas bases de dados. Pequenos desvios na atividade das espécies alvo podem causar efeitos em cascata. Se estas espécies estiverem sub-representadas, a avaliação do perigo fica comprometida. Como, por exemplo, o papel das minhocas nas pastagens. As minhocas contribuem para serviços ecossistêmicos, como a ciclagem de nutrientes, a formação da estrutura do solo, a melhoria da qualidade do solo e a regulação da água (BLOUIN et al. 2013; FONTE; HSIEH; MUELLER, 2023). Na Nova Zelândia, que possui um agroecossistema pecuário baseado em pastagens, a atividade das minhocas é avaliada em US\$ 365-440/ha anualmente (SCHON; GRAY; MACKAY, 2016).

Uma redução na atividade não precisa resultar da perda de um gene específico, mas simplesmente do fardo da mutagênese (MILLER et. al., 1999). As alterações na atividade das minhocas à microescala podem traduzir-se em grandes efeitos à macroescala, particularmente no gado dentro do agroecossistema.

Além disso, os organismos alvo e não-alvo considerados neste estudo limitaram-se aos eucariontes. Se a análise expandir para incluir os procariontes, então seria necessário considerar uma maior profundidade e amplitude dos objetivos de proteção. Em terceiro lugar, os efeitos da tecnologia de edição genética utilizada em ambientes não controlados não se limitam a alterações não intencionais nos ONA decorrentes da atividade pretendida. Um gRNA projetado para uma reação de edição genética pode, após liberação no ambiente e absorção pelas células alvo ou ONAs, tornar-se um reagente para um RNAi não intencional ou reação de edição genética (SHARMA et al., 2022).

Conclusões

Com base no uso das ferramentas *in silico* disponíveis, essa análise previu que o uso da edição do genoma em um contexto em que o organismo alvo é modificado *in situ* no ambiente, poderá resultar em mudanças não intencionais nos genomas dos organismos não-alvos. Estes organismos podem desempenhar funções ecológicas importantes e essenciais ou podem ser organismos indesejáveis. Quer os ONAs sejam desejados ou não, as consequências da sua modificação permanecem imprevisíveis devido ao grande número de modificações *off-target*. Para as utilizações propostas dos reagentes da engenharia genética fora dos laboratórios confinados, são necessárias novas estruturas de avaliação de risco (HEINEMANN, 2019) e que incluam a avaliação da sua persistência no ambiente. Esse trabalho justifica a inclusão dessas análises em instrumentos legais que prevejam o risco e o impacto dos reagentes da tecnologia de edição genética como ameaças à saúde humana e ambiental.

Considerações Finais

Esse trabalho foi financiado pelo Research Council of Norway (301911) e apoiado pelo Norwegian Research Centre (NORCE). Além disso, é parte de um artigo desenvolvido em conjunto com os seguintes pesquisadores: Jack A. Heinemann e Tessa Hiscox, da University of Canterbury, Sarah Zanon Agapito-Tenfen e Philomena Chu, do NORCE, e Caroline Bedin Zanatta, da Universidade Federal de Santa Catarina.

Referências

- AGARWAL, A.; AGASHE, D. The red flour beetle *Tribolium castaneum*: A model for host-microbiome interactions. **PLOS ONE**, v. 15, n. 10, p. e0239051, 2 out. 2020.
- AHTEG; REPORT OF THE AD HOC TECHNICAL EXPERT GROUP ON SYNTHETIC BIOLOGY MONTREAL. **Convention on Biological Diversity**. Canadá, 2019.
- ALKAN, F. et al. CRISPR-Cas9 off-targeting assessment with nucleic acid duplex energy parameters. **Genome Biology**, v. 19, n. 1, 26 out. 2018.
- ALKOORANEE, J. T. et al. Detecting the Hormonal Pathways in Oilseed Rape behind Induced Systemic Resistance by *Trichoderma harzianum* TH12 to *Sclerotinia sclerotiorum*. **PLOS ONE**, v. 12, n. 1, p. e0168850, 3 jan. 2017.
- BACHMAN, P. M. et al. Ecological risk assessment for DvSnf7 RNA: A plant-incorporated protectant with targeted activity against western corn rootworm. **Regulatory Toxicology and Pharmacology**, v. 81, p. 77–88, 1 nov. 2016.
- BAE, S. et al. Microhomology-based choice of Cas9 nuclease target sites. **Nature Methods**, v. 11, n. 7, p. 705–706, 27 jun. 2014.
- BLOUIN, M. et al. A review of earthworm impact on soil function and ecosystem services. **European Journal of Soil Science**, v. 64, n. 2, p. 161–182, 21 fev. 2013.
- BOLAND, G. J.; HALL, R. Index of plant hosts of *Sclerotinia sclerotiorum*. **Canadian Journal of Plant Pathology**, v. 16, n. 2, p. 93–108, jun. 1994.
- BOLOGNESI, R. et al. Characterizing the Mechanism of Action of Double-Stranded RNA Activity against Western Corn Rootworm (*Diabrotica virgifera virgifera* LeConte). **PLoS ONE**, v. 7, n. 10, p. e47534, 11 out. 2012
- BORTESI, L.; FISCHER, R. The CRISPR/Cas9 system for plant genome editing and beyond. **Biotechnology Advances**, v. 33, n. 1, p. 41–52, jan. 2015.
- CAMPBELL, J. F. et al. *Tribolium castaneum*: A Model Insect for Fundamental and Applied Research. **Annual Review of Entomology**, v. 67, n. 1, p. 347–365, 7 jan. 2022.
- CONCORDET, J.-P.; HAEUSSLER, M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. **Nucleic Acids Research**, v. 46, n. W1, p. W242–W245, 14 maio 2018.
- CUI, Y. et al. Review of CRISPR/Cas9 sgRNA Design Tools. **Interdisciplinary sciences, computational life sciences**, v. 10, n. 2, p. 455–465, 2018.
- DEMIRER, G. S. et al. Nanotechnology to advance CRISPR–Cas genetic engineering of plants. **Nature Nanotechnology**, v. 16, n. 3, p. 243–250, 1 mar. 2021.

COMISSÃO EUROPEIA. Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on plants obtained by certain new genomic techniques and their food and feed, and amending Regulation (EU) 2017/625. **Parlamento Europeu**. 2023

FONTE, S. J.; HSIEH, M.; MUELLER, N. D. Earthworms contribute significantly to global food production. **Nature Communications**, v. 14, n. 1, p. 5713, 26 set. 2023.

HANNING, J. E. et al. Lack of correlation between predicted and actual off-target effects of short-interfering RNAs targeting the human papillomavirus type 16 E7 oncogene. **British Journal of Cancer**, v. 108, n. 2, p. 450–460, 1 fev. 2013.

HEINEMANN, J. A. Should dsRNA treatments applied in outdoor environments be regulated? **Environment International**, v. 132, p. 104856, nov. 2019.

HEINEMANN, J. Submission on information and supporting documentation relevant to the trends in new technological developments in synthetic biology and any additional trend that should be considered. **ir.canterbury.ac.nz**, 2023.

HEINEMANN, J. A. et al. Are null segregants new combinations of heritable material and should they be regulated? **Frontiers in genome editing**, v. 4, 10 jan. 2023.

HEINEMANN, J. A.; WALKER, S. Environmentally applied nucleic acids and proteins for purposes of engineering changes to genes and other genetic material. **Biosafety and Health**, v. 1, n. 3, p. 113–123, 1 dez. 2019.

HÖIJER, I. et al. CRISPR-Cas9 induces large structural variants at on-target and off-target sites in vivo that segregate across generations. **Nature Communications**, v. 13, n. 1, 2 fev. 2022.

HSU, P. D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. **Nature Biotechnology**, v. 31, n. 9, p. 827–832, 21 jul. 2013.

KANEHISA, M. et al. KEGG: integrating viruses and cellular organisms. **Nucleic Acids Research**, v. 49, n. D1, p. D545–D551, 30 out. 2020.

Interim Regulatory Impact Statement: Improving our GMO regulations for laboratory and biomedical research | The Treasury New Zealand. Disponível em: <<https://www.treasury.govt.nz/publications/risa/interim-regulatory-impact-statement-improving-our-gmo-regulations-laboratory-and-biomedical-research>>. Acesso em: 29 abr. 2024..

- KNORR, E. et al. Gene silencing in *Tribolium castaneum* as a tool for the targeted identification of candidate RNAi targets in crop pests. **Scientific Reports**, v. 8, n. 1, 1 fev. 2018.
- LI, J. et al. Advances in Synthetic Biology and Biosafety Governance. **Frontiers in Bioengineering and Biotechnology**, v. 9, 30 abr. 2021.
- MEINKE, L. J. et al. Western corn rootworm (*Diabrotica virgifera virgifera*LeConte) population dynamics. **Agricultural and Forest Entomology**, v. 11, n. 1, p. 29–46, 1 fev. 2009.
- MEINKE, L. J.; SOUZA, D.; SIEGFRIED, B. D. The Use of Insecticides to Manage the Western Corn Rootworm, *Diabrotica virgifera virgifera*, LeConte: History, Field-Evolved Resistance, and Associated Mechanisms. **Insects**, v. 12, n. 2, p. 112, 28 jan. 2021.
- MILLER, J. H. et al. Direct Selection for Mutators in *Escherichia coli*. **Journal of Bacteriology**, v. 181, n. 5, p. 1576–1584, mar. 1999.
- MILLER, N. et al. Multiple Transatlantic Introductions of the Western Corn Rootworm. **Science**, v. 310, n. 5750, p. 992–992, 11 nov. 2005.
- MUELLER, S. Facing the 2020 Pandemic: What does Cyberbiosecurity want us to know to safeguard the future? **Biosafety and Health**, set. 2020.
- PARK, J.; BAE, S.; KIM, J.-S. Cas-Designer: a web-based tool for choice of CRISPR-Cas9 target sites. **Bioinformatics**, p. btv537, 10 set. 2015.
- RANGWALA, S. H. et al. Accessing NCBI data using the NCBI Sequence Viewer and Genome Data Viewer (GDV). **Genome Research**, p. gr.266932.120, 25 nov. 2020.
- SAPPINGTON, T. W.; SIEGFRIED, B. D.; GUILLEMAUD, T. Coordinated *Diabrotica* Genetics Research: Accelerating Progress on an Urgent Insect Pest Problem. **American Entomologist**, v. 52, n. 2, p. 90–97, 1 abr. 2006
- SCHON, N. L.; GRAY, R. A.; MACKAY, A. D. Earthworms stimulate pasture production in sheep and beef systems: their economic value. **Journal of New Zealand Grasslands**, v. 78, p. 89–92, 1 jan. 2016.
- SEGAL, D. J.; MECKLER, J. F. Genome Engineering at the Dawn of the Golden Age. **Annual Review of Genomics and Human Genetics**, v. 14, n. 1, p. 135–158, 31 ago. 2013.
- SHARMA, V. K. et al. CRISPR guides induce gene silencing in plants in the absence of Cas. **Genome Biology**, v. 23, n. 1, 3 jan. 2022.
- SHI, J. et al. ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. **Plant biotechnology journal**, v. 15, n. 2, p. 207–216, 2017.

SZKLARCZYK, D. et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. **Nucleic Acids Research**, v. 51, n. D1, p. D638–D646, 12 nov. 2022.

SMITH, R. H. et al. Genome-Wide Analysis of Off-Target CRISPR/Cas9 Activity in Single-Cell-Derived Human Hematopoietic Stem and Progenitor Cell Clones. **Genes**, v. 11, n. 12, p. 1501, 13 dez. 2020.

TAO, J.; BAUER, D. E.; CHIARLE, R. Assessing and advancing the safety of CRISPR-Cas tools: from DNA to RNA editing. **Nature Communications**, v. 14, n. 1, p. 212, 13 jan. 2023.

VATS, S. et al. Opportunity and challenges for nanotechnology application for genome editing in plants. **Plant Nano Biology**. v.1, n. 10001, 2 jun. 2022

WANG, S. et al. Efficient targeted mutagenesis in potato by the CRISPR/Cas9 system. **Plant Cell Reports**, v. 34, n. 9, p. 1473–1476, 17 jun. 2015.

WYTINCK, N. et al. Host induced gene silencing of the *Sclerotinia sclerotiorum* ABHYDROLASE-3 gene reduces disease severity in *Brassica napus*. **PLOS ONE**, v. 17, n. 8, p. e0261102–e0261102, 26 ago. 2022.

ZHANG, F.; WEN, Y.; GUO, X. CRISPR/Cas9 for genome editing: progress, implications and challenges. **Human Molecular Genetics**, v. 23, n. R1, p. R40–R46, 20 mar. 2014.