



República Federativa do Brasil
Ministério da Economia
Instituto Nacional da Propriedade Industrial

(11) BR 112015023489-5 B1



(22) Data do Depósito: 14/03/2014

(45) Data de Concessão: 07/06/2022

(54) Título: MÉTODOS PARA AUMENTAR A ESPECIFICIDADE DE EDIÇÃO DE GENOMA ORIENTADO POR RNA EM UMA CÉLULA, DE INDUÇÃO DE UMA RUPTURA EM UMA REGIÃO ALVO DE UMA MOLÉCULA DE DNA DE FITA DUPLA EM UMA CÉLULA E DE MODIFICAÇÃO DE UMA REGIÃO ALVO DE UMA MOLÉCULA DE DNA DE FITA DUPLA EM UMA CÉLULA

(51) Int.Cl.: C12Q 1/68; C12N 15/00; C07H 21/04.

(30) Prioridade Unionista: 15/03/2013 US 61/799,647; 21/06/2013 US 61/838,178; 21/06/2013 US 61/838,148; 26/12/2013 US 61/921,007.

(73) Titular(es): THE GENERAL HOSPITAL CORPORATION.

(72) Inventor(es): J. KEITH JOUNG; JEFFRY D. SANDER; YANFANG FU; MORGAN MAEDER.

(86) Pedido PCT: PCT US2014029068 de 14/03/2014

(87) Publicação PCT: WO 2014/144592 de 18/09/2014

(85) Data do Início da Fase Nacional: 15/09/2015

(57) Resumo: USO DE RNAs GUIA TRUNCADOS (tru-gRNAs) PARA AUMENTAR A ESPECIFICIDADE DE EDIÇÃO DE GENOMA COM RNA GUIADO. Métodos para aumentar a especificidade de edição do genoma orientada por RNA, por exemplo, edição usando sistemas CRISPR/Cas9, usando RNAs guia truncados (tru-gRNAs).

Relatório Descritivo da Patente de Invenção para
"MÉTODOS PARA AUMENTAR A ESPECIFICIDADE DE EDIÇÃO DE GENOMA ORIENTADO POR RNA EM UMA CÉLULA, DE INDUÇÃO DE UMA RUPTURA EM UMA REGIÃO ALVO DE UMA MOLÉCULA DE DNA DE FITA DUPLA EM UMA CÉLULA E DE MODIFICAÇÃO DE UMA REGIÃO ALVO DE UMA MOLÉCULA DE DNA DE FITA DUPLA EM UMA CÉLULA".

REIVINDICAÇÃO DE PRIORIDADE

[0001] O presente pedido reivindica o benefício dos Pedidos de Patente dos Estados Unidos N^{os} de Série 61/799.647, depositado em 15 de Março de 2013; 61/838.178, depositado em 21 de Junho de 2013; 61/838.148, depositado em 21 de Junho de 2013 e 61/921.007, depositado em 26 de Dezembro de 2013. O conteúdo total dos precedentes é aqui incorporado por referência.

PESQUISA OU DESENVOLVIMENTO PATROCINADO PELO GOVERNO FEDERAL

[0002] A presente invenção foi feita com o apoio do Governo sob a Cessão N^o DP1 GM105378 concedida pelo National Institutes of Health. O Governo tem certos direitos sobre a invenção.

CAMPO TÉCNICO

[0003] Métodos para aumentar a especificidade de edição de genoma orientada por RNA, por exemplo, edição usando sistemas CRISPR/Cas9, usando RNAs guia truncados (TRU-gRNAs).

ANTECEDENTES

[0004] Trabalhos recentes têm demonstrado que sistemas de repetições palindrômicas curtas, agrupadas, regularmente intercaladas (CRISPR)/associados a CRISPR (Cas) (Wiedenheft *et al.*, Nature 482, 331-338 (2012); Horvath *et al.*, Science 327, 167-170 (2010); Nhece *et al.*, Curr Opin Microbiol 14, 321-327 (2011)) podem servir como a base para realização de edição genoma em células de bactérias, leveduras e

humanas, bem como *in vivo* em organismos completos, tais como moscas da fruta, peixe-zebra e camundongos (Wang *et al.*, Cell 153, 910-918 (2013); Shen *et al.*, Cell Res (2013); Dicarlo *et al.*, Nucleic Acids Res (2013); Jiang *et al.*, Nat. Biotechnol 31, 233-239 (2013); Jinek *et al.*, Elife 2, e00471 (2013); Hwang *et al.*, Nat Biotechnol 31, 227-229 (2013); Cong *et al.*, Science 339, 819-823 (2013); Mali *et al.*, Science 339, 823-826 (2013c); Cho *et al.*, Nat Biotechnol 31, 230-232 (2013); Gratz *et al.*, Genetics 194 (4): 1029-1035 (2013)). A nuclease Cas9 de *S. pyogenes* (daqui em diante simplesmente Cas9) pode ser orientada através de complementaridade de pares de bases entre os 20 primeiros nucleotídeos de RNA de um RNA guia manipulado (gRNA) e a fita complementar de uma sequência de DNA genômico alvo de interesse que se encontra após um motivo protoespaçador adjacente (**Protospacer Adjacent Motif** - PAM), por exemplo, um PAM que corresponde à sequência NGG ou NAG (Shen *et al.*, Cell Res (2013); Dicarlo *et al.*, Nucleic Acids Res (2013); Jiang *et al.*, Nat Biotechnol 31, 233 -239 (2013); Jinek *et al.*, Elife 2, e00471 (2013); Hwang *et al.*, Nat Biotechnol 31, 227-229 (2013); Cong *et al.*, Science 339, 819-823 (2013); Mali *et al.*, Science 339, 823-826 (2013c); Cho *et al.*, Nat Biotechnol 31, 230-232 (2013); Jinek *et al.*, Science 337, 816-821 (2012)). Estudos anteriores realizados *in vitro* (Jinek *et al.*, Science 337, 816-821 (2012)) em células de bactérias (Jiang *et al.*, Nat Biotechnol 31, 233-239 (2013)) e em células humanas (Cong *et al.*, Science 339, 819-823 (2013)) demonstraram que a clivagem mediada por Cas9 pode, em alguns casos, ser abolida por desemparelhamentos individuais na interface gRNA/sítio alvo, particularmente nos últimos 10-12 nucleotídeos (nts) localizados na extremidade 3' da região de complementaridade de gRNA de 20 nt.

SUMÁRIO

[0005] A edição de genoma por CRISPR-Cas usa RNA guia, o qual

inclui tanto uma região de complementaridade (a qual se liga ao DNA alvo através de pareamento de bases) quanto uma região de ligação a Cas9, para dirigir uma nuclease Cas9 a um DNA alvo (vide Figura 1). A nuclease pode tolerar um determinado número de desemparelhamentos (até cinco, conforme mostrado aqui) na região de complementaridade e ainda clivar; é difícil prever os efeitos de um desemparelhamento individual ou uma combinação de desemparelhamentos sobre a atividade. Tomadas em conjunto, estas nucleases podem mostrar efeitos fora do alvo significativos, mas pode ser um desafio prever estes sítios. São descritos aqui métodos para aumentar a especificidade de edição de genoma usando o sistema CRISPR/Cas, por exemplo, usando Cas9 ou proteínas de fusão com base em Cas9. Em particular, são fornecidos RNAs guia truncados (tru-gRNAs) que incluem uma região de complementaridade ao alvo reduzida (ou seja, menos de 20 nts, por exemplo, 17-19 ou 17-18 nts de complementaridade ao alvo, por exemplo, 17, 18 ou 19 nts de complementaridade ao alvo), e métodos de uso dos mesmos. Conforme usado aqui, "17-18 ou 17-19" inclui 17, 18 ou 19 nucleotídeos.

[0006] Em um aspecto, a invenção proporciona uma molécula de RNA de guia (por exemplo, um RNA guia individual ou um crRNA) tendo uma região de complementaridade ao alvo de 17-18 ou 17-19 nucleotídeos, por exemplo, a região de complementaridade ao alvo consiste em 17-18 ou 17-19 nucleotídeos, por exemplo, a região de complementaridade ao alvo consiste em 17-18 ou 17-19 nucleotídeos de complementaridade ao alvo consecutivos. Em algumas modalidades, o RNA guia inclui uma região de complementaridade que consiste em 17-18 ou 17-19 nucleotídeos que são complementares a 17-18 ou 17-19 nucleotídeos consecutivos da fita complementar de uma sequência genômica alvo selecionada. Em algumas modalidades, a região de complementaridade ao alvo consiste em 17-18 nucleotídeos (de

complementaridade ao alvo). Em algumas modalidades, a região de complementaridade é complementar a 17 nucleotídeos consecutivos da fita complementar de uma sequência alvo selecionada. Em algumas modalidades, a região de complementaridade é complementar a 18 nucleotídeos consecutivos da fita complementar de uma sequência alvo selecionada.

[0007] Em outro aspecto, a invenção proporciona um ácido ribonucleico que consiste na sequência:

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUA (SEQ ID NO: 2404);

(X₁₇₋₁₈ ou X₁₇₋₁₉) GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 2407); ou

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCU (SEQ ID NO: 2408);

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUC CG(X_N) (SEQ ID NO: 1);

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUC(X_N) (SEQ ID NO: 2);

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGC AAGUUAAAAUAAGGCUAGUCCGUUAUC(X_N) (SEQ ID NO: 3);

(X₁₇₋₁₈)GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAG GCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(X_N) (SEQ ID NO: 4),

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUC CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID NO: 5);

(X₁₇₋₁₈ ou

X₁₇₋₁₉)GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAU
AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGGUCGGU
GC (SEQ ID NO: 6); ou

(X₁₇₋₁₈ ou

X₁₇₋₁₉)GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAU
AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG
UGC (SEQ ID NO: 7);

em que X₁₇₋₁₈ ou X₁₇₋₁₉ é uma sequência (de 17-18 ou 17-19 nucleotídeos) complementar à fita complementar de uma sequência alvo selecionada, de preferência uma sequência alvo imediatamente 5' de um motivo protoespaçador adjacente (**Protospacer Adjacent Motif** - PAM), por exemplo, NGG, NAG ou NNGG (vide, por exemplo, a configuração na Figura 1) e X_N é qualquer sequência, em que N (no RNA) pode ser 0-200, por exemplo, 0-100, 0-50 ou 0-20, que não interfira com a ligação do ácido ribonucleico à Cas9. Em nenhum caso, X₁₇₋₁₈ ou X₁₇₋₁₉ é idêntica a uma sequência que ocorre naturalmente adjacente ao resto do RNA. Em algumas modalidades, o RNA inclui um ou mais U, por exemplo, 1 a 8 ou mais Us (por exemplo, U, UU, UUU, UUUU, UUUUU, UUUUUU, UUUUUUU, UUUUUUUU, UUUUUUUUU) na extremidade 3' da molécula, como um resultado da presença opcional de um ou mais Ts usados como um sinal de término para terminar a transcrição de RNA PolIII. Em algumas modalidades, o RNA inclui um ou mais, por exemplo, até 3, por exemplo, um, dois ou três, nucleotídeos adicionais na extremidade 5' da molécula de RNA que não são complementares à sequência alvo. Em algumas modalidades, a região de complementaridade ao alvo consiste em 17-18 nucleotídeos (de complementaridade ao alvo). Em algumas modalidades, a região de complementaridade é complementar a 17 nucleotídeos consecutivos da fita complementar de uma sequência alvo selecionada. Em algumas modalidades, a região de complementaridade é complementar a 18

nucleotídeos consecutivos.

[0008] Em outro aspecto, a invenção fornece moléculas de DNA que codificam os ácidos ribonucleicos descritos aqui e células hospedeiras que abrigam ou expressam os ácidos ribonucleicos ou vetores.

[0009] Em um aspecto adicional, a invenção proporciona métodos para aumentar a especificidade de edição de genoma orientada por RNA em uma célula, o método compreendendo contato da célula com um RNA guia que inclui uma região de complementaridade que consiste em 17-18 ou 17-19 nucleotídeos que são complementares a 17-18 ou 17-19 nucleotídeos consecutivos da fita complementar de uma sequência genômica alvo selecionada, conforme descrito aqui.

[0010] Em ainda outro aspecto, a invenção proporciona métodos para indução de uma ruptura de fita simples ou dupla em uma região alvo de uma molécula de DNA fita dupla, por exemplo, em uma sequência genômica em uma célula. Os métodos incluem expressão em ou introdução, na célula, de:

uma nuclease Cas9 ou nickase; e

um RNA guia que inclui uma sequência que consiste em 17 ou 18 ou 19 nucleotídeos que são complementares à fita complementar de uma sequência alvo selecionada, de preferência uma sequência alvo imediatamente 5' de um motivo protoespaçador adjacente (**P**rotospacer **A**djacent **M**otif - PAM), por exemplo, NGG, NAG ou NNGG, por exemplo, um ácido ribonucleico conforme descrito aqui.

[0011] Também são fornecidos aqui métodos para modificação de uma região alvo de uma molécula de DNA fita dupla em uma célula. Os métodos incluem expressão ou introdução, na célula, de:

uma proteína de fusão com domínio funcional dCas9-heterólogo (dCas9-HFD); e

um RNA guia que inclui uma região de complementaridade

que consiste em 17-18 ou 17-19 nucleotídeos que são complementares a 17-18 ou 17-19 nucleotídeos consecutivos da fita complementar de uma sequência genômica alvo selecionada, conforme descrito aqui.

[0012] Em algumas modalidades, o RNA guia é (i) um RNA guia individual que inclui uma região de complementaridade que consiste em 17-18 ou 17-19 nucleotídeos que são complementares a 17-18 ou 17-19 nucleotídeos consecutivos da fita complementar de uma sequência genômica alvo selecionada ou (ii) um crRNA que inclui uma região de complementaridade que consiste em 17-18 ou 17-19 nucleotídeos que são complementares a 17-18 ou 17-19 nucleotídeos consecutivos da fita complementar de uma sequência genômica alvo selecionada, e um tracrRNA.

[0013] Em algumas modalidades, a região de complementaridade ao alvo consiste em 17-18 nucleotídeos (de complementaridade ao alvo). Em algumas modalidades, a região de complementaridade é complementar a 17 nucleotídeos consecutivos da fita complementar de uma sequência alvo selecionada. Em algumas modalidades, a região de complementaridade é complementar a 18 nucleotídeos consecutivos.

[0014] Em nenhum caso, X_{17-18} ou X_{17-19} de qualquer das moléculas descritas aqui é idêntica a uma sequência que ocorre naturalmente adjacente ao resto do RNA. Em algumas modalidades, o RNA inclui um ou mais U, por exemplo, 1 a 8 ou mais Us (por exemplo, U, UU, UUU, UUUU, UUUUU, UUUUUU, UUUUUUU, UUUUUUUU, UUUUUUUUU) na extremidade 3' da molécula, como um resultado da presença opcional de um ou mais Ts usados como um sinal de término para terminar a transcrição de RNA PolIII. Em algumas modalidades, o RNA inclui um ou mais, por exemplo, até 3, por exemplo, um, dois, ou três, nucleotídeos adicionais na extremidade 5' da molécula de RNA que não são complementares à sequência alvo.

[0015] Em algumas modalidades, um ou mais dos nucleotídeos do RNA são modificados, por exemplo, bloqueados (ponte de metileno 2'-O-4'-C), é 5'-metilcitidina, é 2'-O-metil-pseudouridina ou no qual a estrutura de fosfato de ribose foi substituída por uma cadeia de poliamida, por exemplo, um ou mais dos nucleotídeos dentro ou fora da região de complementaridade ao alvo X_{17-18} ou X_{17-19} . Em algumas modalidades, alguns ou todos os tracrRNA ou crRNA, por exemplo, dentro ou fora da região de complementaridade ao alvo X_{17-18} ou X_{17-19} , compreende desoxirribonucleotídeos (por exemplo, é DNA total ou parcial, por exemplo, híbridos de DNA/RNA).

[0016] Em um aspecto adicional, a invenção proporciona métodos para modificação de uma região alvo de uma molécula de DNA fita dupla, por exemplo, em uma sequência genômica em uma célula. Os métodos incluem expressão ou introdução, na célula, de:

uma proteína de fusão com domínio funcional dCas9-heterólogo (dCas9-HFD); e

um RNA guia que inclui uma sequência que consiste em 17-18 ou 17-19 nucleotídeos que são complementares à fita complementar de uma sequência alvo selecionada, de preferência, uma sequência alvo imediatamente 5' de um motivo protoespaçador adjacente (**P**roto**s**pacer **A**djacent **M**otif - PAM), por exemplo, NGG, NAG ou NNGG, por exemplo, um ácido ribonucleico conforme descrito aqui. Em nenhum caso X_{17-18} ou X_{17-19} é idêntica a uma sequência que ocorre naturalmente adjacente ao resto do RNA. Em algumas modalidades, o RNA inclui um ou mais, por exemplo, até 3, por exemplo, um, dois, ou três, nucleotídeos adicionais na extremidade 5' da molécula de RNA que não são complementares à sequência alvo.

[0017] Em outro aspecto, a invenção proporciona métodos para modificação, por exemplo, introdução de uma ruptura de sequência específica em uma região alvo de uma molécula de DNA fita dupla, por

exemplo, em uma sequência genômica em uma célula. Os métodos incluem expressão ou introdução, na célula, de:

uma nuclease Cas9 ou nickase, ou uma proteína de fusão com domínio funcional dCas9-heterólogo (dCas9-HFD);

um tracrRNA, por exemplo, que compreende ou consiste na sequência:

GGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGG
CUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC
(SEQ ID NO: 8) ou uma porção ativa da mesma;

UAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUG
AAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 2405) ou uma
porção ativa da mesma;

AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA
CUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 2407) ou
uma porção ativa da mesma;

CAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGU
UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:
2409) ou uma porção ativa da mesma;

UAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUG
AAAAAGUG (SEQ ID NO: 2410) ou uma porção ativa da mesma;

UAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCA (SEQ ID
NO: 2411) ou uma porção ativa da mesma; ou

UAGCAAGUUAAAAUAAGGCUAGUCCG (SEQ ID NO:
2412) ou uma porção ativa da mesma; e

um crRNA que inclui uma sequência que consiste em 17-18
ou 17-19 nucleotídeos que são complementares à fita complementar de
uma sequência alvo selecionada, de preferência uma sequência alvo
imediatamente 5' de um motivo protoespaçador adjacente (**Protospacer
Adjacent Motif - PAM**), por exemplo, NGG, NAG ou NNGG; em algumas
modalidades da crRNA tem a sequência:

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUA (SEQ ID NO: 2404);
 (X₁₇₋₁₈ ou X₁₇₋₁₉) GUUUUAGAGCUAUGCUGUUUUG (SEQ
 ID NO: 2407); ou
 (X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCU (SEQ ID NO:
 2408).

[0018] Em algumas modalidades, o crRNA é (X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 2407) e o tracrRNA é

GGAACCAUUCAAAACAGCAUAGCAAGUUAUUAAAGGCUAGUCC
 GUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:
 8); o cRNA é (X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUA (SEQ ID NO: 2404) e
 o tracrRNA é

UAGCAAGUUAUUAAAGGCUAGUCCGUUAUCAACUUGAAAAAGU
 GGCACCGAGUCGGUGC (SEQ ID NO: 2405); ou o cRNA é (X₁₇₋₁₈ ou
 X₁₇₋₁₉) GUUUUAGAGCUAUGCU (SEQ ID NO:2408) e o tracrRNA é
 AGCAUAGCAAGUUAUUAAAGGCUAGUCCGUUAUCAACUUGAAA
 AAGUGGCACCGAGUCGGUGC (SEQ ID NO: 2406).

[0019] Em nenhum caso, X₁₇₋₁₈ ou X₁₇₋₁₉ é idêntica a uma sequência que ocorre naturalmente adjacente ao resto do RNA. Em algumas modalidades, o RNA (por exemplo, tracrRNA ou crRNA) inclui um ou mais U, por exemplo, 2 a 8 ou mais Us (por exemplo, U, UU, UUU, UUUU, UUUUU, UUUUUU, UUUUUUU, UUUUUUUU) na extremidade 3' da molécula como um resultado da presença opcional de um ou mais Ts usados como um sinal de término para terminar a transcrição de RNA PolIII. Em algumas modalidades, o RNA (por exemplo, tracrRNA ou crRNA) inclui um ou mais, por exemplo, até 3, por exemplo, um, dois ou três, nucleotídeos adicionais na extremidade 5' da molécula de RNA que não são complementares à sequência alvo. Em algumas modalidades, um ou mais dos nucleotídeos do crRNA ou tracrRNA são modificados, por exemplo, bloqueados (ponte de metileno

2'-O-4'-C), é 5'-metilcitidina, é 2'-O-metil-pseudouridina ou em que a estrutura de fosfato de ribose foi substituída por uma cadeia de poliamida, por exemplo, um ou mais dos nucleotídeos dentro ou fora da sequência X_{17-18} ou X_{17-19} . Em algumas modalidades, alguns ou todos dos tracrRNA ou crRNA, por exemplo, dentro ou fora da região de complementaridade ao alvo X_{17-18} ou X_{17-19} , compreendem desoxirribonucleotídeos (por exemplo, é total ou parcialmente DNA, por exemplo, híbridos de DNA/RNA).

[0020] Em algumas modalidades, a proteína de fusão com domínio funcional dCas9-heterólogo (dCas9-HFD) compreende um HFD que modifica a expressão do gene, histonas ou DNA, por exemplo, domínio de ativação de transcrição, repressor de transcrição (por exemplo, silenciadores, tais como Proteína 1 de Heterocromatina (HP1), por exemplo, HP1 α ou HP1 β), enzimas que modificam o estado de metilação do DNA (por exemplo, DNA metiltransferase (DNMT) ou proteínas TET, por exemplo, TET1) ou enzimas que modificam a subunidade histona (por exemplo, histona acetiltransferases (HAT), histona desacetilases (HDAC) ou histona desmetilases). Em modalidades preferidas, o domínio funcional heterólogo é um domínio de ativação transcricional, por exemplo, um domínio de ativação transcricional VP64 ou NF- κ B p65; uma enzima que catalisa a desmetilação de DNA, por exemplo, um membro da família de proteínas TET ou o domínio catalítico de um destes membros da família; ou modificação de histona (por exemplo, LSD1, histona metiltransferase, HDACs ou HATs) ou um domínio de silenciamento de transcrição, por exemplo, de uma Proteína 1 de Heterocromatina (HP1), por exemplo, HP1 α ou HP1 β ; ou um grupo de ancoragem, por exemplo, MS2, CRISPR/proteína 4 Cas Subtipo Ypest (Csy4) ou proteína lambda N. dCas9-HFD são descritos nos Pedidos Provisórios de Patente dos Estados Unidos USSN 61/799.647, depositado em 15 de Março de

2013, registro de procurador N° 00786-0882P02, USSN 61/838.148, depositado em 21 de Junho de 2013 e Publicação Internacional PCT N° PCT/US14/27335, todos os quais são aqui incorporados por referência na íntegra.

[0021] Em algumas modalidades, os métodos descritos aqui resultam em uma mutação indel ou alteração de sequência na sequência genômica alvo selecionada.

[0022] Em algumas modalidades, a célula é uma célula eucariota, por exemplo, uma célula de mamífero, por exemplo, uma célula humana.

[0023] Salvo definição em contrário, todos os termos técnicos e científicos usados aqui têm o mesmo significado conforme normalmente entendido por aqueles versados na técnica à qual a presente invenção pertence. Métodos e materiais são descritos aqui para uso na presente invenção; outros métodos e materiais adequados conhecidos na especialidade também podem ser usados. Os materiais, métodos e exemplos são apenas ilustrativos e não se destinam a ser limitativos. Todas as publicações, pedidos de patentes, patentes, sequências, entradas de dados e outras referências mencionadas aqui são incorporados por referência na íntegra. Em caso de conflito, o presente relatório descritivo, incluindo definições, prevalecerá.

[0024] Outras características e vantagens da invenção serão evidentes a partir da descrição detalhada a seguir e figuras, e a partir das reivindicações.

DESCRIÇÃO DOS DESENHOS

[0025] Figura 1: Diagrama esquemático que ilustra um complexo de gRNA/nuclease Cas9 ligado ao seu sítio de DNA alvo. Tesouras indicam pontos de clivagem aproximados da nuclease Cas9 no sítio alvo de DNA genômico. Note que a numeração de nucleotídeos sobre o RNA guia ocorre de forma inversa a partir de 5' para 3'.

[0026] Figura 2A: Diagrama esquemático que ilustra uma lógica para truncamento da região de complementaridade 5' de um gRNA. Linhas cinzas grossas = sítio de DNA alvo, estrutura em linha cinza escuro fina = gRNA, linhas pretas mostram pareamentos de bases (ou ausência dos mesmos) entre o gRNA e o sítio de DNA alvo.

[0027] Figura 2B: Visão geral esquemática do ensaio de ruptura EGFP. Reparo de rupturas de fita dupla mediada por Cas9 objetivada em um único gene repórter EGFP-PEST integrado por meio de reparo mediado por NHEJ propenso a erros leva a mutações de desvio de quadro que rupturam a sequência de codificação e perda associada de fluorescência em células.

[0028] Figuras 2C a F: Atividades de nucleases RNA-orientadas (RGNs) que abrigam RNAs guia individuais (gRNAs) trazendo (C) um desemparelhamento individual, (D) desemparelhamentos duplos adjacentes, (E) desemparelhamentos duplos de espaço variável e (F) números crescentes de desemparelhamentos adjacentes ensaiados sobre três sítios alvo diferentes na sequência do gene repórter EGFP. Atividades médias das réplicas são mostradas, normalizadas para a atividade de um gRNA individual perfeitamente pareado. As barras de erro indicam o erro padrão da média. Posições com desemparelhamento em cada gRNA individual são destacadas em cinza abaixo. As sequências dos três sítios alvo de EGFP foram como segue:

Sítio 1 de EGFP GGGCACGGGCAGCTTGCCGGTGG
(SEQ ID NO: 9)

Sítio 2 de EGFP GATGCCGTTCTTCTGCTTGTCGG
(SEQ ID NO: 10)

Sítio 3 de EGFP GGTGGTGCAGATGAACTTCAGGG
(SEQ ID NO: 11)

[0029] Figura 2G: Desemparelhamentos na extremidade 5' do

gRNA tornam CRISPR/Cas mais sensível a mais desemparelhamentos 3'. O par de bases de Watson-Crick de gRNAs entre o RNA e DNA com a exceção das posições indicadas com um "m", as quais tem desemparelhamentos com a transversão de Watson-Crick (isto é, Sítio #2 de EGFP, M18-19, tem desemparelhamento ao trocar o gRNA para seu parceiro de Watson-Crick nas posições 18 e 19. Embora as posições próximas de 5' do gRNA sejam geralmente muito bem toleradas, desemparelhamentos nestas posições são importantes para a atividade de nuclease quando outros resíduos são incompatíveis. Quando todas as quatro posições têm desemparelhamentos, a atividade de nuclease não é mais detectável. Isto demonstra ainda que desemparelhamentos nesta posição 5' podem ajudar a compensar os desemparelhamentos em outras posições mais 3'. Note que estes experimentos foram realizados com uma versão sem códons otimizados de Cas9 a qual pode mostrar menores níveis absolutos de atividade de nuclease quando comparado com a versão com códons otimizados.

[0030] Figura 2H: Eficiência de atividades nuclease Cas9 dirigida por gRNAs trazendo regiões de complementaridade de comprimento variável que vão de 15 a 25 nts em um ensaio de ruptura por EGFP U2OS com base em célula humana. A expressão de um gRNA a partir do promotor U6 requer a presença de uma G 5' e, portanto, só foi possível avaliar gRNAs trazendo determinados comprimentos de complementaridade com o sítio de DNA alvo (15, 17, 19, 20, 21, 23 e 25 nts).

[0031] Figura 3A: Eficiências de ruptura por EGFP em células humanas mediada por Cas9 e gRNAs de comprimento completo ou reduzidos para quatro sítios alvo no gene repórter *EGFP*. Comprimentos de regiões de complementaridade e sítios de DNA alvo correspondentes são mostrados. Ctrl = gRNA de controle carecendo de uma região de complementaridade.

[0032] Figura 3B: Eficiências de mutações indel objetivadas introduzidas em sete genes alvos endógenos humanos diferentes por RGNs padrões pareadas (Cas9 e gRNAs de comprimento completo padrões) e tru-RGNs (Cas9 e gRNAs trazendo truncamentos em suas regiões de complementaridade 5'). Comprimentos de regiões de complementaridade de gRNA e sítios de DNA alvo correspondentes são mostrados. Frequências de indel foram medidas por meio do ensaio de T7EI. Ctrl = gRNA de controle carecendo de uma região de complementaridade.

[0033] Figura 3C: Sequências de DNA de mutações indel induzidas por RGNs usando um tru-gRNA ou um gRNA de comprimento completo pareado objetivado ao sítio EMX1. A porção do sítio do DNA alvo que interage com a região de complementaridade de gRNA está destacada em cinza com a primeira base da sequência PAM mostrada em letras minúsculas. As exclusões são indicadas por traços destacados em cinza e inserções por letras em itálico destacadas em cinza. O número líquido de bases eliminadas ou inseridas e o número de vezes em que cada sequência foi isolada são mostrados à direita.

[0034] Figura 3D: Eficiências de alterações mediadas por HDR/ssODN precisas introduzidas em dois genes humanos endógenos pelo padrão pareado e tru-RGNs. % HDR foi medida usando um ensaio de digestão de restrição por *BamHI* (vide os Procedimentos Experimentais para o Exemplo 2). gRNA de controle = vetor de promotor U6 vazio.

[0035] Figura 3E: Células U2OS.EGFP foram transfectadas com quantidades variáveis de plasmídeos de expressão de gRNA de comprimento completo (superior) ou plasmídeos de expressão de tru-gRNA (inferior), juntamente com uma quantidade fixa de plasmídeo de expressão de Cas9e, então, ensaiadas quanto à percentagem de células expressão reduzida de EGFP. Os valores médios obtidos a partir

de experimentos em duplicata são mostrados com erros padrões da média. Note que os dados obtidos com tru-gRNA correspondem estreitamente aos dados obtidos em experimentos realizados com plasmídeos de expressão de gRNA de comprimento completo em vez de plasmídeos de tru-gRNA para estes três sítios alvo de EGFP.

[0036] Figura 3F: Células U2OS.EGFP foram transfectadas com quantidade variável de plasmídeo de expressão de Cas9 juntamente com quantidades fixas de plasmídeos de expressão de gRNA de comprimento completo (superior) ou plasmídeos de expressão de tru-gRNA (inferior) para cada alvo (quantidades determinadas para cada tru-gRNA a partir dos experimentos da Figura 3E). Os valores médios obtidos a partir de experimentos em duplicata são mostrados com erros padrões da média. Note que os dados obtidos com tru-gRNA correspondem estreitamente aos dados obtidos em experimentos realizados com plasmídeos de expressão de gRNA de comprimento completo em vez de plasmídeos de tru-gRNA para estes três sítios alvo de EGFP. Os resultados destas titulações determinaram as concentrações de plasmídeos usados nos ensaios de ruptura por EGFP realizados nos Exemplos 1 e 2.

[0037] Figura 4A: Diagrama esquemático que ilustra localizações de sítios 1 e 4 de VEGFA objetivados por gRNAs para nicks duplos pareados. Sítios alvo para os gRNAs de comprimento completo estão sublinhados, com a primeira base na sequência PAM mostrada em letras minúsculas. A localização do sítio de restrição *Bam*HI inserido por HDR com um doador ssODN é mostrada.

[0038] Figura 4B: Um tru-gRNA pode ser usado com uma estratégia de nickase pareada para induzir de forma eficiente a mutações indel. Substituição de um gRNA de comprimento completo para o sítio VEGFA 1 por um tru-gRNA não reduz a eficácia de mutações indel observadas com um gRNA de comprimento completo pareado

para o sítio VEGFA 4 e nickases Cas9-D10A. gRNA de controle usado é um que carece de uma região de complementaridade.

[0039] Figura 4C: Um tru-gRNA pode ser usado com uma estratégia de nickase pareada para induzir de forma eficiente a alterações de sequência mediadas por HDR/ssODN precisas. Substituição de um gRNA de comprimento completo para o sítio 1 de VEGFA por um tru-gRNA não reduz a eficácia de mutações indel observadas com um gRNA de comprimento completo pareado para o sítio 4 de VEGFA e nickases Cas9-D10A com um modelo doador ssODN. gRNA de controle usado é um que carece de uma região de complementaridade.

[0040] Figura 5A: Atividades de RGNs objetivadas a três sítios em EGFP usando gRNA de comprimento completo (superior) ou tru-gRNAs (inferior) com desemparelhamentos individuais em cada posição (exceto na base mais 5', a qual deve permanecer um G para expressão eficiente do promotor U6). Caixas cinzas na grade abaixo representam posições dos desemparelhamentos de transversão de Watson-Crick. gRNA de controle vazio usado é um gRNA que carece de uma região de complementaridade. Atividades de RGN foram medidas usando o ensaio de ruptura por EGFP e os valores mostrados representam a percentagem de EGFP-negativa observada em relação a uma RGN que usa um gRNA perfeitamente correspondente. Os experimentos foram realizados em duplicata e as médias com barras de erro que representam erros padrões da média são mostradas.

[0041] Figura 5B: Atividades de RGNs objetivadas a três sítios em EGFP usando gRNA de comprimento completo (superior) ou tru-gRNAs (inferior) com desemparelhamentos duplos adjacentes em cada posição (exceto na base mais 5', a qual deve permanecer um G para expressão eficiente a partir do promotor U6). Dados apresentados conforme em 5A.

[0042] Figura 6A: Frequências absolutas de mutações indel sobre e fora do alvo induzidas por RGNs objetivadas a três diferentes sítios de genes humanos endógenos conforme medido por sequenciamento de profundidade. Frequências indel são mostradas para os três sítios alvo a partir de células nas quais RGNs com um gRNA de comprimento completo, um tru-gRNA ou um gRNA de controle que carece de uma região de complementaridade foram expressas. As contagens absolutas de mutações indel usadas para fazer estes gráficos podem ser encontradas na **Tabela 3B**.

[0043] Figura 6B: Vezes de aprimoramento nas especificidades por sítios fora do alvo de três tru-RGNs. Os valores mostrados representam a proporção de atividade sobre/fora do alvo de Tru-RGNs para as atividades sobre/fora do alvo de RGNs padrões para os sites fora do alvo mostrados, calculados usando os dados de (A) e **Tabela 3B**. Para os sítios marcados com um asterisco (*), nenhum indel foi observado com tru-RGN e, portanto, os valores mostrados representam estimativas estatísticas conservadas para vezes de aprimoramento nas especificidades por estes sítios fora do alvo (vide Resultados e Procedimentos Experimentais).

[0044] Figura 6C, superior: Comparação dos sítios sobre o alvo e fora do alvo identificados pelo ensaio de T7EI para a tru-RGN objetivada ao sítio 1 de VEGFA (mais foram identificados por meio de sequenciamento de profundidade). Note que o gRNA de comprimento completo tem desemparelhamentos nos dois nucleotídeos na extremidade 5' do sítio alvo e que estes são os dois nucleotídeos que não estão presentes no sítio alvo de tru-gRNA. Desemparelhamentos no sítio fora do alvo em relação ao sítio sobre o alvo em são destacados no texto sublinhado em **negrito**. Desemparelhamentos entre os gRNAs e o sítio fora do alvo são mostrados com X's.

[0045] Figura 6C, inferior: Frequências de mutações indel

induzidas no sítio fora do alvo por RGNs trazendo gRNAs de comprimento total ou truncados. Frequências de mutação indel foram determinadas por meio do ensaio de T7EI. Observe que o sítio fora do alvo nesta figura é aquele que tínhamos examinado anteriormente quanto a mutações indel induzidas pela RGN padrão objetivada ao sítio 1 de VEGFA e designado como sítio OT1-30 nesse estudo anterior (Exemplo 1 e Fu *et al.*, Nat Biotechnol 31 (9): 822-6 (2013)). É provável que não identificamos mutações fora do alvo neste sítio em nossos experimentos anteriores porque a frequência de mutações indel parece estar no limite de detecção confiável do ensaio de T7EI (2-5%).

[0046] Figuras 7A a D: Sequências de DNA das mutações indel induzidas por RGNs usando tru-gRNAs ou gRNAs de comprimento completo pareados objetivados aos sítios 1 e 3 de VEGFA. Sequências representadas conforme na Figura 3C.

[0047] Figura 7E. Frequências de mutação indel induzidas por tru-gRNAs tendo um nucleotídeo G 5' não pareado. Frequências de mutações indel em células U2OS.EGFP humanas induzidas por Cas9 dirigido por tru-gRNAs trazendo regiões de complementaridade de 17, 18 ou 20 nt para os sítios 1 e 3 de VEGFA e o sítio 1 de EMX1 são mostradas. Três destes gRNAs contêm uma G 5' não pareada (indicado por posições marcadas em negrito). As barras indicam os resultados de experimentos usando gRNA de comprimento completo (20 nt), tru-gRNA (17 ou 18 nt) e tru-gRNA com um nucleotídeo G 5' não pareado (17 ou 18 nt com T em negrito na extremidade 5'). (Note que nenhuma atividade foi detectada pelo tru-gRNA não pareado ao sítio 1 de EMX1).

[0048] Figuras 8A a C: Sequências de mutações indel fora do alvo induzidas por RGNs em células U2OS.EGFP humanas. Sítios fora do alvo genômicos de tipo selvagem reconhecidos por RGNs (incluindo a sequência PAM) são destacados em cinza e numerados conforme na **Tabela 1** e na **Tabela B**. Note que a fita complementar é mostrada para

alguns sítios. Bases eliminadas são mostradas como traços em um fundo cinza. Bases inseridas estão em itálico e destacadas em cinza.

[0049] Figuras 9A a C: Sequências de mutações indel fora do alvo induzidas por RGNs em células HEK293 humanas. Sítios fora do alvo genômicos de tipo selvagem reconhecidos por RGNs (incluindo a sequência PAM) são destacados em cinza e numerados conforme na **Tabela 1** e na **Tabela B**. Note que a fita complementar é mostrada para alguns sítios. Bases eliminadas são mostradas como traços em um fundo cinzento. Bases inseridas estão em itálico e destacadas em cinza.

*Proporcionou um grande número de indels de pb individuais.

DESCRIÇÃO DETALHADA

[0050] Nucleases CRISPR RNA-orientadas (RGNs) surgiram rapidamente como uma plataforma fácil e eficiente para edição de genoma. Embora Marraffini e colaboradores (Jiang *et al.*, Nat Biotechnol 31, 233-239 (2013)) tenham realizado recentemente uma investigação sistemática de especificidade RGN Cas9 em bactérias, as especificidades de RGNs em células humanas não foram extensivamente definidas. Compreensão do âmbito de efeitos fora do alvo mediados por RGN em células humanas e outras eucariotas será extremamente essencial se estas nucleases têm de ser usadas amplamente para pesquisa e aplicações terapêuticas. Os presentes inventores usaram um ensaio repórter com base em células humanas para caracterizar a clivagem fora do alvo de RGNs com base em Cas9. Desemparelhamentos individuais e duplos foram tolerados em diferentes graus, dependendo de sua posição ao longo da interface de DNA-RNA guia (gRNA). Alterações fora do alvo induzidas por quatro de seis RGNs objetivadas a *loci* endógenos em células humanas foram facilmente detectadas ao examinar sítios com parcialmente não pareados. Os sítios fora do alvo identificados trazem até cinco desemparelhamentos e muitos têm mutações com frequências

comparáveis (ou superiores) às aquelas observadas nos sítios sobre o alvo pretendidos. Assim, RGNs são altamente ativas mesmo com interfaces de DNA-RNA não perfeitamente pareadas em células humanas, uma descoberta que poderia confundir seu uso em pesquisa e aplicações terapêuticas.

[0051] Os resultados descritos aqui mostram que prever o perfil de especificidade de uma determinada RGN não é simples nem direto. Os experimentos de ensaio com o repórter *EGFP* mostram que desemparelhamentos individuais e duplos podem ter efeitos variáveis sobre a atividade de RGN em células humanas que não dependem estritamente de sua(s) posição(ões) dentro do sítio alvo. Por exemplo, consistente com relatórios anteriormente publicados, alterações na metade 3' da interface de sgRNA/DNA têm, em geral, efeitos maiores do que aquelas na metade 5' (Jiang *et al.*, Nat Biotechnol 31, 233-239 (2013); Cong *et al.*, Science 339, 819-823 (2013); Jinek *et al.*, Science 337, 816-821 (2012)); no entanto, mutações individuais e duplas na extremidade 3', algumas vezes, também parecem ser bem toleradas, enquanto que mutações duplas na extremidade 5' podem diminuir significativamente as atividades. Além disso, a magnitude destes efeitos para desemparelhamentos em qualquer (quaisquer) posição(ões) parece ser dependente do sítio. Caracterização de perfil abrangente de uma grande série de RGNs com testagem de todas as possíveis substituições de nucleotídeos (além das transversões de Watson-Crick usadas em nossos experimentos com o repórter *EGFP*) pode ajudar a fornecer informações adicionais quanto à faixa de sítios fora do alvo potenciais. A este respeito, o método com base em células bacterianas descrito recentemente de Marraffini e colaboradores (Jiang *et al.*, Nat Biotechnol 31, 233-239 (2013)) ou as metodologias de seleção de sítios por clivagem com base em bibliotecas combinatórias *in vitro* anteriormente aplicado a ZFNs por Liu e colaboradores (Pattanayak *et*

al., Nat Methods 8, 765-770 (2011)) podem ser úteis para a geração de grandes conjuntos de perfis de especificidade de RGN.

[0052] Apesar destes desafios em prever de forma abrangente as especificidades de RGN, foi possível identificar *bona fide* sítios fora do alvo de RGNs ao examinar um subconjunto de sítios genômicos que diferiam do site sobre alvo por uma a cinco desemparelhamentos. Notavelmente, sob as condições destes experimentos, as frequências de mutações induzidas por RGN em muitos destes sítios fora do alvo foram similares (ou superiores) àquelas observadas no sítio sobre o alvo pretendido, permitindo a detecção de mutações nestes sítios usando o ensaio de T7EI (o qual, conforme realizado em nosso laboratório, tem um limite de detecção confiável de ~ 2 a 5% de frequência de mutação). Uma vez que estas taxas de mutação eram muito elevadas, foi possível evitar o uso de métodos de sequenciamento profundo anteriormente necessários para detectar uma frequência muito menor de mutações fora do alvo induzidas por ZFN e TALEN (Pattanayak *et al.*, Methods Nat 8, 765-770 (2011); Perez *et al.*, Nat Biotechnol 26, 808-816 (2008); Gabriel *et al.*, Nat Biotechnol 29, 816-823 (2011); Hockemeyer *et al.*, Nat Biotechnol 29, 731-734 (2011)). Análise de mutagênese fora do alvo por RGN em células humanas também confirmou as dificuldades de previsão das especificidades de RGN - nem todos os sítios fora do alvo com desemparelhamentos individuais e duplos mostram evidências de mutação, enquanto que alguns sítios com até cinco desemparelhamentos também podem mostrar alterações. Além disso, os sítios fora do alvo *bona fide* identificados não exibem qualquer tendência evidente para diferenças de transição ou transversão em relação à sequência alvo pretendida (**Tabela E**; colunas destacadas em cinza).

[0053] Embora sítios fora do alvo fossem observados para um determinado número de RGNs, a identificação destes sítios não foi nem

abrangente nem em larga escala no genoma. Para as seis RGNs estudadas, apenas um pequeno subconjunto do número total muito maior de sequências fora do alvo potenciais no genoma humano (sítios que diferem por três a seis nucleotídeos em relação ao sítio alvo pretendido; comparar **Tabelas E e C**) foi examinado. Embora examinar um número tão grande de *loci* para mutações fora do alvo por meio do ensaio de T7EI não seja nem um prático nem uma estratégia de baixo custo, o uso de sequenciamento de alto rendimento em futuros estudos poderá permitir o exame de um maior número de sítios fora do alvo candidatos e proporcionar um método mais sensível para detecção *bona fide* de mutações fora do alvo. Por exemplo, tal abordagem poderia permitir a divulgação de sítios fora do alvo adicionais para as duas RGNs para as quais nós não conseguimos descobrir eventuais mutações fora do alvo. Além disso, uma melhor compreensão tanto das especificidades de RGN quanto quaisquer fatores epigenômicos (por exemplo, metilação de DNA e estado da cromatina) que podem influenciar as atividades de RGN em células também pode reduzir o número de sítios potenciais que precisam ser examinados e, assim, tornar avaliações de todo o genoma de RGNs fora do alvo mais práticas e acessíveis.

[0054] Conforme descrito aqui, uma série de estratégias podem ser usadas para minimizar as frequências de mutações genômicas fora do alvo. Por exemplo, a escolha específica do sítio alvo de RGN pode ser otimizada; uma vez que sítios fora do alvo que diferem em até cinco posições em relação ao sítio alvo pretendido podem sofrer mutação eficientemente por RGNs, parece improvável que a escolha de sítios alvo com um número mínimo de sítios fora do alvo, conforme avaliado por contagem de desemparelhamentos, seja eficaz; tipicamente, existirão milhares de sítios fora do alvo potenciais, que diferem em quatro ou cinco posições dentro da região de complementaridade de

RNA:DNA de 20 pb, para qualquer dada RGN objetivada a uma sequência no genoma humano (vide, por exemplo, **Tabela C**). Também é possível que o teor de nucleotídeos da região de complementaridade do gRNA possa influenciar a faixa de efeitos fora do alvo potenciais. Por exemplo, foi mostrado que um alto teor de GC estabiliza híbridos de RNA:DNA (Sugimoto *et al.*, Biochemistry 34, 11211-11216 (1995)). E, portanto, pode também ser esperado tornar a hibridização de gRNA/DNA genômico mais estável e mais tolerante a desemparelhamentos. Serão necessários experimentos adicionais com um maior número de gRNAs para avaliar se e como estes dois parâmetros (números de sítios não pareados no genoma e estabilidade dos híbridos de RNA:DNA) influenciam as especificidades de RGNs de todo o genoma. No entanto, é importante notar que, mesmo se tais parâmetros de predição possam ser definidos, o efeito de implementação de tais diretrizes restringiria ainda mais a faixa de objetivação de RGNs.

[0055] Uma estratégia geral para redução potencial de efeitos fora do alvo induzidos por RGN pode ser reduzir as concentrações de gRNA e nuclease Cas9 expressas na célula. Esta ideia foi testada usando as RGNs para sítios 2 e 3 alvo de *VEGFA* em células U2OS.EGFP; transfecção de plasmídeos que expressam menos sgRNA e Cas9 diminuiu a taxa de mutação no sítio sobre o alvo, mas não alterou sensivelmente as taxas relativas de mutações fora do alvo (**Tabelas 2A e 2B**). Consistente com isto, taxas de mutação fora do alvo de alto nível também foram observadas em dois outros tipos de células humanas (células HEK293 e K562), mesmo embora as taxas absolutas de mutagênese sobre o alvo fossem mais baixas do que em células U2OS.EGFP. Assim, não é provável que redução dos níveis de expressão de gRNA e Cas9 em células constituam uma solução para reduzir os efeitos fora do alvo. Além disso, estes resultados também

sugerem que as elevadas taxas de mutagênese fora do alvo observadas em células humanas não são causadas por superexpressão de gRNA e/ou Cas9.

Tabela 2A

Frequências de mutações indel em sítios genômicos sobre e fora do alvo induzidas por diferentes quantidades de plasmídeos de expressão de Cas9 e gRNA individuais para a RGN objetivada ao Sítio 2 do alvo *VEGFA*

Sítio	Sequência	SEQ ID NO:	250 ng de gRNA / 750 ng de Cas9 Frequência média de indel (%) \pm SEM	12,5 ng de gRNA / 50 ng de Cas9 Frequência média de indel (%) \pm SEM
T2 (sobre o alvo)	GACCCCCTCCACCCCGCCTCCGG	12	50,2 \pm 4,9	25,4 \pm 4,8
OT2-1	GACCCCC <u>C</u> CCACCCCGCC <u>C</u> CCGG	13	14,4 \pm 3,4	4,2 \pm 0,2
OT2-2	G <u>GG</u> CCCCCTCCACCCCGCCTCTGG	14	20,0 \pm 6,2	9,8 \pm 1,1
OT2-6	<u>CTA</u> CCCCCTCCACCCCGCCTCCGG	15	8,2 \pm 1,4	6,0 \pm 0,5
OT2-9	G <u>C</u> CCCC <u>AC</u> CCACCCCGCCTCTGG	16	50,7 \pm 5,6	16,4 \pm 2,1
OT2-15	<u>T</u> ACCCCC <u>CA</u> CACCCCGCCTCTGG	17	9,7 \pm 4,5	2,1 \pm 0,0
OT2-17	<u>ACA</u> CCCC <u>CC</u> ACCCCGCCTCAGG	18	14,0 \pm 2,8	7,1 \pm 0,0
OT2-19	<u>ATT</u> CCCC <u>CC</u> ACCCCGCCTCAGG	19	17,0 \pm 3,3	9,2 \pm 0,4
OT2-20	<u>CCC</u> A <u>CC</u> CCACCCCGCCTCAGG	20	6,1 \pm 1,3	N.D.
OT2-23	<u>CG</u> CCCT <u>CC</u> CCACCCCGCCTCCGG	21	44,4 \pm 6,7	35,1 \pm 1,8
OT2-24	<u>CT</u> CCCC <u>AC</u> CCACCCCGCCTCAGG	22	62,8 \pm 5,0	44,1 \pm 4,5
OT2-29	<u>TG</u> CCCT <u>CC</u> CCACCCCGCCTCTGG	23	13,8 \pm 5,2	5,0 \pm 0,2
OT2-34	<u>AGG</u> CCCC <u>CA</u> CACCCCGCCTCAGG	24	2,8 \pm 1,5	N.D.,

Quantidades de plasmídeos de expressão de gRNA e Cas9 transfectados em células U2OS.EGFP para estes ensaios são mostradas na parte superior de cada coluna. (Note que os dados para 250 ng de gRNA/750 ng de Cas9 são os mesmos conforme aqueles apresentados na **Tabela 1**). As frequências médias de indel foram determinadas usando o ensaio de T7EI a partir de amostras replicadas, conforme descrito em **Métodos**. OT = sítios fora do alvo, numerados conforme na **Tabela 1** e **Tabela B**. Desemparelhamentos do sítio sobre o alvo (dentro da região de 20 pb à qual o gRNA hibridiza) estão destacados em negrito, sublinhado. N.D. = nenhum detectado.

Tabela 2B

Frequências de mutações indel em sítios genômicos sobre e fora do alvo induzidas por diferentes quantidades de plasmídeos de expressão de Cas9 e gRNA individuais para a RGN

objetivada ao Sítio 3 do alvo VEGFA

Sítio	Sequência	SEQ ID NO:	250 ng de gRNA / 750 ng de Cas9 Frequência média de indel (%) \pm SEM	12,5 ng de gRNA / 50 ng de Cas9 Frequência média de indel (%) \pm SEM
T3 (sobre o alvo)	GGTGAGTGAGTGTGTGCGTGTGG	25	49,4 \pm 3,8	33,0 \pm 3,7
OT3-1	GGTGAGTGAGTGTGTG <u>I</u> GTGAGG	26	7,4 \pm 3,4	N.D.
OT3-2	<u>A</u> GTGAGTGAGTGTGTG <u>I</u> GTGGGG	27	24,3 \pm 9,2	9,8 \pm 4,2
OT3-4	G <u>C</u> TGAGTGAGTGT <u>A</u> TGCGTGTGG	28	20,9 \pm 11,8	4,2 \pm 1,2
OT3-9	GGTGAGTGAGTG <u>C</u> GTGCG <u>G</u> GTGG	29	3,2 \pm 0,3	N.D.
OT3-17	G <u>I</u> TGAGTGA <u>A</u> TGTGTGCGTGAGG	30	2,9 \pm 0,2	N.D.
OT3-18	<u>I</u> GTG <u>G</u> GTGAGTGTGTGCGTGAGG	31	13,4 \pm 4,2	4,9 \pm 0,0
OT3-20	<u>A</u> G <u>A</u> GAGTGAGTGTGTG <u>C</u> ATGAGG	32	16,7 \pm 3,5	7,9 \pm 2,4

Quantidades de plasmídeos de expressão de gRNA e Cas9 transfectados em células U2OS.EGFP para estes ensaios são mostradas na parte superior de cada coluna. (Note que os dados para 250 ng de gRNA/750 ng de Cas9 são os mesmos conforme aqueles apresentados na **Tabela 1**). As frequências médias de indel foram determinadas usando o ensaio de T7EI a partir de amostras replicadas, conforme descrito em **Métodos**. OT = sítios fora do alvo, numerados conforme na **Tabela 1** e **Tabela B**. Desemparelhamentos do sítio sobre o alvo (dentro da região de 20 pb à qual o gRNA hibridiza) estão destacados em negrito, sublinhado. N.D. = nenhum detectado.

[0056] A descoberta de que mutagênese significativa fora do alvo pode ser induzida por RGNs em três tipos diferentes de células humanas tem implicações importantes para o uso mais difundido desta plataforma de edição de genoma. Para aplicações de pesquisa, os efeitos potencialmente confusos de mutações fora do alvo de alta frequência terão de ser considerados, especialmente para experimentos que envolvem o uso de células cultivadas ou organismos com tempos de geração lentos para os quais cruzamento de alterações indesejadas seria um desafio. Uma forma de controlar tais efeitos pode ser usar várias RGNs objetivadas a diferentes sequências de DNA para induzir à mesma alteração genômica porque os efeitos fora do alvo não são aleatórios, mas sim relacionados ao sítio objetivado. No entanto, para aplicações terapêuticas, estes resultados indicam claramente que as especificidades de RGNs terá de ser cuidadosamente definida e/ou melhorada se estas nucleases devem ser usadas com segurança a

longo prazo para o tratamento de doenças humanas.

Métodos para Aprimoramento de Especificidade

[0057] Conforme mostrado aqui, nucleases RNA-orientadas CRISPR-Cas com base em proteína Cas de *S. pyogenes* podem ter efeitos mutagênicos fora do alvo significativos que são comparáveis a ou mais elevados do que a atividade pretendida sobre o alvo (Exemplo 1). Tais efeitos fora do alvo podem ser problemáticos para pesquisa e, em particular, para aplicações terapêuticas potenciais. Portanto, são necessários métodos para melhorar a especificidade de nucleases RNA-orientadas CRISPR-Cas (RGNs).

[0058] Conforme descrito no Exemplo 1, RGNs Cas9 podem induzir a mutações indel de alta frequência em sítios fora do alvo em células humanas (vide também Cradick *et al.*, 2013; Fu *et al.*, 2013; Hsu *et al.*, 2013; Pattanayak *et al.*, 2013). Estas alterações indesejáveis podem ocorrer em sequências genômicas que diferem em até cinco desemparelhamentos em relação ao sítio sobre o alvo pretendido (vide Exemplo 1). Além disso, embora desemparelhamentos na extremidade 5' da região de complementaridade do gRNA sejam geralmente melhor toleradas do que aqueles que estão na extremidade 3', estas associações não são absolutas e mostram dependência de sítio para sítio (vide Exemplo 1 e Fu *et al.*, 2013; Hsu *et al.*, 2013; Pattanayak *et al.*, 2013). Como um resultado, métodos computacionais que dependem do número e/ou posições de desemparelhamentos têm, atualmente, valor preditivo limitado para a identificação *bona fide* de sítios fora do alvo. Portanto, métodos para redução das frequências de mutações fora do alvo continuam a ser uma prioridade importante se nucleases RNA-orientadas têm de ser usadas para pesquisa e aplicações terapêuticas.

RNAs Guia Truncados (tru-gRNAs) Obtêm Maior Especificidade

[0059] RNAs guia, de um modo geral, vêm em dois sistemas

diferentes: Sistema 1, o qual usa crRNA e tracrRNAs separados que funcionam em conjunto para orientar a clivagem por Cas9, e o Sistema 2, o qual usa um híbrido de crRNA-tracrRNA quimérico que combina os dois RNAs guia separados em um único sistema (dito como um RNA guia individual ou sgRNA; vide também Jinek *et al.*, Science 2012; 337: 816-821). O tracrRNA pode ser variavelmente truncado e foi mostrado que uma série de comprimentos funcionam tanto no sistema separado (Sistema 1) quanto no sistema de gRNA quimérico (Sistema 2). Por exemplo, em algumas modalidades, o tracrRNA pode ser truncado a partir de sua extremidade 3' em pelo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35 ou 40 nts. Em algumas modalidades, a molécula de tracrRNA pode ser truncada a partir de sua extremidade 5' em pelo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35 ou 40 nts. Alternativamente, a molécula de tracrRNA pode ser truncada a partir de ambas as extremidades 5'e 3', por exemplo, em pelo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15 ou 20 nts na extremidade 5' e em pelo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35 ou 40 nts na extremidade 3'. Vide, por exemplo, Jinek *et al.*, Science 2012; 337: 816-821; Mali *et al.*, Science. 15 de Fevereiro de 2013; 339 (6121): 823-6; Cong *et al.*, Science. 15 de Fevereiro de 2013; 339 (6121): 819-23; e Hwang e Fu *et al.*, Nat Biotechnol. Março de 2013; 31 (3): 227-9; Jinek *et al.*, Elife 2, e00471 (2013)). Para o Sistema 2, em geral, gRNAs quiméricos de comprimento mais longo demonstraram uma maior atividade sobre o alvo, mas as especificidades relativas de gRNAs de vários comprimentos atualmente continuam indefinidas e, portanto, pode ser desejável, em determinados casos, usar gRNAs mais curtos. Em algumas modalidades, os gRNAs são complementares a uma região que está dentro de cerca de 100-800 pb a montante do sítio de início de transcrição, por exemplo, está dentro de cerca de 500 pb a montante do sítio de início de transcrição, inclui o sítio de início de transcrição ou está

dentro de cerca de 100 -800 pb, por exemplo, dentro de cerca de 500 pb, a jusante do sítio de início de transcrição. Em algumas modalidades, vetores (por exemplo, plasmídeos) que codificam mais de um gRNA são usados, por exemplo, plasmídeos que codificam 2, 3, 4, 5 ou mais gRNAs objetivados a diferentes sítios da mesma região do gene alvo.

[0060] O presente pedido descreve uma estratégia para melhorar a especificidade de RGNs com base na ideia aparentemente contra-intuitiva de encurtar, em vez de alongar, a região de complementaridade do gRNA. Estes gRNAs mais curtos podem induzir a vários tipos de eventos de edição de genoma sobre o alvo mediada por Cas9 com eficiência comparável (ou, em alguns casos, superior) aos gRNAs de comprimento completo em múltiplos sítios em um único gene repórter *EGFP* integrado e em genes humanos endógenos. Além disso, RGNs usando estes gRNAs encurtados exibem sensibilidade aumentada a um pequeno número de desemparelhamentos na interface gRNA-DNA alvo. Mais importante, o uso de gRNAs encurtados reduz substancialmente as taxas de efeitos fora do alvo genômico em células humanas, proporcionando aprimoramentos de especificidade tão elevados quanto 5000 vezes ou mais nesses sítios. Assim, esta estratégia de gRNA encurtado constitui uma abordagem altamente eficaz para reduzir os efeitos fora do alvo sem comprometer a atividade sobre o alvo e sem a necessidade de expressão de um segundo gRNA potencialmente mutagênico. Esta abordagem pode ser aplicada em si ou em conjunto com outras estratégias, tal como o método de nickase pareada, para reduzir os efeitos fora do alvo de RGNs em células humanas.

[0061] Assim, um método para aumentar a especificidade de nucleases CRISPR/Cas é encurtar o comprimento da espécie de RNA guia (gRNA) usada para dirigir a especificidade de nuclease. A nuclease Cas9 pode ser orientada a alvos genômicos de 17-18 nt específicos que

trazem um motivo protoespaçador adjacente (**Protospacer Adjacent Motif - PAM**) proximal adicional, por exemplo, de sequência NGG, usando uma RNA-guia, por exemplo, um gRNA individual ou um crRNA (pareado com um tracrRNA), tendo 17 ou 18 nt em sua extremidade 5' que são complementares com a fita complementar do sítio alvo do DNA genômico (**Figura 1**).

[0062] Embora se possa esperar que o aumento do comprimento da região de complementaridade do gRNA melhore a especificidade, os presentes inventores (Hwang *et al.*, PLoS One. 09 de Julho de 2013; 8 (7): E68708) e outros (Ran *et al.*, Cell. 12 de Setembro 2013; 154 (6): 1380-9) observaram anteriormente que alongamento da região de complementaridade do sítio alvo na extremidade 5' do gRNA realmente torna sua função menos eficiente no sítio sobre o alvo.

[0063] Por outro lado, os experimentos no Exemplo 1 mostraram que gRNAs tendo vários desemparelhamentos dentro de uma região de objetivação de complementaridade 5' de comprimento padrão ainda poderia induzir à clivagem robusta mediada por Cas9 de seus sítios alvo. Assim, é possível que gRNAs truncados que carecem destes nucleotídeos na extremidade 5' possam exibir atividades comparáveis aos seus homólogos de comprimento completo (Figura 2a). Além disso, foi especulado que estes nucleotídeos 5' normalmente poderiam compensar desemparelhamentos em outras posições ao longo da interface de gRNA-DNA alvo e, portanto, previsto que gRNAs mais curtos poderiam ser mais sensíveis às diferenças e, assim, induzir a níveis mais baixos de mutações fora do alvo (Figura 2A).

[0064] Diminuição do comprimento da sequência de DNA específica poderá também diminuir a estabilidade do híbrido de gRNA:DNA, tornando-o menos tolerante a desemparelhamentos e, deste modo, tornando o alvo mais específico. Isto é, truncamento da sequência de gRNA para reconhecer um alvo de DNA mais curto pode,

na verdade, resultar em uma nuclease RNA-orientada que é menos tolerante mesmo a desemparelhamentos de um único nucleotídeo e, portanto, é mais específica e tem menos efeitos não intencionais fora do alvo.

[0065] Esta estratégia para encurtar a região de complementaridade de gRNA poderia potencialmente ser usada com outras proteínas de RNA-orientadas que não Cas9 de *S. pyogenes*, incluindo outras proteínas Cas de bactérias ou archaea, bem como variantes de Cas9 que nick uma única fita de DNA ou não têm atividade de nuclease, tal como uma dCas9 trazendo mutações de inativação catalítica em um ou ambos os domínios de nuclease. Esta estratégia pode ser aplicada a sistemas que usam um gRNA individual, bem como aqueles que usam gRNAs duplos (por exemplo, o crRNA e tracrRNA encontrado em sistemas de ocorrência natural).

[0066] Assim, é descrito aqui um RNA guia individual que compreende um crRNA fundido a um tracrRNA normalmente transcrito, por exemplo, um RNA guia individual de Cas9, conforme descrito em Mali *et al.*, Science. 15 de Fevereiro de 2013; 339 (6121): 823-6, mas com uma sequência na extremidade 5' que é complementar a menos de 20 nucleotídeos (nts), por exemplo, 19, 18 ou 17 nts, de preferência 17 ou 18 nts, da fita complementar a uma sequência alvo imediatamente 5' de um motivo protoespaçador adjacente (**P**roto**s**pacer **A**djacent **M**otif - PAM), por exemplo, NGG, NAG ou NNGG. Em algumas modalidades, o RNA guia de Cas9 encurtado consiste na sequência: (X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUA (SEQ ID NO:2404); (X₁₇₋₁₈ ou X₁₇₋₁₉) GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:2407); ou (X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCU (SEQ ID NO: 2408); (X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAGAAUAGCAAGUUAAAUAAGGCUAGUC

CG(X_N) (SEQ ID NO: 1);
 (X₁₇₋₁₈ ou
 X₁₇₋₁₉)GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUAAAAUAA
 GGCUAGUCCGUUAUC(X_N) (SEQ ID NO: 2);
 (X₁₇₋₁₈ ou
 X₁₇₋₁₉)GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGC
 AAGUUAAAAUAAGGCUAGUCCGUUAUC(X_N) (SEQ ID NO: 3); (X₁₇₋₁₈
 ou
 X₁₇₋₁₉)GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUC
 CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(X_N) (SEQ
 ID NO: 4),
 (X₁₇₋₁₈ ou
 X₁₇₋₁₉)GUUUAAGAGCUAGAAAUAGCAAGUUUAAAUAAAGGCUAGUC
 CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID
 NO: 5);
 (X₁₇₋₁₈ ou
 X₁₇₋₁₉)GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAU
 AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG
 UGC (SEQ ID NO: 6); ou (X₁₇₋₁₈ ou
 X₁₇₋₁₉)GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAU
 AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG
 UGC (SEQ ID NO: 7); em que X₁₇₋₁₈ ou X₁₇₋₁₉ é a sequência de

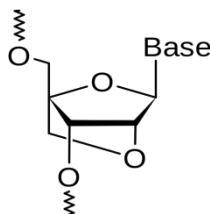
nucleotídeos complementar a 17-18 ou 17-19 nucleotídeos consecutivos da sequência alvo, respectivamente. São também descritos aqui DNAs que codificam os RNA guia de Cas9 encurtados que foram anteriormente descritos na literatura (Jinek *et al.*, Science 337 (6096): 816-21 (2012) e Jinek *et al.*, Elife 2:e00471 (2013)).

[0067] Os RNA guia podem incluir X_N, o qual pode ser qualquer sequência, em que N (no RNA) pode ser 0-200, por exemplo, 0-100, 0-50 ou 0-20, que não interfere com a ligação do ácido ribonucleico à

Cas9.

[0068] Em algumas modalidades, o RNA guia inclui um ou mais nucleotídeos adenina (A) ou uracila (U) sobre a extremidade 3'. Em algumas modalidades o RNA inclui um ou mais Us, por exemplo, 1 a 8 ou mais Us (por exemplo, U, UU, UUU, UUUU, UUUUU, UUUUUU, UUUUUUU, UUUUUUUU) na extremidade 3' da molécula como um resultado da presença opcional de um ou mais Ts usados como um sinal de término para terminar a transcrição de RNA PolIII.

[0069] Foi demonstrado que oligonucleotídeos de RNA modificados, tais como ácidos nucleicos bloqueados (**Locked Nucleic Acids - LNAs**), aumentam a especificidade da hibridização de RNA-DNA ao bloquear os oligonucleotídeos modificados em uma conformação mais favorável (estável). Por exemplo, o RNA 2'-O-metila é uma base modificada, onde há uma ligação covalente adicional entre o oxigênio 2' e o carbono 4' a qual, quando incorporada em oligonucleotídeos, pode melhorar a estabilidade térmica e a seletividade global (**fórmula I**).



fórmula I - ácido nucleico bloqueado

[0070] Assim, em algumas modalidades, os tru-gRNAs descritos aqui podem compreender um ou mais oligonucleotídeos de RNA modificados. Por exemplo, as moléculas de RNAs guia truncados descritas aqui podem ter um, alguns ou todos os 17-18 ou 17-19 nts da região 5' do RNA guia complementar à sequência alvo são modificadas, por exemplo, bloqueados (ponte de 2'-O-4'-C metileno), 5'-metilcitidina, 2'-O-metil-pseudouridina, ou nos quais a estrutura de fosfato de ribose foi substituída por uma cadeia de poliamida (ácido nucleico peptídico), por exemplo, um ácido ribonucleico sintético.

[0071] Em outras modalidades, um, alguns ou todos os nucleotídeos da sequência de tru-gRNA podem ser modificados, por exemplo, bloqueados (ponte de 2'-O-4'-C metileno), 5'-metilcitidina, 2'-O-metil-pseudouridina, ou nos quais a estrutura de fosfato de ribose foi substituída por uma cadeia de poliamida (ácido nucleico peptídico), por exemplo, um ácido ribonucleico sintético.

[0072] Em um contexto celular, complexos de Cas9 com estes gRNAs sintéticos podem ser usados para melhorar a especificidade de todo o genoma do sistema de CRISPR/nuclease Cas9.

[0073] tru-gRNAs modificados ou sintéticos exemplificativos podem compreender, ou consistir, as sequências a seguir:

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUA(X_N) (SEQ ID NO: 2404);

(X₁₇₋₁₈ ou X₁₇₋₁₉) GUUUUAGAGCUAUGCUGUUUUG (X_N)
(SEQ ID NO: 2407);

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCU(X_N) (SEQ ID NO: 2408);

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUC
CG(X_N) (SEQ ID NO:1);

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUAAAAUAA
GGCUAGUCCGUUAUC(X_N) (SEQ ID NO: 2);

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGC
AAGUUAAAAUAAGGCUAGUCCGUUAUC(X_N) (SEQ ID NO: 3);

(X₁₇₋₁₈)GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAG
GCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(X_N) (SEQ ID NO: 4),

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUC

CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID NO:5);

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAU AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG UGC (SEQ ID NO: 6); ou

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAU AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG UGC (SEQ ID NO: 7);

em que X₁₇₋₁₈ ou X₁₇₋₁₉ é uma sua sequência complementar a 17-18 ou 17-19 nts de uma sequência alvo, respectivamente, de preferência uma sequência alvo imediatamente 5' de um motivo protoespaçador adjacente (**P**roto**s**pac**e**r **A**djacent **M**otif - PAM), por exemplo, NGG, NAG ou NNGG, e ainda em que um ou mais dos nucleotídeos são bloqueados, por exemplo, um ou mais dos nucleotídeos dentro da sequência X₁₇₋₁₈ ou X₁₇₋₁₉, um ou mais dos nucleotídeos dentro da sequência X_N ou um ou mais dos nucleotídeos dentro de qualquer sequência de tru-gRNA. X_N é qualquer sequência, em que N (no RNA) pode ser 0-200, por exemplo, 0-100, 0-50 ou 0-20, que não interfere com a ligação do ácido ribonucleico à Cas9. Em algumas modalidades, o RNA inclui um ou mais Us, por exemplo, 1 a 8 ou mais Us (por exemplo, U, UU, UUU, UUUU, UUUUU, UUUUUU, UUUUUUU, UUUUUUUU) na extremidade 3' da molécula, como um resultado da presença opcional de um ou mais Ts usados como um sinal de término para terminar a transcrição de RNA PolIII.

[0074] Embora alguns dos exemplos descritos aqui usem um gRNA individual, os métodos podem também ser usados com gRNAs duplos (por exemplo, o crRNA e tracrRNA encontrados em sistemas de ocorrência natural). Neste caso, um tracrRNA individual seria usado em

conjunto com múltiplos crRNAs diferentes expressos usando o presente sistema, por exemplo, o seguinte: (X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUA (SEQ ID NO:2404); (X₁₇₋₁₈ ou X₁₇₋₁₉) GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO:2407); ou (X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCU (SEQ ID NO:2408); e uma sequência de tracrRNA. Neste caso, o crRNA é usado como o RNA guia nos métodos e moléculas descritos aqui e o tracrRNA pode ser expresso a partir da mesma ou uma molécula de DNA diferente. Em algumas modalidades, os métodos incluem contato da célula com um tracrRNA que compreende ou consiste na sequência de GGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCC GUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 8) ou uma porção ativa da mesma (uma porção ativa é aquela que retém a capacidade de formar complexos com Cas9 ou dCas9). Em algumas modalidades, a molécula de tracrRNA pode ser truncada a partir de sua extremidade 3' em pelo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35 ou 40 nts. Em outra modalidade, a molécula de tracrRNA pode ser truncada a partir de sua extremidade 5' em pelo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35 ou 40 nts. Alternativamente, a molécula de tracrRNA pode ser truncada a partir de ambas as extremidades 5' e 3', por exemplo, em pelo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15 ou 20 nts na extremidade 5' e pelo menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35 ou 40 nts na extremidade 3'. Sequências de tracrRNA exemplificativas, além de SEQ ID NO: 8, incluem as seguintes: UAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC (SEQ ID NO: 2405) ou uma porção ativa da mesma; AGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGC (SEQ ID NO: 2407) ou uma porção

ativa da mesma;
 CAAAACAGCAUAGCAAGUUAUAAAGGCUAGUCCGUUAUCAAC
 UUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 2409) ou uma
 porção ativa da mesma;
 UAGCAAGUUAUAAAGGCUAGUCCGUUAUCAACUUGAAAAAGU
 G (SEQ ID NO: 2410) ou uma porção ativa da mesma;
 UAGCAAGUUAUAAAGGCUAGUCCGUUAUCA (SEQ ID NO: 2411)
 ou uma porção ativa da mesma; ou
 UAGCAAGUUAUAAAGGCUAGUCCG (SEQ ID NO: 2412) ou uma
 porção ativa da mesma.

[0075] Em algumas modalidades em que (X_{17-18} ou X_{17-19})
 GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 2407) é usada como
 um crRNA, o seguinte tracrRNA é usado:
 GGAACCAUUCAAAACAGCAUAGCAAGUUAUAAAGGCUAGUCC
 GUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO:
 8) ou uma porção ativa da mesma. Em algumas modalidades em que
 (X_{17-18} ou X_{17-19}) GUUUUAGAGCUA (SEQ ID NO: 2404) é usada como
 um crRNA, o seguinte tracrRNA é usado:
 UAGCAAGUUAUAAAGGCUAGUCCGUUAUCAACUUGAAAAAGU
 GGCACCGAGUCGGUGC (SEQ ID NO: 2405) ou uma porção ativa da
 mesma. Em algumas modalidades em que (X_{17-18} ou X_{17-19})
 GUUUUAGAGCUAUGCU (SEQ ID NO: 2408) é usada como um crRNA,
 o seguinte tracrRNA é usado:
 AGCAUAGCAAGUUAUAAAGGCUAGUCCGUUAUCAACUUGAAA
 AAGUGGCACCGAGUCGGUGC (SEQ ID NO: 2406) ou uma porção
 ativa da mesma.

[0076] Além disso, em um sistema que usam crRNA e tracrRNA
 separados, um ou ambos podem ser sintéticos e incluem um ou mais
 nucleotídeos ou desoxirribonucleotídeos modificados (por exemplo,
 bloqueados).

[0077] Em algumas modalidades, os RNA guia individuais e/ou crRNAs e/ou tracrRNAs podem incluir um ou mais nucleotídeos adenina (A) ou uracila (U) na extremidade 3'.

[0078] RGNs com base em Cas9 existentes usam a formação de heteroduplas de gRNA-DNA para orientar a objetivação para sítios genômicos de interesse. No entanto, heteroduplas de RNA-DNA podem formar uma variedade de estruturas mais promíscuas do que suas contrapartes de DNA-DNA. Na verdade, duplas de DNA-DNA são mais sensíveis a diferenças, o que sugere que uma nuclease DNA-orientada pode não se ligar tão prontamente à sequências fora do alvo, tornando-a relativamente mais específica do que as nucleases RNA-orientadas. Assim, os RNAs guia truncados descritos aqui podem ser híbridos, isto é, em que um ou mais desoxirribonucleotídeos, por exemplo, um oligonucleotídeo de DNA curto, substitui a totalidade ou parte do gRNA, por exemplo, a totalidade ou parte da região de complementaridade de um gRNA. Esta molécula com base em DNA poderia substituir a totalidade ou parte do gRNA de um sistema de gRNA individual ou, alternativamente, pode substituir a totalidade ou parte do crRNA em um sistema de crRNA/tracrRNA duplo. Tal sistema que incorpora o DNA na região de complementaridade deve objetivar mais confiavelmente as sequências de DNA genômico previstas em virtude de intolerância geral de duplas de DNA-DNA a desemparelhamentos comparado com duplas de RNA-DNA. Métodos para fazer tais duplas são conhecidos na técnica; vide, por exemplo, Barker *et al.*, BMC Genomics. 22 de Abril de 2005; 6: 57; e Sugimoto *et al.*, Biochemistry. 19 de Setembro de 2000; 39 (37): 11270-81.

[0079] tru-gRNAs modificados ou sintéticos exemplificativos podem compreender, ou consistir, as sequências a seguir:

(X₁₇₋₁₈ ou

X₁₇₋₁₉)GUUUUAGAGCUAGAAAUAGCAAGUUAAAUAAGGCUAGUC

CG(X_N) (SEQ ID NO:1);

(X_{17-18} ou

X_{17-19})GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUAAAAUAA
GGCUAGUCCGUUAUC(X_N) (SEQ ID NO:2);

(X_{17-18} ou

X_{17-19})GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGC
AAGUUAAAAUAAGGCUAGUCCGUUAUC(X_N) (SEQ ID NO:3);

(X_{17-18} ou

X_{17-19})GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUC
CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(X_N) (SEQ
ID NO:4),

(X_{17-18} ou

X_{17-19})GUUUAAGAGCUAGAAAUAGCAAGUUUAAAUAAGGCUAGUC
CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID
NO:5);

(X_{17-18} ou

X_{17-19})GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAU
AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG
UGC (SEQ ID NO:6); ou

(X_{17-18} ou

X_{17-19})GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAU
AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG
UGC (SEQ ID NO:7);

em que X_{17-18} ou X_{17-19} é uma sua sequência complementar a 17-18 ou 17-19 nts de uma sequência alvo, respectivamente, de preferência uma sequência alvo imediatamente 5' de um motivo protoespaçador adjacente (**P**roto**s**pacer **A**djacent **M**otif - PAM), por exemplo, NGG, NAG ou NNGG, e ainda em que um ou mais dos nucleotídeos são desoxirribonucleotídeos, por exemplo, um ou mais dos nucleotídeos dentro da sequência X_{17-18} ou X_{17-19} , um ou mais dos

nucleotídeos dentro da sequência X_N ou um ou mais dos nucleotídeos dentro de qualquer sequência do tru-gRNA. X_N é qualquer sequência em que N (no RNA) pode ser 0-200, por exemplo, 0-100, 0-50 ou 0-20, que não interfere com a ligação do ácido ribonucleico à Cas9. Em algumas modalidades, o RNA inclui um ou mais L, por exemplo, 1 a 8 ou mais nós (por exemplo, U, UU, UUU, UUUU, UUUUU, UUUUUU, UUUUUUU, UUUUUUUU) na extremidade 3' da molécula, como um resultado da presença opcional de um ou mais Ts usados como um sinal de término para terminar a transcrição de RNA PolIII.

[0080] Além disso, em um sistema que usa crRNA e tracrRNA separados, um ou ambos podem ser sintéticos e incluem um ou mais desoxirribonucleotídeos.

[0081] Em algumas modalidades, os RNAs guia individuais ou crRNAs ou tracrRNAs incluem um ou mais nucleotídeos adenina (A) ou uracila (U) na extremidade 3'.

[0082] Em algumas modalidades, o gRNA é objetivado para um sítio que tem pelo menos três ou mais desemparelhamentos diferentes a partir de qualquer sequência no resto do genoma de modo a minimizar efeitos fora do alvo.

[0083] Os métodos descritos podem incluir expressão em uma célula, ou contato da célula com, um gRNA de Cas9 encurtado (tru-gRNA) conforme descrito aqui (opcionalmente um tru-gRNA modificado ou híbrido de DNA/RNA), além de uma nuclease que pode ser orientada pelos gRNAs de Cas9 encurtados, por exemplo, uma nuclease Cas9, por exemplo, conforme descrito em Mali *et al.*, uma nickase Cas9 conforme descrito no Jinek *et al.*, 2012; ou uma fusão de dCas9-domínio heterofuncional (dCas9-HFD).

Cas9

[0084] Uma série de bactérias expressam variantes de proteína Cas9. A Cas9 de *Streptococcus pyogenes* é atualmente a mais usada;

algumas das outras proteínas Cas9 têm níveis elevados de identidade de sequência com a Cas9 de *S. pyogenes* e usam os mesmos RNAs guia. Outras são mais diversas, usam diferentes gRNAs e reconhecem sequências PAM diferentes também (a sequência de 2-5 nucleotídeos especificada pela proteína a qual está adjacente à sequência especificada pelo RNA). Chylinski *et al.* classificaram proteínas Cas9 de um grande grupo de bactérias (RNA Biology 10: 5, 1-12; 2013) e um grande número de proteínas Cas9 estão listados na Figura 1 e Tabela 1 suplementar do mesmo, as quais são aqui incorporadas por referência. Proteínas Cas9 adicionais são descritas em Esvelt *et al.*, Methods Nat. Novembro de 2013; 10 (11): 1116-1121 e Fonfara *et al.*, "Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems". Nucleic Acids Res. 22 de Novembro de 2013. [cabeçalho Epub de impressão] doi:10.1093/nar/gkt1074.

[0085] Moléculas Cas9 de uma variedade de espécies podem ser usadas nos métodos e composições descritos aqui. Embora as moléculas de Cas9 de *S. pyogenes* e *S. thermophilus* sejam o assunto da maior parte da descrição aqui, moléculas de Cas9 derivadas de ou com base em proteínas Cas9 de outras espécies indicadas aqui podem ser usadas também. Em outras palavras, embora que a grande parte da descrição aqui use moléculas Cas9 de *S. pyogenes* e *S. thermophilus*, moléculas Cas9 de outras espécies podem substituí-las. Estas espécies incluem aquelas apresentadas na tabela a seguir, a qual foi criada com base na Figura 1 suplementar de Chylinski *et al.*, 2013.

Proteínas Cas9 alternativas	
GenBank Ac. No.	Bactéria
303229466	<i>Veillonella atypica</i> ACS-134-V-Col7a
34762592	<i>Fusobacterium nucleatum subsp. vincentii</i>
374307738	<i>Filifactor alocis</i> ATCC 35896

Proteínas Cas9 alternativas	
GenBank Ac. No.	Bactéria
320528778	<i>Solobacterium moorei</i> F0204
291520705	<i>Coprococcus catus</i> GD-7
42525843	<i>Treponema denticola</i> ATCC 35405
304438954	<i>Peptoniphilus duerdenii</i> ATCC BAA-1640
224543312	<i>Catenibacterium mitsuokai</i> DSM 15897
24379809	<i>Streptococcus mutans</i> UA159
15675041	<i>Streptococcus pyogenes</i> SF370
16801805	<i>Listeria innocua</i> Clip11262
116628213	<i>Streptococcus thermophilus</i> LMD-9
323463801	<i>Staphylococcus pseudintermedius</i> ED99
352684361	<i>Acidaminococcus intestini</i> RyC-MR95
302336020	<i>Olsenella uli</i> DSM 7084
366983953	<i>Oenococcus kitaharae</i> DSM 17330
310286728	<i>Bifidobacterium bifidum</i> S17
258509199	<i>Lactobacillus rhamnosus</i> GG
300361537	<i>Lactobacillus gasseri</i> JV-V03
169823755	<i>Finnegoldia magna</i> ATCC 29328
47458868	<i>Mycoplasma mobile</i> 163K
284931710	<i>Mycoplasma gallisepticum</i> str. F
363542550	<i>Mycoplasma ovipneumoniae</i> SC01
384393286	<i>Mycoplasma canis</i> PG 14
71894592	<i>Mycoplasma synoviae</i> 53
238924075	<i>Eubacterium rectale</i> ATCC 33656
116627542	<i>Streptococcus thermophilus</i> LMD-9
315149830	<i>Enterococcus faecalis</i> TX0012
315659848	<i>Staphylococcus lugdunensis</i> M23590
160915782	<i>Eubacterium dolichum</i> DSM 3991

Proteínas Cas9 alternativas	
GenBank Ac. No.	Bactéria
336393381	<i>Lactobacillus coryniformis</i> subsp. <i>torquens</i>
310780384	<i>Ilyobacter polytropus</i> DSM 2926
325677756	<i>Ruminococcus albus</i> 8
187736489	<i>Akkermansia muciniphila</i> ATCC BAA-835
117929158	<i>Acidothermus cellulolyticus</i> 11B
189440764	<i>Bifidobacterium longum</i> DJO10A
283456135	<i>Bifidobacterium dentium</i> Bd1
38232678	<i>Corynebacterium diphtheriae</i> NCTC 13129
187250660	<i>Elusimicrobium minutum</i> Pei191
319957206	<i>Nitratifractor salsuginis</i> DSM 16511
325972003	<i>Sphaerochaeta globus</i> str. <i>Buddy</i>
261414553	<i>Fibrobacter succinogenes</i> subsp. <i>succinogenes</i>
60683389	<i>Bacteroides fragilis</i> NCTC 9343
256819408	<i>Capnocytophaga ochracea</i> DSM 7271
90425961	<i>Rhodopseudomonas palustris</i> BisB18
373501184	<i>Prevotella micans</i> F0438
294674019	<i>Prevotella ruminicola</i> 23
365959402	<i>Flavobacterium columnare</i> ATCC 49512
312879015	<i>Aminomonas paucivorans</i> DSM 12260
83591793	<i>Rhodospirillum rubrum</i> ATCC 11170
294086111	<i>Candidatus Puniceispirillum marinum</i> IMCC1322
121608211	<i>Verminephrobacter eiseniae</i> EF01-2
344171927	<i>Ralstonia syzygii</i> R24
159042956	<i>Dinoroseobacter shibae</i> DFL 12
288957741	<i>Azospirillum</i> sp- B510
92109262	<i>Nitrobacter hamburgensis</i> X14
148255343	<i>Bradyrhizobium</i> sp- BTAi1

Proteínas Cas9 alternativas	
GenBank Ac. No.	Bactéria
34557790	<i>Wolinella succinogenes</i> DSM 1740
218563121	<i>Campylobacter jejuni</i> subsp. <i>jejuni</i>
291276265	<i>Helicobacter mustelae</i> 12198
229113166	<i>Bacillus cereus</i> Rock1-15
222109285	<i>Acidovorax ebreus</i> TPSY
189485225	Cupim, grupo 1, não cultivado
182624245	<i>Clostridium perfringens</i> D str.
220930482	<i>Clostridium cellulolyticum</i> H10
154250555	<i>Parvibaculum lavamentivorans</i> DS-1
257413184	<i>Roseburia intestinalis</i> L1-82
218767588	<i>Neisseria meningitidis</i> Z2491
15602992	<i>Pasteurella multocida</i> subsp. <i>multocida</i>
319941583	<i>Sutterella wadsworthensis</i> 3 1
254447899	<i>gama proteobactéria</i> HTCC5015
54296138	<i>Legionella pneumophila</i> str. Paris
331001027	<i>Parasutterella excrementihominis</i> YIT 11859
34557932	<i>Wolinella succinogenes</i> DSM 1740
118497352	<i>Francisella novicida</i> U112

[0086] Os construtos e métodos descritos aqui podem incluir o uso de qualquer uma destas proteínas Cas9 e seus RNAs guia correspondentes ou outros RNAs guia que são compatíveis. Também foi mostrado que a Cas9 do sistema CRISPR1 de *Streptococcus thermophilus* LMD-9 funciona em células humanas em Cong *et al.* (Science 339, 819 (2013)). Ortólogos de Cas9 de *N. meningitidis* são descritos em Hou *et al.*, Proc Natl Acad Sci USA: 24 de Setembro de 2013; 110 (39): 15644-9 e Esvelt *et al.*, Methods Nat. Novembro de 2013; 10 (11): 1116-1121. Além disso, Jinek *et al.* mostraram *in vitro* que ortólogos de Cas9 de *S. thermophilus* e *L. innocua* (mas não de *N.*

meningitidis ou *C. jejuni*, as quais provavelmente usam um RNA guia diferente), podem ser orientados por um gRNA duplo de *S. pyogenes* para clivar o DNA de plasmídeo alvo, embora com uma eficiência ligeiramente reduzida.

[0087] Em algumas modalidades, o presente sistema usa a proteína Cas9 de *S. pyogenes*, quer conforme codificada em bactérias ou com códons otimizados para expressão em células de mamíferos, contendo mutações em D10, E762, H983, H840 e D986 ou N863 ou, por exemplo, D10A/D10N e H840A/H840N/H840Y, para tornar a porção de nuclease da proteína cataliticamente inativa; substituições nestas posições podem ser alanina (conforme elas estão em Nishimasu *et al.*, Cell 156, 935-949 (2014)) ou podem ser outros resíduos, por exemplo, glutamina, asparagina, tirosina, serina ou aspartato, por exemplo, E762Q, H983N, H983Y, D986N, N863D, N863S ou N863H (**Figura 1C**). A sequência de Cas9 de *S. pyogenes* cataliticamente inativa que pode ser usada nos métodos e composições descritos aqui é como segue; as mutações exemplificativas de D10A e H840A estão em negrito e sublinhado.

```

10 20 30 40 50 60
MDKKYSIGLAIGTNSVGWAV ITDEYKVPSK KFKVLGNTDR HSIKKNLIGA LLFDSGETAE

70 80 90 100 110 120
ATRLKRTARR RYTRRKNRIC YLQEIFSNEK AKVDDSFHHR LEESFLVEED KKHERHPIFG

130 140 150 160 170 180
NIVDEVAYHE KYPTIYHLRK KLVDSTDKAD LRLIYLALAH MIKFRGHFLI EGDLPDNDSD

190 200 210 220 230 240
VDKLFILQVQ TYNQLFEENP INASGVDAKA ILSARLSKSR RLENLIAQLP GEKKNGLFGN

250 260 270 280 290 300
LIALSLGLTP NFKSNFDLAE DAKLQLSKDT YDDDLNLLA QIGDQYADLF LAAKNLSDAI

310 320 330 340 350 360

```

LLSDILRVNT EITKAPLSAS MIKRYDEHHQ DLTLLKALVR QQLPEKYKEI FFDQSKNGYA

370 380 390 400 410 420

GYIDGGASQE EFYKFIKPIL EKMDGTEELL VKLNREDLLR KQRTFDNGSI PHQIHLGELH

430 440 450 460 470 480

AILRRQEDFY PFLKDNREKI EKILTRIPY YVGPLARGNS RFAWMTRKSE ETITPWNFEE

490 500 510 520 530 540

VVDKGASAQs FIERMTNFDK NLPNEKVLPK HSLLEYFTV YNELTKVKYV TEGMRKPAFL

550 560 570 580 590 600

SGEQKKAIVD LLFKTNRKVT VKQLKEDYFK KIECFDSVEI SGVEDRFNAS LGTYHDLLKI

610 620 630 640 650 660

IKDKDFLDNE ENEDILEDIV LTLTLFEDRE MIEERLKTYA HLFDDKVMKQ LKRRRYTGWG

670 680 690 700 710 720

RLSRKLINGI RDKQSGKTIL DFLKSDGFAN RNFMQLIHDD SLTFKEDIQK AQVSGQGDSL

730 740 750 760 770 780

HEHIANLAGS PAIKKGILQT VKVDELVKV MGRHKPENIV IEMARENQTT QKGQKNSRER

790 800 810 820 830 840

MKRIEEGIKE LGSQILKEHP VENTQLQNEK LYLYYLQNGR DMYVDQELDI NRLSDYDVDAA

850 860 870 880 890 900

IVPQSFLKDD SIDNKVLTRS DKNRGKSDNV PSEEVVKKMK NYWRQLLNAK LITQRKFDNL

910 920 930 940 950 960

TKAERGGLSE LDKAGFIKRQ LVETRQITKH VAQILDSRMN TKYDENDKLI REVKVITLKS

970 980 990 1000 1010 1020

KLVSDFRKDF QFYKVBREINN YHHAHDAYLN AVVGTAIIK YPKLESEFVY GDYKVYDVRK

1030 1040 1050 1060 1070 1080

MAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGEIRKR PLIETNGETG EIVWDKGRDF

1090 1100 1110 1120 1130 1140

ATVRKVL SMP QVNIVKKTEV QTGGFSKESI LPKRNSDKLI ARKKDWDPKK YGGFDSPTVA

1150 1160 1170 1180 1190 1200
YSVLVVAKVE KGKSKKLKSV KELLGITIME RSSFEKNPID FLEAKGYKEV KKDIIKLPK

1210 1220 1230 1240 1250 1260
YSLFELENGR KRMLASAGEL QKGNELALPS KYVNFLYLAS HYEKLKGSPE
DNEQKQLFVE

1270 1280 1290 1300 1310 1320
QHKHYLDEII EQISEFSKRV ILADANLDKV LSAYNKHRDK PIREQAENII HLFTLTNLGA

1330 1340 1350 1360
PAAFKYFDTT IDRKYRSTK EVLDATLIHQ SITGLYETRI DLSQLGGD (SEQ ID NO:33)

[0088] Em algumas modalidades, a nuclease Cas9 usada aqui é pelo menos cerca de 50% idêntica à sequência de Cas9 de *S. pyogenes*, ou seja, pelo menos 50% idêntica à SEQ ID NO: 33. Em algumas modalidades, as sequências de nucleotídeos são cerca de 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% ou 100% idênticas à SEQ ID NO: 33. Em algumas modalidades, quaisquer diferenças a partir de SEQ ID NO: 33 estão em regiões não conservadas, tal como identificado por meio de alinhamento de sequência das sequências apresentadas em Chylinski *et al.*, RNA Biology 10: 5, 1-12; 2013 (por exemplo, na figura 1 suplementar e tabela 1 suplementar); Esvelt *et al.*, Methods Nat. Novembro de 2013; 10 (11): 1116-1121 e Fonfara *et al.*, Nucl. Acids Res. (2014) 42 (4): 2577-2590. [cabeçalho Epub de impressão, 22 de Novembro de 2013] doi:10.1093/nar/gkt1074.

[0089] Para determinar a identidade percentual de duas sequências, as sequências são alinhadas para fins de comparação ótima (lacunas são introduzidas em uma ou ambas de uma primeira e uma segunda sequências de aminoácidos ou ácidos nucleicos, conforme necessário para um alinhamento ótimo, e sequências não homólogas podem ser desconsideradas para fins de comparação). O

comprimento de uma sequência de referência alinhada para fins de comparação é de pelo menos 50% (em algumas modalidades, cerca de 50%, 55%, 60%, 65%, 70%, 75%, 85%, 90%, 95%, ou 100 % do comprimento da sequência de referência são alinhados). Os nucleotídeos ou resíduos nas posições correspondentes são, então, comparados. Quando uma posição na primeira sequência é ocupada pelo mesmo nucleotídeo ou resíduo na posição correspondente na segunda sequência, então, as moléculas são idênticas nesta posição. A percentagem de identidade entre as duas sequências é uma função do número de posições idênticas compartilhadas pelas sequências, levando-se em conta o número de lacunas e o comprimento de cada lacuna que precisa de ser introduzida para alinhamento ótimo das duas sequências.

[0090] A comparação de sequências e determinação da identidade percentual entre duas sequências podem ser realizadas usando um algoritmo matemático. Para fins do presente Pedido, a identidade percentual entre duas sequências de aminoácidos é determinada através do algoritmo de Needleman e Wunsch ((1970) J. Mol Biol 48: 444-453) que foi incorporado no programa GAP no GCG do pacote de software, usando uma matriz de pontuação Blossum 62 com uma penalidade por lacuna de 12, uma penalidade por extensão de lacuna de 4 e uma penalidade por lacuna de desvio de quadro de 5.

Cas9-HFD

[0091] Cas9-HFD são descritas nos Pedidos Provisórios de Patente dos Estados Unidos USSN 61/799.647, depositado em 15 de Março de 2013, USSN 61/838.148, depositado em 21 de Junho de 2013 e Pedido Internacional PCT N° PCT/US14/27335, todos os quais são aqui incorporados por referência na íntegra.

[0092] Cas9-HFD são criadas por fusão de um domínio heterólogo funcional (por exemplo, um domínio de ativação transcricional, por

exemplo, a partir de VP64 ou NF- κ B p65), com a extremidade N-terminal ou C-terminal de uma proteína Cas9 cataliticamente inativa (dCas9). No presente caso, conforme citado acima, o dCas9 pode ser de qualquer espécie mas é, de preferência, de *S. pyogenes*. Em algumas modalidades, a Cas9 contém mutações nos resíduos D10 e H840, por exemplo, D10N/D10A e H840A/H840N/H840Y, para tornar a porção de nuclease da proteína cataliticamente inativa, por exemplo, conforme mostrado em SEQ ID NO: 33, supra.

[0093] Os domínios de ativação de transcrição podem ser fundidos no N-término ou C-término de Cas9. Além disso, embora a presente descrição exemplifique domínios de ativação de transcrição, outros domínios funcionais heterólogos (por exemplo, repressores transcricionais (por exemplo, KRAB, ERD, SID e outros, por exemplo, aminoácidos 473-530 do domínio repressor (ERD) de fator repressor ETS2 (FER), aminoácidos 1-97 do domínio KRAB de KOX1 ou aminoácidos 1-36 do domínio de interação mSIN3 de Mad (SID); vide Beerli *et al.*, PNAS USA 95: 14628-14633 (1998)) ou silenciadores, tais como proteína 1 de Heterocromatina (HP1, também conhecida como swi6), por exemplo, HP1 α ou HP1 β ; proteínas ou peptídeos que podem recrutar RNAs não codificadores longos (lncRNAs) fundidos com uma sequência de ligação a RNA fixa, tais como aqueles ligados pela proteína de revestimento MS2, endorribonuclease Csy4 ou a proteína N lambda; enzimas que modificam o estado de metilação de DNA (por exemplo, DNA metiltransferase (DNMT) ou proteínas TET); ou enzimas que modificam subunidades de histona (por exemplo, histona acetiltransferases (HAT), histona desacetilases (HDAC), histona metiltransferases (por exemplo, para metilação dos resíduos de lisina ou arginina) ou desmetilases histona (por exemplo, para desmetilação de resíduos de lisina ou arginina)), conforme conhecido na técnica também podem ser usados. Um determinado número de sequências de

tais domínios são conhecidas na técnica, por exemplo, um domínio que catalisa a hidroxilação de citosinas metiladas em DNA. Proteínas exemplificativas incluem a família de enzimas Ten-Eleven-Translocation (TET) 1-3, as quais convertem 5-metilcitosina (5-MC) em 5-hidroximetilcitosina (5-hmC) em DNA.

[0094] Sequências para TET1-3 humana são conhecidas na técnica e são mostradas na tabela a seguir:

Gene	Nos. Acesso GenBank	
	Aminoácido	Ácido Nucleico
TET1	NP_085128.2	NM_030625.2
TET2*	NP_001120680.1 (var 1)	NM_001127208.2
	NP_060098.3 (var 2)	NM_017628.4
TET3	NP_659430.1	NM_144993.1

*Variante (1) representa o transcrito mais longo e codifica a isoforma mais longa (a). Variante (2) difere quanto a 5' UTR e 3' UTR e a sequência codificadora comparado com a variante 1. A isoforma (b) resultante é mais curta e tem um C-término distinto comparado com a isoforma a.

[0095] Em algumas modalidades, a totalidade ou parte da sequência de comprimento completo do domínio catalítico pode ser incluída, por exemplo, um módulo catalítico que compreende a extensão rica em cisteína e o domínio 2OGFeDO codificado por 7 éxons altamente conservados, por exemplo, o domínio catalítico Tet1 compreendendo os aminoácidos 1580-2052, TET2 compreendendo os aminoácidos 1290-1905 e Tet3 compreendendo os aminoácidos 966-1678. Vide, por exemplo, a Figura 1 de Iyer *et al.*, Cell Cycle. 01 de Junho de 2009; 8 (11): 1698-710. Epub de 27 de Junho de 2009, para um alinhamento que ilustra os resíduos catalíticos-chave em todas as três proteínas Tet e os materiais complementares do mesmo (disponível no [site ftp
ftp.ncbi.nih.gov/pub/aravind/DONS/supplementary_material_DONS.html](http://ftp.ncbi.nih.gov/pub/aravind/DONS/supplementary_material_DONS.html)) para sequências de comprimento total (vide, por exemplo, seq 2c);

em algumas modalidades, a sequência inclui os aminoácidos 1418-2136 de Tet1 ou a região correspondente em Tet2/3.

[0096] Outros módulos catalíticos podem ser a partir das proteínas identificadas por Iyer *et al.*, 2009.

[0097] Em algumas modalidades, o domínio funcional heterólogo é um grupo de ancoragem biológica e compreende toda ou parte (por exemplo, domínio de ligação a DNA) da proteína de revestimento MS2, endorribonuclease Csy4 ou a proteína lambda N. Estas proteínas podem ser usadas para recrutar moléculas de RNA que contêm uma estrutura de haste-alça específica para uma região especificada pelas sequências de objetivação de gRNA de dCas9. Por exemplo, um dCas9 fundido a proteína de revestimento MS2, endorribonuclease Csy4 ou lambda N pode ser usado para recrutar um RNA não codificador longo (lncRNA), tal como XIST ou HOTAIR; vide, por exemplo, Keryer-Bibens *et al.*, Biol. Celular 100: 125-138 (2008), que está ligado à sequência de ligação de Csy4, MS2 ou lambda N. Alternativamente, a sequência de ligação de proteína Csy4, MS2 ou lambda N pode ser ligada a uma outra proteína, por exemplo, conforme descrito em Keryer-Bibens *et al.*, supra, e a proteína pode ser orientada ao sítio de ligação dCas9 usando os métodos e composições descritos aqui. Em algumas modalidades, a Csy4 é cataliticamente inativa.

[0098] Em algumas modalidades, as proteínas de fusão incluem um ligante entre dCas9 e os domínios funcionais heterólogos. Os ligantes que podem ser usados nestas proteínas de fusão (ou entre proteínas de fusão em uma estrutura concatenada) podem incluir qualquer sequência que não interfira com a função das proteínas de fusão. Em modalidades preferidas, os ligantes são curtos, por exemplo, 2-20 aminoácidos e são, tipicamente, flexíveis (isto é, compreendendo aminoácidos com um elevado grau de liberdade, tais como glicina, alanina e serina). Em algumas modalidades, o ligante compreende uma

ou mais unidades compostas de GGGS (SEQ ID NO: 34) ou GGGGS (SEQ ID NO: 35), por exemplo, duas, três, quatro ou mais repetições da unidade GGGS (SEQ ID NO: 34) ou GGGGS (SEQ ID NO: 35). Outras sequências ligantes também podem ser usadas.

Sistemas de Expressão

[0099] De modo a usar os RNA guia descritos, pode ser desejável expressá-los a partir de um ácido nucleico que os codifica. Isto pode ser realizado em uma variedade de maneiras. Por exemplo, o ácido nucleico que codifica o RNA guia pode ser clonado em um vetor intermediário para transformação em células procariotas ou eucariotas para replicação e/ou expressão. Vetores intermediários são, tipicamente, vetores procariotas, por exemplo, plasmídeos ou vetores de vaivém, ou vetores de inseto, para armazenamento ou manipulação do ácido nucleico que codifica o RNA guia para produção do RNA guia. O ácido nucleico que codifica o RNA guia também pode ser clonado em um vetor de expressão, para administração a uma célula de planta, célula animal, de preferência uma célula de mamífero ou uma célula humana, célula fúngica, célula bacteriana ou célula de protozoário.

[0100] Para obter expressão, uma sequência que codifica um RNA guia é, tipicamente, subclonada em um vetor de expressão que contém um promotor para comandar a transcrição direta. Promotores bacterianos e eucariotas adequados são bem conhecidos na técnica e descritos, por exemplo, em Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (2001, 3ª ed.); Gene Transfer and Expression: A Laboratory Manual (1990); e Current Protocols in Molecular Biology (Ausubel *et al.*, eds., 2010). Sistemas de expressão bacterianos para expressar a proteína manipulada estão disponíveis, por exemplo, em *E. coli*, *Bacillus sp.* e *Salmonella* (Paiva *et al.*, 1983, Gene. 22: 229-235). Kits para tais sistemas de expressão estão comercialmente disponíveis. Sistemas de expressão eucariotas para células de mamífero, células de

levedura e células de inseto são bem conhecidos na técnica e também estão comercialmente disponíveis.

[0101] O promotor usado para expressão direta de um ácido nucleico depende da aplicação particular. Por exemplo, um promotor constitutivo forte é, tipicamente, usado para expressão e purificação de proteínas de fusão. Em contraste, quando o RNA guia tem de ser administrado *in vivo* para regulação de genes, um promotor constitutivo ou um promotor indutível pode ser usado, dependendo o uso particular do RNA guia. Além disso, um promotor preferido para administração do RNA guia pode ser um promotor fraco, tal como TK de HSV, ou um promotor tendo atividade similar. O promotor também pode incluir elementos que respondem à transativação, por exemplo, elementos de resposta à hipóxia, elementos de resposta a Gal4, elementos de resposta ao repressor lac e sistemas de controle de pequenas molécula, tais como os sistemas regulados por tetraciclina e o sistema RU-486 (vide, por exemplo, Gossen & Bujard, 1992, Proc. Natl. Acad. Sci. USA, 89: 5547; Oligino *et al.*, 1998, Gene Ther., 5: 491-496; Wang *et al.*, 1997, Gene Ther., 4: 432-441; Neering *et al.*, 1996, Blood, 88: 1147-55; e Rendahl *et al.*, 1998, Nat. Biotechnol., 16: 757-761).

[0102] Além do promotor, o vetor de expressão contém, tipicamente, uma unidade de transcrição ou cassete de expressão que contém todos os elementos adicionais necessários para expressão do ácido nucleico nas células hospedeiras, quer procariotas ou eucariotas. Uma cassete de expressão típico contém, assim, um promotor operativamente ligado, por exemplo, à sequência de ácido nucleico que codifica o gRNA, e todos os sinais necessários, por exemplo, para poliadenilação eficaz do transcrito, sítios de término de transcrição, sítios de ligação ao ribossomo ou sítios de término de tradução. Elementos adicionais do cassete podem incluir, por exemplo, potencializadores e sinais intrônicos de *splicing* heterólogos.

[0103] O vetor de expressão particular usado para transportar a informação genética para a célula é selecionado tendo em vista o uso pretendido do gRNA, por exemplo, expressão em plantas, animais, bactérias, fungos, protozoários, etc. Vetores de expressão bacterianos padrões incluem plasmídeos, tais como plasmídeos com base em pBR322, pSKF, pET23D, e sistemas de expressão fusão-tag comercialmente disponíveis, tais como GST e LacZ.

[0104] Vetores de expressão contendo elementos reguladores de vírus eucariotas são, muitas vezes, usados em vetores de expressão eucariotas, por exemplo, vetores SV40, vetores de papiloma vírus e vetores derivados do vírus de Epstein-Barr. Outros vetores eucariotas exemplificativos incluem pMSG, pAV009/A+, pMTO10/A+, pMAMneo-5, pDSVE de baculovírus e qualquer outro vetor que permita a expressão de proteínas sob o comando do promotor precoce de SV40, promotor tardio de SV40, promotor de metalotioneína, promotor do vírus de tumor mamário de murino, promotor do vírus do sarcoma de Rous, promotor de poliedrina ou outros promotores mostrados eficazes para expressão em células eucariotas.

[0105] Os vetores para expressar os RNA guia podem incluir promotores de RNA Pol III para acionar a expressão dos RNAs guia, por exemplo, os promotores H1, U6 ou 7SK. Estes promotores humanos permitem a expressão de gRNAs em células de mamífero após transfecção de plasmídeo. Alternativamente, um promotor T7 pode ser usado, por exemplo, para transcrição *in vitro*, e o RNA pode ser transcrito *in vitro* e purificado. Vetores adequados para expressão de RNAs curtos, por exemplo, siRNAs, shRNAs ou outros RNAs pequenos, podem ser usados.

[0106] Alguns sistemas de expressão têm marcadores para seleção de linhagens de células estavelmente transfectadas, tais como timidina quinase, higromicina B fosfotransferase e di-hidrofolato

redutase. Sistemas de expressão de elevado rendimento também são adequados, tal como usando um vetor de baculovírus em células de inseto, com a sequência que codifica o gRNA sob a direção do promotor de poliedrina ou outros promotores fortes de baculovírus.

[0107] Os elementos que são, tipicamente, incluídos nos vetores de expressão também incluem um replicon que funciona em *E. coli*, um gene que codifica resistência a antibiótico para permitir seleção de bactérias que abrigam os plasmídeos recombinantes e sítios de restrição únicos em regiões não essenciais do plasmídeo para permitir inserção de sequências recombinante.

[0108] Métodos de transfecção convencionais são usados para produzir linhagens de células bacterianas, de mamífero, levedura ou inseto que expressam grandes quantidades de proteína as quais são, então, purificadas usando técnicas convencionais (vide, por exemplo, Colley *et al.*, 1989, J. Biol. Chem., 264: 17619-22; Guide to Protein Purification, em Methods in Enzymology, vol. 182 (Deutscher, ed., 1990)). A transformação de células eucariotas e procariotas é realizada de acordo com técnicas convencionais (vide, por exemplo, Morrison, 1977, J. Bacteriol 132: 349-351; Clark-Curtiss Curtiss &, Methods in Enzymology 101: 347-362 (Wu *et al.*, eds, 1983).

[0109] Qualquer um dos procedimentos conhecidos para introdução de sequências de nucleotídeos estranhas em células hospedeiras pode ser usado. Estes incluem o uso de transfecção com fosfato de cálcio, polibreno, fusão de protoplastos, eletroporação, nucleofecção, lipossomas, microinjeção, DNA nu, vetores de plasmídeo, vetores virais, tanto epissômicos quanto integrativos, e qualquer um dos outros métodos bem conhecidos para introdução de DNA genômico clonado, cDNA, DNA sintético ou outro material genético estranho em uma célula hospedeira (vide, por exemplo, Sambrook *et al.*, supra). É necessário apenas que o procedimento de engenharia genética usado

em particular seja capaz de introduzir com sucesso pelo menos um gene na célula hospedeira capaz de expressar o gRNA.

[0110] A presente invenção inclui os vetores e células que compreendem os vetores.

EXEMPLOS

[0111] A invenção é ainda descrita nos exemplos a seguir, os quais não limitam o âmbito da invenção descrita nas reivindicações.

Exemplo 1. Avaliação de especificidade de endonucleases RNA-orientadas

[0112] Nucleases RNA-orientadas CRISPR (RGNs) surgiram rapidamente como uma plataforma fácil e eficiente para edição de genoma. Este exemplo descreve o uso de um ensaio de repórter com base em células humanas para caracterizar a clivagem fora do alvo de RGNs com base em Cas9.

Materiais e Métodos

[0113] Os materiais e métodos a seguir foram usados no Exemplo 1.

Construção de RNAs guia

[0114] Oligonucleotídeos de DNA (**Tabela A**) que abrigam 20 sequências de nt variáveis para objetivação de Cas9 foram recozidos para gerar fragmentos de DNA fita dupla curtos com saliências de 4 pb compatíveis com ligação em plasmídeo pMLM3636 digerido com *BsmBI*. A clonagem de tais oligonucleotídeos recozidos gera plasmídeos que codificam um RNA guia fita simples +103 quimérico com 20 nucleotídeos 5' variáveis sob a expressão de um promotor U6 (Hwang *et al.*, Nat Biotechnol 31, 227-229 (2013); Mali *et al.*, Science 339, 823-826 (2013)). pMLM3636 e o plasmídeo de expressão pJDS246 (que codifica uma versão com códons otimizados de Cas9) usado neste estudo estão disponíveis através do serviço de distribuição de plasmídeo sem fins lucrativos Addgene (addgene.org/crispr-cas).

Tabela A																								
Posição de Sequência Alvo de gRNA																				Oligos para Geração de plasmídeo de expressão de gRNA				
Sítio 1 de EGFP alvo																								
2	1	1	1	1	1	1	1	1	1	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')	#	oligonucleotídeo 2 (5' a 3')	#
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1		ACACCGGGCACGG GCAGCTTGCCGGG	36	AAAACCCGGCAAG CTGCCCGTGCCCG	236
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	37	AAAACCCGGCAAG CTGCCCGTGCCCG	237
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	c		ACACCGGGCACGG GCAGCTTGCCCGG	38	AAAACCCGGCAAG CTGCCCGTGCCCG	238
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	c	G		ACACCGGGCACGG GCAGCTTGCCCGG	39	AAAACCCGGCAAG CTGCCCGTGCCCG	239
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	g	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	40	AAAACCCGGCAAG CTGCCCGTGCCCG	240
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	g	g	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	41	AAAACCCGGCAAG CTGCCCGTGCCCG	241
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	c	C	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	42	AAAACCCGGCAAG CTGCCCGTGCCCG	242
G	G	G	C	A	C	G	G	G	C	A	G	C	T	a	G	C	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	43	AAAACCCGGCAAG CTGCCCGTGCCCG	243
G	G	G	C	A	C	G	G	G	C	A	G	g	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	44	AAAACCCGGCAAG CTGCCCGTGCCCG	244
G	G	G	C	A	C	G	G	G	C	A	c	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	45	AAAACCCGGCAAG GTGCCCGTGCCCG	245
G	G	G	C	A	C	G	G	G	C	t	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCTGCTTGCCGGG	46	AAAACCCGGCAAG CAGCCGTGCCCG	246
G	G	G	C	A	C	G	G	G	g	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GGAGCTTGCCGGG	47	AAAACCCGGCAAG CTCCCGTGCCCG	247
G	G	G	C	A	C	G	G	c	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	48	AAAACCCGGCAAG CTGCCGTGCCCG	248
G	G	G	C	A	C	G	c	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	49	AAAACCCGGCAAG CTGCCGTGCCCG	249
G	G	G	C	A	C	c	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	50	AAAACCCGGCAAG CTGCCGTGCCCG	250
G	G	G	C	A	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	51	AAAACCCGGCAAG CTGCCCTGCCCG	251
G	G	G	C	t	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	52	AAAACCCGGCAAG CTGCCGTGCCCG	252
G	G	G	g	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	53	AAAACCCGGCAAG CTGCCGTGCCCG	253
G	G	c	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	54	AAAACCCGGCAAG CTGCCGTGCCCG	254
G	c	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGCTTGCCGGG	55	AAAACCCGGCAAG CTGCCGTGCCCG	255
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	c	c		ACACCGGGCACGG GCAGCTTGCCCGG	56	AAAACCCGGCAAG CTGCCGTGCCCG	256
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	g	g	G	G		ACACCGGGCACGG GCAGCTTGCGGGG	57	AAAACCCCGCAAGC TGCCCGTGCCCG	257
G	G	G	C	A	C	G	G	G	C	A	G	C	T	a	c	C	C	G	G		ACACCGGGCACGG GCAGTACCCGGG	58	AAAACCCGGGTAG CTGCCGTGCCCG	258
G	G	G	C	A	C	G	G	G	C	A	G	g	a	T	G	C	C	G	G		ACACCGGGCACGG GCAGGATCCGGG	59	AAAACCCGGCATCC TGCCCGTGCCCG	259
G	G	G	C	A	C	G	G	G	C	t	c	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCTCTTGCCGGG	60	AAAACCCGGCAAG GAGCCGTGCCCG	260
G	G	G	C	A	C	G	G	c	g	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG CGAGTTGCCGGG	61	AAAACCCGGCAAG CTGCCGTGCCCG	261
G	G	G	C	A	C	c	c	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGTTGCCGGG	62	AAAACCCGGCAAG CTGCCGTGCCCG	262
G	G	G	C	t	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGGTGGG GCAGTTGCCGGG	63	AAAACCCGGCAAG CTGCCAGCCCG	263
G	G	c	g	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGTTGCCGGG	64	AAAACCCGGCAAG CTGCCGTGCCCG	264
G	c	c	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGTTGCCGGG	65	AAAACCCGGCAAG CTGCCGTGCCCG	265
G	c	C	g	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGCACGG GCAGTTGCCGGG	66	AAAACCCGGCAAG CTGCCGTGCCCG	266
G	c	c	g	t	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGCGTCGG GCAGTTGCCGGG	67	AAAACCCGGCAAG CTGCCGTGCCCG	267
G	c	c	g	t	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGCGTGGG GCAGTTGCCGGG	68	AAAACCCGGCAAG CTGCCGTGCCCG	268
G	c	c	g	t	g	c	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGCGTGGG GCAGCGCGTGGG	69	AAAACCCGGCAAG CTGCCGTGCCCG	269
G	c	c	g	t	g	c	G	C	A	G	C	T	T	G	C	C	G	G	G		ACACCGCGTGGG GCAGCGCGTGGG	70	AAAACCCGGCAAG	270

Tabela A																																					
Posição de Sequência Alvo de gRNA																		Oligos para Geração de plasmídeo de expressão de gRNA																			
																		GCAGCTTGCCGGG		CTGCCCGTGCCCG																	
G	c	c	g	t	g	c	c	c	C	A	G	C	T	T	G	C	C	G	G		ACACCGCGTGCC	71.	AAAACCCGGCAAG	268													
																				CTGCCCGTGCCCG																	
G	c	c	g	t	g	c	c	c	g	A	G	C	T	T	G	C	C	G	G		ACACCGCGTGCC	72.	AAAACCCGGCAAG	268													
																				CGAGCTTGCCGGG		CTGCCCGTGCCCG															
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	g	G	c		ACACGGGCACGG	73.	AAAACCCGGCAAG	268													
																				CGAGCTTGCCGGG		CTGCCCGTGCCCG															
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	c	C	g	G	G		ACACGGGCACGG	74.	AAAACCCGGCAAG	268													
																				GCAGCTTCCGGGG		CTGCCCGTGCCCG															
G	G	G	C	A	C	G	G	G	C	A	G	C	a	T	G	C	g	G	G		ACACGGGCACGG	75.	AAAACCCGGCATGC	268													
																				GCAGCATGCGGGG		TGCCCGTGCCCG															
G	G	G	C	A	C	G	G	G	C	A	c	C	T	T	G	C	g	G	G		ACACGGGCACGG	76.	AAAACCCGGCAAG	270													
																				GCACCTTGCCGGG		GTGCCCGTGCCCG															
G	G	G	C	A	C	G	G	G	g	A	G	C	T	T	G	C	g	G	G		ACACGGGCACGG	77.	AAAACCCGGCAAG	270													
																				GGAGCTTGCGGGG		CTCCCGTGCCCG															
G	G	G	C	A	C	G	c	G	C	A	G	C	T	T	G	C	g	G	G		ACACGGGCACGC	78.	AAAACCCGGCAAG	270													
																				GCAGCTTGCCGGG		CTGCGCGTGCCCG															
G	G	G	C	A	g	G	G	G	C	A	G	C	T	T	G	C	g	G	G		ACACCGGCAGGG	79.	AAAACCCGGCAAG	270													
																				GCAGCTTGCGGGG		CTGCCCTTGCCCG															
G	G	G	g	A	C	G	G	G	C	A	G	C	T	T	G	C	g	G	G		ACACGGGGACGG	80.	AAAACCCGGCAAG	270													
																				GCAGCTTGCGGGG		CTGCCCTGCCCCG															
G	c	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	g	G	G		ACACCGGCACGG	81.	AAAACCCGGCAAG	270													
																				GCAGCTTGCGGGG		CTGCCCTGTGCGG															
G	G	G	C	A	C	G	G	G	g	A	G	C	T	T	G	C	C	G	c		ACACGGGCACGG	82.	AAAACCGGGCAAG	270													
																				GGAGCTTGCCCGG		CTCCCGTGCCCG															
G	G	G	C	A	C	G	G	G	g	A	G	C	T	T	c	C	C	G	G		ACACGGGCACGG	83.	AAAACCGGGGAAG	270													
																				GGAGCTTCCCGGG		CTCCCGTGCCCG															
G	G	G	C	A	C	G	G	G	g	A	G	C	a	T	G	C	C	G	G		ACACGGGCACGG	84.	AAAACCGGGCATG	270													
																				GGAGCATGCCGGG		CTCCCGTGCCCG															
G	G	G	C	A	C	G	G	G	g	A	c	C	T	T	G	C	C	G	G		ACACGGGCACGG	85.	AAAACCGGGCAAG	270													
																				GGACCTTGCCGGG		GTCCCGTGCCCG															
G	G	G	C	A	C	G	c	G	g	A	G	C	T	T	G	C	C	G	G		ACACGGGCACGC	86.	AAAACCGGGCAAG	268													
																				GGAGCTTGCCGGG		CTCCCGTGCCCG															
G	G	G	C	A	g	G	G	G	g	A	G	C	T	T	G	C	C	G	G		ACACCGGGACGG	87.	AAAACCGGGCAAG	268													
																				GGAGCTTGCCGGG		CTCCCGTGCCCG															
G	G	G	g	A	C	G	G	G	g	A	G	C	T	T	G	C	C	G	G		ACACGGGGACGG	88.	AAAACCGGGCAAG	268													
																				GGAGCTTGCCGGG		CTCCCGTGCCCG															
G	c	G	C	A	C	G	G	G	g	A	G	C	T	T	G	C	C	G	G		ACACCGGCACGG	89.	AAAACCGGGCAAG	268													
																				GGAGCTTGCCGGG		CTCCCGTGCGCG															
G	c	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	c		ACACCGGCACGG	90.	AAAACCGGGCAAG	268													
																				GCAGCTTGCCGCG		CTGCCGTGCGCG															
G	c	G	C	A	C	G	G	G	C	A	G	C	T	T	c	C	C	G	G		ACACCGGCACGG	91.	AAAACCGGGGAAG	268													
																				GCAGCTTCCCGGG		CTGCCGTGCGCG															
G	c	G	C	A	C	G	G	G	C	A	G	C	a	T	G	C	C	G	G		ACACCGGCACGG	92.	AAAACCGGGCATG	268													
																				GCAGCATGCCGGG		CTGCCCGTGCGCG															
G	c	G	C	A	C	G	G	G	C	A	c	C	T	T	G	C	C	G	G		ACACCGGCACGG	93.	AAAACCGGGCAAG	268													
																				GCACCTTGCCGGG		GTGCCGTGCGCG															
G	c	G	C	A	C	G	c	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGCACGC	94.	AAAACCGGGCAAG	268													
																				GCAGCTTGCCGGG		CTGCCGTGCGCG															
G	c	G	C	A	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGCAGGG	95.	AAAACCGGGCAAG	268													
																				GCAGCTTGCCGGG		CTGCCCTGCGCG															
G	c	G	g	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G		ACACCGGGACGG	96.	AAAACCGGGCAAG	268													
																				GCAGCTTGCCGGG		CTGCCCTGCCCG															
Sítio 2 de EGFP alvo																																					
2	1	1	1	1	1	1	1	1	1	1	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')				oligonucleotídeo 2 (5' a 3')											
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1			ACACCGATGCCGTT	97.	AAAACACAAGCAGA		AGAACGGCATCG											
																						CTTCTGCTTGTG		AAAACACAAGCAGA		AGAACGGCATCG											
G	A	T	G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T			ACACCGATGCCGTT	98.	AAAACACAAGCAGA		AGAACGGCATCG											
																						CTTCTGCTTGAG		AAAACACAAGCAGA		AGAACGGCATCG											
G	A	T	G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	a			ACACCGATGCCGTT	99.	AAAACACAAGCAGA		AGAACGGCATCG											
																						CTTCTGCTCTTG		AAAACACAAGCAGA		AGAACGGCATCG											
G	A	T	G	C	C	G	T	T	C	T	T	C	T	G	C	T	a	G	T			ACACCGATGCCGTT	100.	AAAACACAAGCAGA		AGAACGGCATCG											
																						CTTCTGCTAGTG		AAAACACAAGCAGA		AGAACGGCATCG											
G	A	T	G	C	C	G	T	T	C	T	T	C	T	G	C	a	T	G	T			ACACCGATGCCGTT	101.	AAAACACAAGCAGA		AGAACGGCATCG											
																						CTTCTGCATGTG		AAAACACAAGCAGA		AGAACGGCATCG											
G	A	T	G	C	C	G	T	T	C	T	T	C	T	G	g	T	T	G	T			ACACCGATGCCGTT	102.	AAAACACAAGCAGA		AGAACGGCATCG											
																						CTTCTGTTTGTG		AAAACACAAGCAGA		AGAACGGCATCG											
G	A	T	G	C	C	G	T	T	C	T	T	C	T	c	C	T	T	G	T			ACACCGATGCCGTT	103.	AAAACACAAGCAGA		AGAACGGCATCG											
																						CTTCTCCTTGTG		AAAACACAAGCAGA		AGAACGGCATCG											
G	A	T	G	C	C	G	T	T	C	T	T	C	a	G	C	T	T	G	T			ACACCGATGCCGTT	104.	AAAACACAAGCAGA		AGAACGGCATCG											
																						CTTCAGCTTGTG		AAAACACAAGCAGA		AGAACGGCATCG											

Tabela A																							
Posição de Sequência Alvo de gRNA																		Oligos para Geração de plasmídeo de expressão de gRNA					
G	A	T	G	C	C	G	T	T	C	T	T	g	T	G	C	T	T	G	T	ACACCGATGCCGTT CTTGTGCTTGTG	108	AAAACACAAGCAGA AGAACGGCATCG	298
G	A	T	G	C	C	G	T	T	C	T	a	C	T	G	C	T	T	G	T	ACACCGATGCCGTT CTACTGCTTGTG	109	AAAACACAAGCAGA AGAACGGCATCG	300
G	A	T	G	C	C	G	T	T	C	a	T	C	T	G	C	T	T	G	T	ACACCGATGCCGTT CATCTGCTTGTG	107	AAAACACAAGCAGA AGAACGGCATCG	301
G	A	T	G	C	C	G	T	T	g	T	T	C	T	G	C	T	T	G	T	ACACCGATGCCGTT GTTCTGCTTGTG	108	AAAACACAAGCAGA AGAACGGCATCG	302
G	A	T	G	C	C	G	T	a	C	T	T	C	T	G	C	T	T	G	T	ACACCGATGCCGTA CTTCTGCTTGTG	109	AAAACACAAGCAGA AGAACGGCATCG	303
G	A	T	G	C	C	G	a	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGATGCCGAT CTTCTGCTTGTG	110	AAAACACAAGCAGA AGAACGGCATCG	304
G	A	T	G	C	C	c	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGATGCCCTT CTTCTGCTTGTG	111	AAAACACAAGCAGA AGAACGGCATCG	305
G	A	T	G	C	g	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGATGCCGTT CTTCTGCTTGTG	112	AAAACACAAGCAGA AGAACGGCATCG	306
G	A	T	G	g	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGATGGCGTT CTTCTGCTTGTG	113	AAAACACAAGCAGA AGAACGGCATCG	307
G	A	T	c	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGATCCCGTT CTTCTGCTTGTG	114	AAAACACAAGCAGA AGAACGGCATCG	308
G	A	t	G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGAAGCCGTT CTTCTGCTTGTG	115	AAAACACAAGCAGA AGAACGGCATCG	309
G	a	T	G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGTTGCCGTT CTTCTGCTTGTG	116	AAAACACAAGCAGA AGAACGGCATCG	310
G	A	T	G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	c	a	ACACCGATGCCGTT CTTCTGCTTCAG	117	AAAACACAAGCAGA AGAACGGCATCG	311
G	A	T	G	C	C	G	T	T	C	T	T	C	T	G	C	a	a	G	T	ACACCGATGCCGTT CTTCTGCAAGTG	118	AAAACACAAGCAGA AGAACGGCATCG	312
G	A	T	G	C	C	G	T	T	C	T	T	C	T	c	g	T	T	G	T	ACACCGATGCCGTT CTTCTGCTTGTG	119	AAAACACAAGCAGA AGAACGGCATCG	313
G	A	T	G	C	C	G	T	T	C	T	T	g	a	G	C	T	T	G	T	ACACCGATGCCGTT CTTGAGCTTGTG	120	AAAACACAAGCTCA AGAACGGCATCG	314
G	A	T	G	C	C	G	T	T	C	a	a	C	T	G	C	T	T	G	T	ACACCGATGCCGTT CAACTGCTTGTG	121	AAAACACAAGCAGA AGAACGGCATCG	315
G	A	T	G	C	C	G	T	a	g	T	T	C	T	G	C	T	T	G	T	ACACCGATGCCGTA GTTCTGCTTGTG	122	AAAACACAAGCAGA AGAACGGCATCG	316
G	A	T	G	C	C	c	a	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGATGCCCAT CTTCTGCTTGTG	123	AAAACACAAGCAGA AGAACGGCATCG	317
G	A	T	G	g	g	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGATGGGGT TCTTCTGCTTGTG	124	AAAACACAAGCAGA AGAACGGCATCG	318
G	A	a	c	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGAACCCGTT CTTCTGCTTGTG	125	AAAACACAAGCAGA AGAACGGCATCG	319
G	t	a	G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGTAGCCGTT CTTCTGCTTGTG	126	AAAACACAAGCAGA AGAACGGCAAGG	320
G	t	a	c	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGTACCCGTT CTTCTGCTTGTG	127	AAAACACAAGCAGA AGAACGGGTACG	321
G	t	a	c	g	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGTACGCGTT CTTCTGCTTGTG	128	AAAACACAAGCAGA AGAACGGGTACG	322
G	t	a	c	g	g	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGTACGGGTT CTTCTGCTTGTG	129	AAAACACAAGCAGA AGAACGGGTACG	323
G	t	a	c	g	g	c	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGTACGGCTT CTTCTGCTTGTG	130	AAAACACAAGCAGA AGAACGGGTACG	324
G	t	a	c	g	g	c	a	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGTACGGCAT CTTCTGCTTGTG	131	AAAACACAAGCAGA AGATGCCGTACG	325
G	t	a	c	g	g	c	a	a	C	T	T	C	T	G	C	T	T	G	T	ACACCGTACGGCA ACTTCTGCTTGTG	132	AAAACACAAGCAGA AGTTGCCGTACG	326
G	t	a	c	g	g	c	a	a	g	T	T	C	T	G	C	T	T	G	T	ACACCGTACGGCA AGTTCTGCTTGTG	133	AAAACACAAGCAGA ACTTGCCGTACG	327
G	A	T	G	C	C	G	T	T	C	T	T	C	T	G	C	T	a	G	a	ACACCGATGCCGTT CTTCTGCTAGAG	134	AAAACACTAGCAGA AGAACGGCATCG	328
G	A	T	G	C	C	G	T	T	C	T	T	C	T	G	g	T	a	G	T	ACACCGATGCCGTT CTTCTGGTAGTG	135	AAAACACTAGCAGA AGAACGGCATCG	329
G	A	T	G	C	C	G	T	T	C	T	T	C	a	G	C	T	a	G	T	ACACCGATGCCGTT CTTCAGCTAGTG	136	AAAACACTAGCTGA AGAACGGCATCG	330
G	A	T	G	C	C	G	T	T	C	T	a	C	T	G	C	T	a	G	T	ACACCGATGCCGTT CTACTGCTAGTG	137	AAAACACTAGCAGT AGAACGGCATCG	331
G	A	T	G	C	C	G	T	T	g	T	T	C	T	G	C	T	a	G	T	ACACCGATGCCGTT GTTCTGCTAGTG	138	AAAACACTAGCAGA ACAACGGCATCG	332
G	A	T	G	C	C	G	a	T	C	T	T	C	T	G	C	T	a	G	T	ACACCGATGCCGAT CTTCTGCTAGTG	139	AAAACACTAGCAGA AGATCGGCATCG	333
G	A	T	G	C	g	G	T	T	C	T	T	C	T	G	C	T	a	G	T	ACACCGATGCCGTT CTTCTGCTAGTG	140	AAAACACTAGCAGA AGAACGGCATCG	334
G	A	T	c	C	C	G	T	T	C	T	T	C	T	G	C	T	a	G	T	ACACCGATCCCGTT	141	AAAACACTAGCAGA	335

Tabela A																													
Posição de Sequência Alvo de gRNA																				Oligos para Geração de plasmídeo de expressão de gRNA									
																				CTTCTGCTAGTG		AGAACGGGATCG							
G	t	T	G	C	C	G	T	T	C	T	T	C	T	G	C	T	a	G	T		ACACCGTTGCCGTT CTTCTGCTAGTG	142	AAAACACTAGCAGA AGAACGGCAACG	338					
G	A	T	G	C	C	G	T	T	g	T	T	C	T	G	C	T	T	G	a		ACACCGATGCCGTT GTTCTGCTTGAG	143	AAAACCTAAGCAGA ACAACGGCATCG	339					
G	A	T	G	C	C	G	T	T	g	T	T	C	T	G	g	T	T	G	T		ACACCGATGCCGTT GTTCTGGTTGTG	144	AAAACACAACCAGA ACAACGGCATCG	340					
G	A	T	G	C	C	G	T	T	g	T	T	C	a	G	C	T	T	G	T		ACACCGATGCCGTT GTTCACTGTTGTG	145	AAAACACAAGCTGA ACAACGGCATCG	341					
G	A	T	G	C	C	G	T	T	g	T	a	C	T	G	C	T	T	G	T		ACACCGATGCCGTT GTACTGCTTGTG	146	AAAACACAAGCAGT ACAACGGCATCG	342					
G	A	T	G	C	C	G	a	T	g	T	T	C	T	G	C	T	T	G	T		ACACCGATGCCGAT GTTCTGCTTGTG	147	AAAACACAAGCAGA ACATCGGCATCG	343					
G	A	T	G	C	g	G	T	T	g	T	T	C	T	G	C	T	T	G	T		ACACCGATGCCGTT GTTCTGCTTGTG	148	AAAACACAAGCAGA ACAACCGCATCG	344					
G	A	T	c	C	C	G	T	T	g	T	T	C	T	G	C	T	T	G	T		ACACCGATCCCGTT GTTCTGCTTGTG	149	AAAACACAAGCAGA ACAACGGGATCG	345					
G	t	T	G	C	C	G	T	T	g	T	T	C	T	G	C	T	T	G	T		ACACCGTTGCCGTT GTTCTGCTTGTG	150	AAAACACAAGCAGA ACAACGGCAACG	346					
G	t	T	G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	a		ACACCGTTGCCGTT CTTCTGCTTGTG	151	AAAACCTAAGCAGA AGAACGGCAACG	347					
G	t	T	G	C	C	G	T	T	C	T	T	C	T	G	g	T	T	G	T		ACACCGTTGCCGTT CTTCTGTTGTG	152	AAAACACAACCAGA AGAACGGCAACG	348					
G	t	T	G	C	C	G	T	T	C	T	T	C	a	G	C	T	T	G	T		ACACCGTTGCCGTT CTTCACTGTTGTG	153	AAAACACAAGCTGA AGAACGGCAACG	349					
G	t	T	G	C	C	G	T	T	C	T	a	C	T	G	C	T	T	G	T		ACACCGTTGCCGTT CTACTGCTTGTG	154	AAAACACAAGCAGT AGAACGGCAACG	350					
G	t	T	G	C	C	G	a	T	C	T	T	C	T	G	C	T	T	G	T		ACACCGTTGCCGAT CTTCTGCTTGTG	155	AAAACACAAGCAGA AGATCGGCAACG	351					
G	t	T	G	C	g	G	T	T	C	T	T	C	T	G	C	T	T	G	T		ACACCGTTGCCGTT CTTCTGCTTGTG	156	AAAACACAAGCAGA AGAACCGCAACG	352					
G	t	T	c	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T		ACACCGTTCGCGTT CTTCTGCTTGTG	157	AAAACACAAGCAGA AGAACGGGAACG	353					
Sítio 3 de EGFP alvo																													
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	Oligonucleotídeo 1 (5' a 3')	1	Oligonucleotídeo 2 (5' a 3')							
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1										
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A		ACACCGGTGGTGC AGATGAACCTCAG	158	AAAACCTGAAGTTCA TCTGCACCACCG	354					
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	t		ACACCGGTGGTGC AGATGAACCTCTG	159	AAAACAGAAGTTCA TCTGCACCACCG	355					
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	g	A		ACACCGGTGGTGC AGATGAACCTTGTG	160	AAAACCTGAAGTTCA TCTGCACCACCG	356					
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	a	C	A		ACACCGGTGGTGC AGATGAACCTACAG	161	AAAACCTGAGTTCT TCTGCACCACCG	357					
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	C	a	T	C	A		ACACCGGTGGTGC AGATGAACATCAG	162	AAAACCTGATGTTCA TCTGCACCACCG	358					
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	g	T	T	C	A		ACACCGGTGGTGC AGATGAAGTTCTG	163	AAAACCTGAACTTCA TCTGCACCACCG	359					
G	G	T	G	G	T	G	C	A	G	A	T	G	A	t	C	T	T	C	A		ACACCGGTGGTGC AGATGATCTTCAG	164	AAAACCTGAAGATCA TCTGCACCACCG	360					
G	G	T	G	G	T	G	C	A	G	A	T	G	t	A	C	T	T	C	A		ACACCGGTGGTGC AGATGACTTCTG	165	AAAACCTGAAGTACA TCTGCACCACCG	361					
G	G	T	G	G	T	G	C	A	G	A	T	c	A	A	C	T	T	C	A		ACACCGGTGGTGC AGATCAACTTCAG	166	AAAACCTGAAGTTGA TCTGCACCACCG	362					
G	G	T	G	G	T	G	C	A	G	A	a	G	A	A	C	T	T	C	A		ACACCGGTGGTGC AGAAGAATCTCAG	167	AAAACCTGAAGTTCT TCTGCACCACCG	363					
G	G	T	G	G	T	G	C	A	G	t	T	G	A	A	C	T	T	C	A		ACACCGGTGGTGC AGTTGAATCTCAG	168	AAAACCTGAAGTTCA ACTGCACCACCG	364					
G	G	T	G	G	T	G	C	A	c	A	T	G	A	A	C	T	T	C	A		ACACCGGTGGTGC ACATGAACTTCTG	169	AAAACCTGAAGTTCT TGTGCACCACCG	365					
G	G	T	G	G	T	G	C	t	G	A	T	G	A	A	C	T	T	C	A		ACACCGGTGGTGC TGATGAATCTCAG	170	AAAACCTGAAGTTCA TCAGCACCACCG	366					
G	G	T	G	G	T	G	g	A	G	A	T	G	A	A	C	T	T	C	A		ACACCGGTGGTGG AGATGAACCTCAG	171	AAAACCTGAAGTTCT TCTCCACCACCG	367					
G	G	T	G	G	T	c	C	A	G	A	T	G	A	A	C	T	T	C	A		ACACCGGTGGTCC AGATGAACCTCAG	172	AAAACCTGAAGTTCA TCTGGACCACCG	368					
G	G	T	G	G	a	G	C	A	G	A	T	G	A	A	C	T	T	C	A		ACACCGGTGGAGC AGATGAACCTCAG	173	AAAACCTGAAGTTCT TCTGCTCCACCG	369					
G	G	T	G	c	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A		ACACCGGTGCTGC AGATGAACCTCAG	174	AAAACCTGAAGTTCA TCTGCAGCACCG	370					
G	G	T	c	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A		ACACCGGTCGTGC AGATGAACCTCAG	175	AAAACCTGAAGTTCT TCTGCAGCACCG	371					

Tabela A																							
Posição de Sequência Alvo de gRNA																			Oligos para Geração de plasmídeo de expressão de gRNA				
G	G	a	G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGAGGTGC	176	AAAAGTGAAGTTCA	376
																				AGATGAACCTCAG		TCTGCACCTCCG	
G	c	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGTGGTGC	177	AAAAGTGAAGTTCA	377
																				AGATGAACCTCAG		TCTGCACCAGCG	
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	g	t	ACACCGTGGTGC	178	AAAACACAAGTTCA	378
																				AGATGAACCTGTG		TCTGCACCACCG	
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	C	a	a	C	A	ACACCGTGGTGC	179	AAAAGTGTGTTCA	379
																				AGATGAACACAG		TCTGCACCACCG	
G	G	T	G	G	T	G	C	A	G	A	T	G	A	t	g	T	T	C	A	ACACCGTGGTGC	180	AAAAGTGAACATCA	380
																				AGATGAACCTCAG		TCTGCACCACCG	
G	G	T	G	G	T	G	C	A	G	A	T	c	t	A	C	T	T	C	A	ACACCGTGGTGC	181	AAAAGTGAAGTAGA	381
																				AGATCTACTTCA		TCTGCACCACCG	
G	G	T	G	G	T	G	C	A	G	t	a	G	A	A	C	T	T	C	A	ACACCGTGGTGC	182	AAAAGTGAAGTTCT	382
																				AGTAGAATTCAG		ACTGCACCACCG	
G	G	T	G	G	T	G	C	t	c	A	T	G	A	A	C	T	T	C	A	ACACCGTGGTGC	183	AAAAGTGAAGTTCA	383
																				TCATGAACCTCAG		TGAGCACCACCG	
G	G	T	G	G	T	c	g	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGTGGTGC	184	AAAAGTGAAGTTCA	384
																				AGATGAATTCAG		TCTGCACCACCG	
G	G	T	G	c	a	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGTGCACG	185	AAAAGTGAAGTTCA	385
																				AGATGAACCTCAG		TCTGCTGCACCG	
G	G	a	c	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGACGTGC	186	AAAAGTGAAGTTCA	386
																				AGATGAACCTCAG		TCTGCACGTCAG	
G	c	a	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A		ACACCGCAGGTGC	187	AAAAGTGAAGTTCA	387
																				AGATGAACCTCAG		TCTGCACCAGGG	
G	c	a	c	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCACGTGC	188	AAAAGTGAAGTTCA	388
																				AGATGAACCTCAG		TCTGCACGTGCG	
G	c	a	c	c	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCACCTGC	189	AAAAGTGAAGTTCA	389
																				AGATGAACCTCAG		TCTGCAGGTGCG	
G	c	a	c	c	a	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCACACG	190	AAAAGTGAAGTTCA	390
																				AGATGAACCTCAG		TCTGCTGGTGCG	
G	c	a	c	c	a	c	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCACACC	191	AAAAGTGAAGTTCA	391
																				AGATGAACCTCAG		TCTGCTGGTGCG	
G	c	a	c	c	a	c	g	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCACACG	192	AAAAGTGAAGTTCA	392
																				AGATGAACCTCAG		TCTGCTGGTGCG	
G	c	a	c	c	a	c	g	t	G	A	T	G	A	A	C	T	T	C	A	ACACCGCACACG	193	AAAAGTGAAGTTCA	393
																				TGATGAACCTCAG		TCACGTGGTGCG	
G	c	a	c	c	a	c	g	t	c	A	T	G	A	A	C	T	T	C	A	ACACCGCACACG	194	AAAAGTGAAGTTCA	394
																				TCATGAACCTCAG		TGACGTGGTGCG	
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	a	C	t	ACACCGTGGTGC	195	AAAACAGTAGTTCA	395
																				AGATGAACACTG		TCTGCACCACCG	
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	g	T	a	C	A	ACACCGTGGTGC	196	AAAAGTGTACTTCA	396
																				AGATGAAGTACAG		TCTGCACCACCG	
G	G	T	G	G	T	G	C	A	G	A	T	G	t	A	C	T	a	C	A	ACACCGTGGTGC	197	AAAAGTGTAGTACA	397
																				AGATGTACTACAG		TCTGCACCACCG	
G	G	T	G	G	T	G	C	A	G	A	a	G	A	A	C	T	a	C	A	ACACCGTGGTGC	198	AAAAGTGTAGTTCT	398
																				AGAAGAACTACAG		TCTGCACCACCG	
G	G	T	G	G	T	G	C	A	c	A	T	G	A	A	C	T	a	C	A	ACACCGTGGTGC	199	AAAAGTGTAGTTCA	399
																				ACATGAACACTACAG		TGTGCACCACCG	
G	G	T	G	G	T	G	g	A	G	A	T	G	A	A	C	T	a	C	A	ACACCGTGGTGG	200	AAAAGTGTAGTTCA	399
																				AGATGAACACTACAG		TCTCCACCACCG	
G	G	T	G	G	a	G	C	A	G	A	T	G	A	A	C	T	a	C	A	ACACCGTGGAGC	201	AAAAGTGTAGTTCA	399
																				AGATGAACACTACAG		TCTGCTCCACCG	
G	G	T	c	G	T	G	C	A	G	A	T	G	A	A	C	T	a	C	A	ACACCGTGGTGC	202	AAAAGTGTAGTTCA	399
																				AGATGAACACTACAG		TCTGCACGACCG	
G	c	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	a	C	A	ACACCGTGGTGC	203	AAAAGTGTAGTTCA	399
																				AGATGAACACTACAG		TCTGCACCAGCG	
G	G	T	G	G	T	G	C	A	c	A	T	G	A	A	C	T	T	C	t	ACACCGTGGTGC	204	AAAACAGAAGTTCA	399
																				ACATGAACCTCTG		TGTGCACCACCG	
G	G	T	G	G	T	G	C	A	c	A	T	G	A	A	g	T	T	C	A	ACACCGTGGTGC	205	AAAAGTGAACCTCA	399
																				ACATGAAGTTTCA		TGTGCACCACCG	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGTGGTGC	206	AAAAGTGAAGTACA	400
																				ACATGTACTTCA		TGTGCACCACCG	
G	G	T	G	G	T	G	C	A	c	A	a	G	A	A	C	T	T	C	A	ACACCGTGGTGC	207	AAAAGTGAAGTTCT	401
																				ACAAGAACTTCA		TGTGCACCACCG	
G	G	T	G	G	T	G	g	A	c	A	T	G	A	A	C	T	T	C	A	ACACCGTGGTGG	208	AAAAGTGAAGTTCA	402
																				ACATGAACCTCAG		TGTCCACCACCG	
G	G	T	G	G	a	G	C	A	c	A	T	G	A	A	C	T	T	C	A	ACACCGTGGAGC	209	AAAAGTGAAGTTCA	403
																				ACATGAACCTCAG		TGTGCTCCACCG	
G	G	T	c	G	T	G	C	A	c	A	T	G	A	A	C	T	T	C	A	ACACCGTGGTGC	210	AAAAGTGAAGTTCA	404
																				ACATGAACCTCAG		TGTGCACGACCG	
G	c	T	G	G	T	G	C	A	c	A	T	G	A	A	C	T	T	C	A	ACACCGTGGTGC	211	AAAAGTGAAGTTCA	405
																				ACATGAACCTCAG		TGTGCACCAGCG	
G	c	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	t	ACACCGTGGTGC	212	AAAACAGAAGTTCA	406

Sequências de oligonucleotídeos usadas para gerar plasmídeos de expressão que codificam gRNAs individuais/gRNAs individuais variantes objetivados aos sítios no gene repórter EGFP e gRNAs individuais objetivados a seus genes humanos alvo #, SEQ ID NO:.

[0115] Células U2OS.EGFP que trazem uma cópia única integrada de um gene de fusão *EGFP-PEST* foram cultivadas conforme anteriormente descrito (Reyon *et al.*, Nat Biotech 30, 460-465 (2012)). Para transfecções, 200.000 células foram nucleofectadas com as quantidades indicadas de plasmídeo de expressão de sgRNA e pJDS246 em conjunto com 30 ng de um plasmídeo de codificação de

Td-tomate usando o kit SE Cell Line 4D-Nucleofector™ X (Lonza), de acordo com o protocolo do fabricante. As células foram analisadas 2 dias a pós transfecção usando um citômetro de fluxo BD LSR II. As transfecções para otimização da concentração de plasmídeo de gRNA/Cas9 foram realizadas em triplicata e todas as outras as transfecções foram realizadas em duplicata.

Amplificação por PCR e verificação da sequência de sítios genômicos humanos endógenos

[0116] As reações de PCR foram realizadas usando DNA polimerase Phusion Hot Start II High-Fidelity (NEB) com iniciadores de PCR e as condições listadas na Tabela B. A maioria dos *loci* amplificaram com sucesso usando PCR "touchdown" (98 °C, 10 s; 72-62 °C, -1 °C/ciclo, 15 s; 72 °C, 30 s] 10 ciclos, [98 °C, 10 s; 62 °C, 15 s; 72 °C, 30 s] 25 ciclos). PCR para os alvos restantes foi realizada com 35 ciclos em uma temperatura constante de recozimento de 68 °C ou 72 °C e DMSO a 3% ou betaína a 1M, se necessário. Os produtos de PCR foram analisados em um sistema de eletroforese capilar QIAxCEL para verificar o tamanho e pureza. Os produtos validados foram tratados com ExoSap-TI (Affymetrix) e sequenciados por meio do método de Sanger (MGH DNA Sequencing Core) para verificar cada sítio alvo.

TABELA B								
Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SE Q ID N O	Iniciador de PCR reverso	SE Q ID N O	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	TCCAGATGGCACATTGTCAG	436.	AGGGAGCAGGAAAGTGAGGT	748.	DMSO			
	GGGGCCCACTCTTCTTCAT	437.	ACCCAGACTCCTGGTGTGGC	749.	No DMSO	0	0	1
	GCTAAGCAGAGATGCCTATGCC	438.	ACCACCCCTTCCCCCAGAAA	750.	DMSO	2	0	0
	ACCCACAGCCAGGTTTTCA	439.	GAATCACTGCACCTGGCCATC	751.	DMSO	0	0	2
	TGCGGCAACTTCAGACAAAC	440.	TAAAGGGCGTCTGGGAGAG	752.	DMSO	1	1	0
	GCATGTCAGGATCTGACCCC	441.	TGCAGGGCCATCTTGTGTGT	753.	DMSO	0	2	0
	CCACCACATGTTCTGGGTGC	442.	CTGGGTCTGTTCCCTGTGGG	754.	DMSO	1	1	1
	GGCTCTCCCTGCCCTAGTTT	443.	GCAGGTCAAGTTGGAACCCG	755.	DMSO	0	2	1
	GGGGCTGAGAACACATGAGATGCA	444.	AGATTTGTGCACTGCCTGCCT	756.	DMSO	1	0	2

TABELA B								
Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SEQ ID NO	Iniciador de PCR reverso	SEQ ID NO	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	CCCGACCTCCGCTCCAAAGC	445.	GGACCTCTGCACACCCTGGC	757.	DMSO	2	1	0
	TGCAAGGTGCGATAGTCCCA	446.	CAGGAGGGGGAAGTGTGTCC	758.	DMSO	1	1	1
	GCCCATTCTTTTGCAGTGGGA	447.	GAGAGCAAGTTGTTCGCCAGG	759.	DMSO	0	1	2
	GCCCCAGCCCTCTGTTC	448.	GCTGCTGGTAGGGGAGCTGG	760.	DMSO	1	2	0
	CGGCTGCCTTCCCTGAGTCC	449.	GGGTGACGCTTGCCATGAGC	761.	72C recoze r, DMSO a 3%	1	2	0
	TGACCCTGGAGTACAAA TGTTCCCA	450.	GCTGAGACAACCAGCCCA GCT	762.	72C recoze r, DMSO a 3%	2	1	0
	TGCTCCACCCTTAGCCCT	451.	GCAGCCGATCCACACTGGGG	763.	DMSO	1	0	2
	AACTCAGGACAACACTGCTGT	452.	CCCAGGAGCAGGGTACAA TGC	764.	DMSO	0	1	2
	TCCTCCTTGGAGAGGGCCC	453.	CCTTGGAAGGGGCTTGG TGG	765.	DMSO	0	3	0
	CCGAGGGCATGGGAATCCT	454.	GGCTGCTGCGAGTTGCCAAC	766.	DMSO	0	1	3
	TGCTTTGCATGGGTCTCAGACA	455.	GGGTGCTTGCCCTCTGTGT	767.	DMSO	0	2	2
	AGCTCCTTCTCATTTCTCTCTGCTGT	456.	CACAGAAGGATGTGTGCA GGT	768.	DMSO	0	2	2
	AGCAGACACAGTGAATGCTGCT	457.	GGTCAGGTGTGCTGCTAGGCA	769.	DMSO	1	1	2
	CCTGTGGGCTCTCAGGTGC	458.	ACTGCCTGCCAAAGTGGGTGT	770.	Sem DMSO TD	1	1	2
	AGCTGCACTGGGGAATGAGT	459.	TGCCGGGTAATAGCTGGCTT	771.	DMSO	0	1	3
	CCAGCCTGGGCAACAAAGCG	460.	GGGGGCTTCCAGGTCACAGG	772.	72C recoze r, DMSO a 3%, 6% DMSO	0	3	1
	TACCCCCACTGCCCAATTGC	461.	ACAGGTCCATGCTTAGCAGAGGG	773.	DMSO	0	1	3
GGGTGATTGAAGTTTGCTCCAGG GGGTGATTGAAGTTTGCTGCAGG (SEQ ID NO:424)	ACGGATTACGACGGAGGTGC	462.	CCGAGTCCGTGGCAGAGAGC	774.	DMSO	0 / 1	2	2
	TGTGGTTGAAGTAGGGGACAGGT	463.	TGGCCCAATTGGAAGTGA TTTCGT	775.	DMSO	3	1	0
	TGGGATGGCAGAGTCATCAACGT	464.	GGCCCAATCGGTAGAGGATGCA	776.	DMSO	0	3	1
	ATGGGGCGCTCCAGTCTGTG	465.	TGCACCCACACAGCCAGCAA	777.	DMSO	0	3	1
	GGGGAGGGAGGACCAGGGAA	466.	AATTAGCTGGGCGCGGTGTGT	778.	72C recoze r, DMSO a 3%	0	1	3
	ATCCCGTGCAGGAAGTCGCC	467.	CAGGCGGGCCCTTGAGGAT	779.	DMSO	3	1	0
	CCCCAACCTTTGCTCAGCG	468.	TGAGGAGAACACCACAGGCAGA	780.	DMSO	1	2	1
	ATCGACGAGGAGGGGGCCTT	469.	CCCCTCACTCAAGCAGGCC	781.	DMSO	0	3	1
	TGCTCAAGGGGCTGTTCCA	470.	CAGGGGCAGTGGCAGGAGTC	782.	No DMSO	1	3	0
	TGCTGGCACGCAGTAGGTG	471.	GGGAAGGGGAACAGTGCA	783.	DMSO	0	0	5

TABELA B								
Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SE Q ID N O	Iniciador de PCR reverso	SE Q ID N O	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	Não otimizado					1	1	3
	ACCTGGGCTTGCCACTA GGG	472.	GCTGCTCGCAGTTAAGCA CCA	784.	DMSO	1	3	1
	GTGGCCGGGCTACTGCT ACC	473.	GGTCCACAAGCTGGGGG CA	785.	DMSO	3	2	0
	Não otimizado					1	3	1
	GCAAGAGCGGAGGAGA CCC	474.	AGAGTCATCCATTTCCTGG GGGC	786.	DMSO	2	3	0
	GGGGTCAGTGGTGATAT CCCCCT	475.	AGGAATCCTTTTCCATT GCTTGTTT	787.	Betain a a 1M, TD	1	4	0
	AGAGAGGCCACGTGGAG GGT	476.	GCCTCCCTCCTCCTTCC CA	788.	DMSO	1	3	1
	GACAGTGCCTTGCGATG CAC	477.	TCTGACCGGTATGCCTGA CG	789.	DMSO	3	2	0
	TGTGTGAACGCAGCCTG GCT	478.	TGGTCTAGTACTTCTCCA GCCTT	790.	DMSO	3	1	1
	GGTTCTCCCTTGGCTCCT GTGA	479.	CCCACGTCTCCTAGCCCT GC	791.	DMSO	1	3	1
	TGAAGTCAACAATCTAAG CTTCCACCT	480.	AGCTTTGGTAGTTGGAGT CTTTGAAGG	792.	DMSO	3	1	2
	TGATTGGGCTGCAGTTCA TGTACA	481.	GCACAGCCTGCCCTTGGA AG	793.	DMSO	2	1	3
	TCCATGGGCCCTCTGA AAGA	482.	AGCGGCTTCTGCTTCTGC GA	794.	DMSO	1	0	5
	GCGGTTGGTGGGGTTGA TGC	483.	GAGTTCCTCCTCCGCCA GT	795.	DMSO	2	0	4
	AGGCAAGATTTCCAGTG TGCAAGA	484.	GCTTTGCTGGGACTCC GC	796.	DMSO	2	0	4
	GCTGCTGGTCGGGCTCT CTG	485.	GCTCTGTCCCACCTCCCC TGG	797.	Sem DMSO TD	3	1	2
	GCTGCGAGGCTTCCGTG AGA	486.	CGCCCTAGAGCTAAGGG GGT	798.	DMSO	3	2	1
	CCAGGAGCCTGAGAGCT GCC	487.	AGGGCTAGGACTGCAGTG AGC	799.	DMSO	1	3	2
	CTGTGCTCAGCCTGGGT GCT	488.	GCCTGGGGCTGTGAGTAG TTT	800.	DMSO	2	3	1
	AGCTCGCGCCAGATCTG TGG	489.	ACTTGGCAGGCTGAGGCA GG	801.	72C recoze r, DMSO a 3%	4	2	0
	AGAGAAGTCGAGGAAGA GAGAG	490.	CAGCAGAAAGTTCATGGT TTCG	802.	DMSO			
	TGGACAGCTGCAGTACT CCCTG	491.	ACTGATCGATGATGGCCT ATGGGT	803.	DMSO	0	0	2
	CAAGATGTCACTTGGG CTA	492.	GCAGCCTATTGTCTCCTG GT	804.	DMSO	1	0	1
	GTCCAGTGCCTGACCCT GGC	493.	AGCATCATGCCTCCAGCT TCA	805.	DMSO	1	1	1
	GCTCCCGATCCTCTGCC ACC	494.	GCAGCTCCACCACCCTC AG	806.	DMSO	1	2	0
	GGGGACAGGCAGGCAAG GAG	495.	GTGCGTGTCCGTTACCC CT	807.	DMSO	1	1	1
	AAGGGGCTGCTGGGTAG GAC	496.	CGTGATTCGAGTTCCTGG CA	808.	DMSO	2	1	0
	GACCCTCAGGAAGCTGG GAG	497.	CTGCGAGATGCCCCAAAT CG	809.	betain a a 1M, TD	1	0	2
	CCGCGGCGCTCTGCTAG A	498.	TGCTGGGATTACAGGCGC GA	810.	DMSO	1	1	1
	CCAGGTGGTGTGAGCGG AGG	499.	TGCCTGGCCCTCTGTAG TCT	811.	DMSO	0	2	1
	CGACTCCACGGCGTCTC AGG	500.	CAGCGCAGTCCAGCCGA TG	812.	betain a a 1M, TD	2	1	0
	CTTCCCTCCCCAGCAC CAC	501.	GCTACAGGTTGCACAGTG AGAGGT	813.	DMSO	1	1	1

TABELA B								
Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SEQ ID NO	Iniciador de PCR reverso	SEQ ID NO	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	CCCCGGGGAGTCTGTCC TGA	502.	CCCAGCCGTTCCAGGTCT TCC	814.	72C Recoz er, DMSO a 3%	1	0	2
	GAAGCGCGAAAACCCGG CTC	503.	TCCAGGGTCTTCTCGGCC	815.	DMSO	1	0	2
	AGGGTGGTCAGGGAGGC CTT	504.	CATGGGGCTCGGACCTCG TC	816.	DMSO	2	0	1
	GGGAAGAGGCAGGGCTG TCG	505.	TGCCAGGAAGGAAGCTGG CC	817.	72C Recoz er, DMSO a 3%	0	2	1
	GAGTGACGATGAGCCCC GGG	506.	CCCTTAGCTGCAGTCGCC CC	818.	68C Recoz er, DMSO a 3%	0	1	3
	CCCATGAGGGTTTGAG TGC	507.	TGAAGATGGGCAGTTGG GG	819.	DMSO	0	2	2
	CACCTGGGGCATCTGGG TGG	508.	ACTGGGGTTGGGAGGG GAT	820.	DMSO	2	0	2
	TCATGATCCCCAAAGGG CT	509.	CCATTGTGCTGATCTGTG GGT	821.	DMSO	1	0	3
	TGGTGCCCGAATAGTG GCCA	510.	AGGAAATGTGTTGTGCCA GGGC	822.	DMSO	1	2	1
	GCCTCAGACAACCCTGC CCC	511.	GCCAAGTGTACTCATCAA GAAAGTGG	823.	Sem DMSO TD	2	1	1
	GCCGGGACAAGACTGAG TTGGG	512.	TCCCGAACTCCCGCAAAA CG	824.	DMSO	1	2	1
	TGCTGCAGGTGGTTCCG GAG	513.	CTGGAACCGCATCCTCCG CA	825.	Sem DMSO TD	1	0	3
	ACACTGGTCCAGGTCCC GTCT	514.	GGCTGTGCCTTCCGATGG AA	826.	DMSO	2	1	1
CTCTCCCCCACCCCCCTCTGG (SEQ ID NO:425)	ATCGCGCCCAAGCACAG GT	515.	AGGCTTCTGAAAAGTCC TCAATGCA	827.	DMSO	3	0	2
	Não otimizado					1	1	2
	CCCTCATGGTGGTCTTAC GGCA	516.	AGCCACACATCTTCTGGT AGGG	828.	DMSO	1	1	2
	TGCGTCGCTCATGCTGG GAG	517.	AGGGTGGGGTGTACTGGC TCA	829.	DMSO	0	3	1
	GAGCTGAGACGGCACCA CTG	518.	TGGCCTTGAACCTCTGGG CT	830.	betain a a 1M, TD	0	1	3
	Não otimizado					1	2	1
	AGTGAGAGTGGCACGAA CCA	519.	CAGTAGGTGGTCCCTTCC GC	831.	DMSO	2	1	1
	Não otimizado			832.		1	1	3
	GGGAGAACCCTGTCCAG CCT	520.	AAGCCGAAAAGCTGGGCA AA	833.	DMSO	0	2	3
	CTTCCAGTGTGGCCCG TCC	521.	ACACAGTCAGAGCTCCGC CG	834.	DMSO	1	1	3
	Não otimizado					1	0	4
	CTGAGAGGGGAGGGG GAGG	522.	TGACTGGTCTTGCCTC CCA	835.	68C Recoz er, DMSO a 3%	3	0	2
	CAGCCTGCTGCATCGGA AAA	523.	TGCAGCCAAGAGAAAAAG CCT	836.	betain a a 1M, TD	1	0	4
	TCCCTCTGACCCGGAAC CCA	524.	ACCCGACTTCTCCCAATT GC	837.	DMSO	2	1	2
	TGGGGTTGCGTGCTTG TCA	525.	GCCAGGAGGACACCAGGA CC	838.	DMSO	4	1	0
	ATCAGGTGCCAGGAGGA CAC	526.	GGCCTGAGAGTGGAGAGT GG	839.	DMSO	4	1	0

TABELA B								
Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SE Q ID N O	Iniciador de PCR reverso	SE Q ID N O	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	Não otimizado					1	4	0
	TGAGCCACATGAATCAAG GCCTCC	527.	ACCTCTCCAAGTCTCAGTA ACTCTCT	840.	DMSO	1	3	1
	GGTCCCTCTGTGCAGTG GAA	528.	CTTTGGTGGACCTGCACAGC	841.	DMSO	2	2	2
	GCGAGGCTGCTGACTTC CCT	529.	GCTGGGACTACAGACATG TGCCA	842.	DMSO	2	2	2
ATTTCTCCCCCCC-CCTCAGG (SEQ ID NO:426)	ATTGCAGGCGTGTCAG GCA	530.	AAATCCTGCATGGTGATG GGAGT	843.	DMSO	1	1	5
	TGCTCTGCCATTTATGTC CTATGAAC	531.	ACAGCCTCTTCTCCATGAC TGAGC	844.	DMSO	1	3	2
	TCCGCCCAACAGGAGG CAG	532.	GCGGTGGGAAGCCATTG AG	845.	DMSO	2	3	1
	GGGGGTCTGGCTCACCT GGA	533.	CCTGTGCGGAGAGTGCCT GC	846.	DMSO	3	1	2
	TCCTGGTTCAATTTGCTAG AACTCTGGA	534.	ACTCCAGATGCAACCAGG GCT	847.	DMSO	3	2	1
	CGTGTGGTGAGCCTGAG TCT	535.	GCTTACCGTAGAGGCTG CT	848.	DMSO	3	0	3
	AGGCCCTGATAATTCATG CTACCAA	536.	TCAGTGACAACCTTTTGTA TTCGGCA	849.	DMSO	0	2	4
	Não otimizado					2	2	2
		537.						
	TCCAGATGGCACATTGTC AG	538.	AGGGAGCAGGAAAGTGAG GT	850.	DMSO			
	GCAGGCAAGCTGTCAAG GGT	539.	CACCGACACACCCACTCA CC	851.	DMSO	0	0	1
	GAGGGGGAAGTCACCGA CAA	540.	TACCCGGGCCGTCTGTTA GA	852.	DMSO	0	0	2
	GACACCCACACACTCTC ATGC	541.	TGAATCCCTTACCCCCAA G	853.	DMSO	1	0	1
	TCCTTTGAGGTTATCCC CC	542.	CCAATCCAGGATGATTCC GC	854.	DMSO	1	0	1
	CAGGGCCAGGAACACAG GAA	543.	GGGAGGTATGTGCGGGA GTG	855.	DMSO	1	1	0
	TGCAGCCTGAGTGAGCA AGTGT	544.	GCCCAGGTGCTAAGCCCC TC	856.	DMSO	1	0	1
	TACAGCCTGGGTGATGG AGC	545.	TGTGTCATGGACTTTCCCA TTGT	857.	betain a a 1M, TD	1	1	0
	GGCAGGCATTAACATCAT CAGGTCC	546.	TCTCCCCAAGGTATCAG AGAGCT	858.	DMSO	1	1	0
	GGGCCTCCCTGCTGGTT CTC	547.	GCTGCCGTCCGAACCCAA GA	859.	DMSO	0	1	1
	ACAAACGCAGGTGGACC GAA	548.	ACTCCGAAAATGCCCGC AGT	860.	DMSO	1	1	0
	AGGGGAGGGACATTGCT CT	549.	TTGAGAGGGTTCAAGTGGT TGC	861.	DMSO	1	0	1
	CTAATGCTTACGGCTGCG GG	550.	AGCCAACGGCAGATGCAA AT	862.	DMSO	1	0	1
	GAGCGAAGTTAACCCAC CGC	551.	CACACATGCACATGCCCC TG	863.	68C, DMSO a 3%	2	0	0
	GCATGTGTCTAACTGGAG ACAATAGCA	552.	TCCCCATATCAACACACA CA	864.	DMSO	2	0	0
	GCCCCTCCCGCCTTTTGT GT	553.	TGGGCAAGGACATGAAA CAGACA	865.	DMSO	2	0	0
	GCCTCAGCTCTGCTCTTA AGCCC	554.	ACGAACAGATCATTTTCA TGGCTTCC	866.	DMSO	2	0	0
	CTCCAGAGCCTGGCCTA CCA	555.	CCCTCCTCGGAAGTGCCT TG	867.	DMSO	0	1	1
	TCTGTCAACACACAGTTA CCACC	556.	GTTGCCTGGGGATGGGGT AT	868.	DMSO	0	1	1
	GGGGACCCTCAAGAGGC ACT	557.	GGGCATCAAAGGATGGGG AT	869.	DMSO	2	0	1
	TGTGGAGGGTGGGACCT GGT	558.	ACAGTGAGGTGCGGTCTT TGGG	870.	DMSO	1	0	2
	CGGGGTGGCAGTGACGT CAA	559.	GGTGCAGTCCAAGAGCCC CC	871.	DMSO	0	0	3

TABELA B								
Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SEQ ID NO	Iniciador de PCR reverso	SEQ ID NO	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	AGCTGAGGCAGAGTCCC CGA	560.	GGGAGACAGAGCAGCGC CTC	872.	DMSO	1	1	1
	ACCACCAGACCCCACT CCA	561.	AGGACGACTTGTGCCCCA TTCA	873.	72C Recoz er, DMSO a 3%	1	1	1
	GGGTCAGGACGCAGGTC AGA	562.	TCCACCCACCCACCCATC CT	874.	72C Recoz er, DMSO a 3%	2	0	1
	ACACTCTGGGCTAGGTG CTGGA	563.	GCCCCCTCACCACATGAT GCT	875.	DMSO	2	0	1
	GGGGCCATTCTCTGCT GCA	564.	TGGGGATCCTTGCTCATG GC	876.	DMSO	3	0	0
	ACACACTGGCTCGCATT ACCA	565.	CCTGCACGAGGCCAGGTG TT	877.	DMSO	2	1	0
	TGGGCACGTAGTAACT GCACCA	566.	CTCGCCGCCGTGACTGTA GG	878.	DMSO	0	3	1
	TCAGCTGGTCTGGGCT TGG	567.	AGAGCACTGGGTAGCAGT CAGT	879.	DMSO	2	1	0
	AGACACAGCCAGGGCCT CAG	568.	GGTGGGCGTGTGTGTGTA CC	880.	68C, DMSO a 3%	1	1	1
	ACACTCTCACACGCGAC CAA	569.	GAGAAGTCAGGGCTGGCG GG	881.	72C Recoz er, DMSO a 3%	1	2	0
	ACTGCCTGCATTCCCGG GT	570.	TGGTGAGGGCTTCAGGGA GC	882.	DMSO	1	1	1
	GCCAGGTTCACTGACTGC CC	571.	TCCTTCTACACATCGGCG GC	883.	DMSO	2	1	0
	CGAGGGAGCCGAGTTCC TAA	572.	CTGACCTGGGCTCTGGT AC	884.	DMSO	1	2	0
	TCCTCGGGAAGTCATGG CTTCA	573.	GCACTGAGCAACCAGGAG CAC	885.	DMSO	2	1	0
	Não otimizado					1	0	3
	TAAACCGTTGCCCGCG CTC	574.	GCTCCCTGCCAGGTGAA CC	886.	DMSO	2	1	1
	CCTGCTGAGACTCCAGG TCC	575.	CTGCGGAGTGGCTGGCTA TA	887.	DMSO	2	0	2
	CTCGGGGACTGACAAGC CGG	576.	GGAGCAGCTCTCCAGGG CC	888.	DMSO	3	0	1
	CCCCGACCAAGCAGGA GCA	577.	CTGGCAGCCTCTGGATGG GG	889.	DMSO	1	2	1
	Não otimizado					0	3	1
	ATTTAGAGCCCCGGGG AAA	578.	AGGCCGCGGTGTTATGGT TA	890.	DMSO	1	2	1
	GCCAGTGGCTTAGTGTCT TTGTGT	579.	TGACATATTTCTGGGCC ATGGGT	891.	DMSO	2	1	1
	TGCCAGAAAGACATGGG CCAGA	580.	CCATGCTGACATCATATAC TGGGAAGC	892.	DMSO	3	1	0
	GCGTGTCTCTGTGTGCG TGC	581.	CCAGGCTGGGCACACAGG TT	893.	DMSO	3	1	0
	Não otimizado					2	2	0
	TGCCAGTCCAATATTT AGCAGCT	582.	AGGATGAGTTCATGTCCTT TGTGGGG	894.	DMSO	2	2	0
	GGGTGAAAATTTGGTACT GTTAGCTGT	583.	AATGACTCATTCCCTGGGT ATCTCCCA	895.	DMSO	2	2	0
	TGCCCATCAATCACCTC GGC	584.	CAAGTCGGCAGGGCAGT GA	896.	DMSO	1	2	2
	GCCTCCTCTGCCGCTGG TAA	585.	TGAGAGTTCCTGTTGCTC CACACT	897.	DMSO	1	2	2
	Não otimizado					2	2	1
	GCCACCAAAATAGCCAG CGT	586.	ACATGCATCTGTGTGTC GT	898.	DMSO	3	0	2
	ACAGACTGACCCCTTGAAA AATACCAGT	587.	TGTATCTTTCTTGCCAATG GTTTTCCC	899.	DMSO	2	1	2

TABELA B								
Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SE Q ID N O	Iniciador de PCR reverso	SE Q ID N O	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	AGCCAAATTTCTCAACAG CAGCACT	588.	TCCTGGAGAGCAGGCATT TTTGT	900.	DMSO	3	1	1
	ACCTCCTTGTGCTGCCTG GC	589.	GGCGGGAAGGTAACCCCTG GG	901.	DMSO	2	1	2
	CACAAAGCTCTACCTTTC CAGTAGTGT	590.	TGATCCGATGGTTGTTAC AGCT	902.	DMSO	3	1	1
	TGTGGGGATTACCTGCCT GGC	591.	ACGCACAAAATGCCCTT GTCA	903.	DMSO	2	2	1
	TGAGGCAGACCAGTCAT CCAGC	592.	GCCCAGCAGCAGTGTAGG GC	904.	DMSO	2	3	0
	ATTAGCTGGGCGTGGCG GAG	593.	ACTGCATCTCATCTCAGG CAGCT	905.	DMSO	2	1	3
	TGAAGCAGAAGGAGTGG AGAAGGA	594.	TCAGCTTCACATCTGTTTC AGTTCAGT	906.	DMSO	4	0	2
	TGGTGGAGTGTGTGTGT GGT	595.	AGAGCAGAAAGAGAGTGC CCA	907.	DMSO	1	3	2
	GCCCCTGTACGTCCTGA CAGC	596.	TGCACAAGCCACTTAGCC TCTCT	908.	DMSO	3	1	2
	AGCGCAGGTAAACAGGC CCA	597.	TCTCTCGCCCCGTTTCCTT GT	909.	DMSO	3	1	2
	ATGGGTGCCAGGTACCA CGC	598.	ACAGCAGGAAGGAGCCGC AG	910.	DMSO	2	3	1
	CGGGCGGGTGGACAGAT GAG	599.	AGGAGGTCTCGAGCCAGG GG	911.	DMSO	2	3	1
	TCAACCTAGTGAACACAG ACCACTGA	600.	GTCTATATACAGCCACAA CCTCATGT	912.	DMSO	1	2	3
	GCCAGGGCCAGTGGATT GCT	601.	TGTCATTCTTAGTATGTC AGCCGGA	913.	DMSO	2	4	0
	GAGCCCCACCGTTTCA G TCC	602.	GCCAGAGCTACCCACTCG CC	914.	DMSO	1	3	2
		603.						
	GGAGCAGCTGGTCAGAG GGG	604.	GGAAGGGGGGACACTGG GGA	915.	DMSO			
	TCTCTCCTTCAACTCATG ACCAGCT	605.	ATCTGCACATGTATGTACA GGAGTCAT	916.	DMSO	0	1	1
AAGACAGAGGAGAAGAAGGG (SEQ ID NO:427)	TGGGGAATCTCCAAGAA CCCCC	606.	AGGGTGTACTGTGGGAAC TTTGCA	917.	DMSO	2	1	1
	GATGGCCCCACTGAGCA CGT	607.	ACTTCGTAGAGCCTTAAAC ATGTGGC	918.	DMSO	1	0	2
	AGGATTAATGTTTAAAGT CACTGGTGG	608.	TCAAACAAGGTGCAGATA CAGCA	919.	betain a a 1M, TD	1	0	2
	TCCAAGCCACTGTTTTCT CAGTCA	609.	TGCTCTGTGGATCATATTT TGGGGGA	920.	DMSO	0	1	2
	ACTTTCAGAGCTTGGGG CAGGT	610.	CCCACGCTGAAGTGCAAT GGC	921.	DMSO	1	1	1
	CAAAGCATGCCTTTCAGC CG	611.	GGCTCTCGATTGGCAC CT	922.	betain a a 1M, TD	1	1	1
	Não otimizado					1	0	2
	GGACTCCCTGCAGCTCC AGC	612.	AGGAACACAGGCCAGGCT GG	923.	72C Recoz er, 6% DMSO	0	0	3
	CCCTTTAGGCACCTTCCC CA	613.	CCGACCTTCATCCCTCCT GG	924.	DMSO	0	1	2
	TGATTCTGCCTTAGAGTC CCAGGT	614.	TGGGCTCTGTGCCCTAC CCA	925.	DMSO	0	3	0
	Não otimizado					2	1	0
	AGGCAGGAGAGCAAGCA GGT	615.	ACCCTGACTACTGACTGA CCGCT	926.	DMSO	0	1	2
	CTCCCCATTGCGACCCG AGG	616.	AGAGGCATTGACTTGGAG CACCT	927.	DMSO	1	2	0
	CTGGAGCCCAGCAGGAA GGC	617.	CCTCAGGGAGGGGGCCT GAT	928.	DMSO	1	2	0
	ACTGTGGGCGTTGTCCC CAC	618.	AGGTGCGTGCAGGGTTTA AGGA	929.	DMSO	1	0	3
	GGCGCTCCCTTTTCCCT TTGT	619.	CGTCACCCATCGTCTCGT GGA	930.	DMSO	2	0	2

TABELA B								
Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SEQ ID NO	Iniciador de PCR reverso	SEQ ID NO	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	TGCCATCTATAGCAGCCCCT	620.	GCATCTTGCTAACCGTACTCTTCTGA	931.	DMSO	1	0	3
	GTGGAGACGCTAAACCTGTGAGGT	621.	GCTCCTGGCCTCTTCCTACAGC	932.	DMSO	1	2	1
	CCGAACCTCTGCTGAGCTTGATGC	622.	CCAAGTCAATGGGCAACAAGGGA	933.	DMSO	0	2	2
	Não otimizado					1	1	2
	TGCCCCCAAGACCTTTCTCC	623.	ATGGCAGGCAGAGGAGGAGA	934.	DMSO	2	0	2
	GGGTGGGGCCATTGTGGTT	624.	CTGGGGCCAGGGTTTCTGCC	935.	DMSO	3	0	1
	TGGAGAACATGAGAGGCTTGCAA	625.	TCCTTCTGTAGGCAATGGGAACAA	936.	DMSO	3	0	1
	GCCACATGGTAGAAGTCGGC	626.	GGCAGATTTCCCCCATGCTG	937.	betain a a 1M, TD	1	2	1
	TGTACACCCCAAGTCCTCC	627.	AAGGGGAGTGTGCAAGCCTC	938.	DMSO	3	1	0
	AGGTCTGGCTAGAGATGCAGCA	628.	AGTCCAACTCAGGTGAGACCT	939.	DMSO	3	1	0
	CCAAGAGGACCCAGCTGTTGGA	629.	GGGTATGGAATCTGGATTAGCAGAGC	940.	DMSO	0	2	2
	ACCATCTCTTCATTGATGAGTCCCAA	630.	ACACTGTGAGTATGCTTGGCGT	941.	DMSO	2	2	0
	GGCTGCGGGGAGATGAGCTC	631.	TCGGATGCTTTTCCACAGGGCT	942.	DMSO	2	2	1
	TCTTCCAGGAGGGCAGCTCC	632.	CCAATCTGAGCTCCTACAAGGCT	943.	DMSO	1	0	4
	GAGCTGCACTGGATGGCACT	633.	TGCTGGTTAAGGGGTGTTTGGGA	944.	DMSO	1	1	3
	TCTGGGAAGGTGAGGAGGCCA	634.	TGGGGGACAATGGAAAAGCAATGA	945.	DMSO	0	2	3
	CTTGCTCCCAGCCTGACCCC	635.	AGCCCTTGCCATGCAGGACC	946.	DMSO	3	1	1
	GGGATTTTTATCTGTTGGTGCGAA	636.	AACCACAGATGTACCCTCAAAGCT	947.	DMSO	2	2	1
	ACCCATCAGGACCGCAGCAC	637.	TCTGGAACCTGGGAGGCGGA	948.	72C Recoz er, DMSO a 3%	3	1	1
	CGTCCCTCACAGCCAGCTC	638.	CCTCCTTGGGCCTGGGGTTC	949.	DMSO	1	3	1
	CCCTCTGCAAGGTGGAGTCTCC	639.	AGATGTTCTGTCCCAGGCCT	950.	DMSO	1	3	1
	GGCTTCCACTGCTGAAGGCCT	640.	TGCCCTCCACATACCCTCC	951.	DMSO	2	1	2
	AGCATTCGCTGTCGGTGATGT	641.	AGCACCTATTGGACACTGTCTCTCT	952.	DMSO	1	3	1
	TCTAGAGCAGGGGCACATGCA	642.	TGGAGATGGAGCCTGGTGGA	953.	DMSO	2	2	1
	GGTCTCAGAAAATGGAGAGAAAGCAGC	643.	CCCACAGAAACCTGGGCCCT	954.	DMSO	1	2	3
	GGTTGCTGATACCAAAACGTTTGCT	644.	TGGGTCCTCTCCACCTCTGCA	955.	DMSO	0	3	3
	ACTCTCCTTAAGTACTGATATGGCTGT	645.	CAGAACTTGTCTGTGTTGCCA	956.	DMSO	0	4	2
	Não otimizado					2	2	2
	Não otimizado					2	2	2
	CAATGCCTGCAGTCCTCAGGA	646.	TCCCAAGAGAAAACCTGTCTGTGACA	957.	DMSO	4	1	1
	GCATTGGCTGCCAGGGAAA	647.	TGGCTGTGCTGGGCTGTGTT	958.	DMSO	2	2	2
	CCACAAGCCTCAGCCTACCCG	648.	ACAGGTGCCAAAACACTGCCT	959.	DMSO	2	1	3
TCATTGCAGCAGAAGAAGAAAGGTCATTGTAGCAGAAGAAGAAAGG (SEQ ID NO:428)	GCCTCTTGCAAATGAGACTCCTTTT	649.	CGATCAGTCCCCTGGCGTCC	960.	DMSO	2 / 1	2 / 3	2
	TCCAGAAATCTGCCTCCGCA	650.	AGGGGTTTCCAGGCACATGGG	961.	DMSO	0	4	2

TABELA B								
Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SEQ ID NO	Iniciador de PCR reverso	SEQ ID NO	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
		651.		962.				
	TCCTAAAAATCAGTTTG AGATTACTTCC	652.	AAAGTGTTAGCCAACATAC AGAAGTCAGGA	963.	DMSO			
GGTATCTAAGTCATTACCTGTGG GGTATCTAAGTCAATACCTGTGG (SEQ ID NO:429)	ACATCTGGGGAAAGCAA AAGTCAACA	653.	TGTCTGAGTATCTAGGCTA AAAGTGGT	964.	DMSO	1 / 2	1	1
	ACGATCTTGCTTCATTTCC CCTGTACA	654.	AGTGCTTTGTGAAGTGA AGCAAAACA	965.	DMSO	0	3	0
	GCACCTTGGTGCTGCTAA ATGCC	655.	GGGCAACTGAACAGGCAT GAATGG	966.	DMSO	1	2	0
	AACTGTCTGCATCCCCG CC	656.	GGTGCACCTGGATCCACC CA	967.	DMSO	1	1	1
	Não otimizado					1	1	1
	CATCACCTCCACCAGG CCC	657.	ACCACTGCTGCAGGCTCC AG	968.	72C Recoz er, DMSO a 3%	0	3	0
	Não otimizado					2	0	2
	CCTGACCCGTGGTTCCC GAC	658.	TGGTGCCTGGTGTGTGTG GT	969.	72C Recoz er, DMSO a 3%	1	2	1
	TGGGAACATTGGAGAAG TTTCCTGA	659.	CCATGTGACTACTGGGCT GCCC	970.	DMSO	1	1	2
	AGCCTTGGAAGCAACT CCCT	660.	GTTTCTCTCTCTCAGAAAA GAAAGAGG	971.	DMSO	1	0	3
	GGCAGCGGACTTCAGAG CCA	661.	GCCAGAGGCTCTCAGCAG TGC	972.	DMSO	1	0	3
	CCAGCCTGGTCAATATG GCA	662.	ACTGTGCCCCAGCCCCATA TT	973.	DMSO	2	1	1
	ATGCCAACACTCGAGGG GCC	663.	CGGGTTGTGGCACC GG TA	974.	DMSO	2	1	1
	TTGCTCTAGTGGGGAGG GGG	664.	AGAGTTCAGGCATGAAAA GAAGCAACA	975.	DMSO	3	0	1
	AGCTGAAGATAGCAGTGT TTAAGCCT	665.	TGCAATTTGAGGGGCTCT CTTCA	976.	DMSO	1	1	2
	AGTCACTGGAGTAAGCCT GCCT	666.	TGCCAGCCAAAAGTTGTTA GTGTGT	977.	DMSO	2	0	2
	GGGTCTCCCTCAGTGCC CTG	667.	TGTGTGGTAGGGAGCAAA ACGACA	978.	DMSO	2	0	2
	TGGGGGCTGTTAAGAGG CACA	668.	TGACCACACACCCCCCA CG	979.	DMSO	1	2	1
	TCAAAACAGATTGACCAA GGCCAAAT	669.	TGTGTTTTTAAGCTGCACC CCAGG	980.	DMSO	1	0	3
	TCTGGCACCAGGACTGA TTGTACA	670.	GCACGCAGCTGACTCCCA GA	981.	DMSO	1	2	1
	Não otimizado					1	0	3
	AGCATCTGTGATACCCTA CCTGTCT	671.	ACCAGGGCTGCCACAGAG TC	982.	DMSO	1	0	3
	TAGTCTTTGCCCCAGGC TG	672.	CTCGGCCCTGAGAGTTC AT	983.	DMSO	1	2	1
TCCATCTCACTATTACCTGAGGTCCAT CTCACTCATTACCTGATG (SEQ ID NO:430)	CTGCAACCAGGGCCCTT ACC	673.	GAGCAGCAGCAAAGCCAC CG	984.	DMSO	1	1	2
	GCCTGGAGAGCAAGCCT GGG	674.	AGCCGAGACAATCTGCC CG	985.	DMSO	1	1	2
TTTATATTAGTGATTACCTGCGG (SEQ ID NO:431)	AGTGAACAAACAAGCAG CAGTCTGA	675.	GGCAGGTCTGACCAGTGG GG	986.	Sem DMSO TD	1	2	1
	AGGCTCAGAGAGGTAAG CAATGGA	676.	TGAGTAGACAGAAATGTTA CCG GTGT	987.	DMSO	3	0	2
	TCAGAGATGTTAAAGCCT TGGTGGG	677.	AGTGAACCAAGGGAATGG GGGA	988.	DMSO	3	0	2
	TGTGCTTTCTGGGGTAGT GGCA	678.	CACCTCAGCCCTGTAGTC CTGG	989.	DMSO	0	4	1
	CCATTGGGTGACTGAATG CACA	679.	GCCACTGTCCCCAGCCTA TT	990.	betain a a 1M, TD	1	3	1

TABELA B								
Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SEQ ID NO	Iniciador de PCR reverso	SEQ ID NO	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	ACCAAGAAAGTAAAAGGAAACCC	680.	TGAGATGGCATACGATTACCCA	991.	DMSO	1	2	2
	AGGGTGGGACTGAAAGGAGCT	681.	TGGCATCACTCAGAGATTGGAACACA	992.	DMSO	3	1	1
	ACCAGTGTGTGTGACCTTGGA	682.	TCCTATGGGAGGGGAGGCTTCT	993.	DMSO	3	1	1
	CCAGGTGTGGTGGTTCA TGAC	683.	GCATACGGCAGTAGAATGAGCC	994.	68C, DMSO a 3%	4	0	1
	CAGGCGCTGGGTTCTTAGCCT	684.	CCTTCCTGGGCCCATGGTG	995.	DMSO	2	3	0
	TGGGGTCCAAGATGTCCCT	685.	TGAACTGCTTGATGAGGTGTGGA	996.	DMSO	1	2	2
	GCTGGGCTTGGTGGTATATGC	686.	ACTTGCAAAGCTGATAACTGACTGA	997.	DMSO	5	0	1
	AGTTGGTGTCACTGACAA TGGGA	687.	CGCAGCGCACGAGTTCATCA	998.	DMSO	3	0	3
	AGAGGAGGCACAATTCAACCCCT	688.	GGCTGGGAGGCCCTCACAT	999.	DMSO	1	1	4
	GGGAAAGTTTGGGAAAGTCAGCA	689.	AGGACAAGCTACCCACACC	1000.	DMSO	1	3	2
	TGGTGATCAAGGGTTGCTTCT	690.	TCATTCCAGCACGCCGGGAG	1001.	DMSO	0	3	3
	CCCAGGCTGCCATCACACT	691.	TGGAGTAAGTATACCTTGGGACCT	1002.	DMSO	1	3	2
	TCAGTGCCCTGGGTCC TCA	692.	TGTGCAATACCTAGCACGGTGC	1003.	DMSO	4	2	0
	AGCACTCCCTTTTGAATT TGGTGCT	693.	ACTGAAGTCCAGCCTCTTCCATTCA	1004.	DMSO	2	1	3
	GAAACCGTCCCTGGTGCCA	694.	GGGGAGTAGGGGTAGTTTGCC	1005.	DMSO	2	0	4
	TTGCGGGTCCCTGTGGA GTC	695.	AGGTGCCGTGTTGTGCCCCAA	1006.	DMSO	1	2	3
		696.		1007.				
	GCCCTACATCTGCTCTCCCTCCA	697.	GGGCCGGGAAAGAGTTGTG	1008.	DMSO			
	TTGGAGTGTGGCCGGGTTG	698.	ACCTCTCTTCTCTGCCTC ACTGT	1009.	DMSO	0	1	1
	CACACCATGCTGATCCAGGC	699.	GCAGTACGGAAGCACGAAGC	1010.	DMSO	1	1	1
	CTCCAGGGCTCGCTGTC CAC	700.	CTGGGCTCTGCTGTTCCCC	1011.	DMSO	0	2	1
	CTGTGGTAGCCGTGGCCAGG	701.	CCCCATACCCTCTCCGGGA	1012.	DMSO	0	2	1
	GGTGCGGGACTTGAATGAG	702.	CCAGCGTGTTCGAAGGAT	1013.	betain a a 1M, TD	0	1	2
GGAATCCCTCTCCAGCCCTGGGGAATCCCTCTCCAGCCTCTGG (SEQ ID NO:432)	CCAGAGTGGGGCCCTGTGA	703.	TTTCCACACTCAGTTCTGCAGGA	1014.	DMSO	1	1	1 / 2
GGAATCTCTCTTGGCATCTGG (SEQ ID NO:433)	TGTGACTGGTTGTCTGCTTTTCT	704.	GCAGTGTTTTGTGGTGATGGGCA	1015.	betain a a 1M, TD	0	1	5
	CTGGCCAAGGGTGAGTGGG	705.	TGGGACCCAGCAGCCAA TG	1016.	DMSO	1	0	2
	ACGGTGTGCTGGCTGCTCTT	706.	ACAGTGTGACCGTGCTGG	1017.	DMSO	1	1	1
	TGGTTTGGGCTCAGGGATGG	707.	TGCCTCCACAAAAATGCTACCT	1018.	DMSO	0	0	3
	TGGTTTGGGCTCAGGGATGG	708.	ACCCCTTATCCAGAACCATGA	1019.	DMSO	0	0	3
	TCCAAGTCAGCATGAGGGCT	709.	TGGGAGCTGTCTTTTTTGCCA	1020.	DMSO	0	3	0
	CACCCCTCTCAGCTTCCCAA	710.	GCTAGAGGGTCTGCTGCC TT	1021.	DMSO	1	2	0
	AGACCCCTTGGCCAAGCACA	711.	CTTGCTCTCACCCCGCCTCC	1022.	DMSO	2	1	0
	ACATGTGGGAGCGGACAGA	712.	TCTCACTTGTGTTACCGATGTCG	1023.	DMSO	0	1	3

TABELA B								
Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SEQ ID NO	Iniciador de PCR reverso	SEQ ID NO	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	GGACGACTGTGCCTGGGACA	713.	AGTGCCAGAGTGTGTGTA ACTGCT	1024.	72C Recocer, DMSO a 3%	0	1	3
	GGAGAGCTCAGCGCCAGGTC	714.	CAGCGTGGCCCGTGGGAATA	1025.	DMSO	1	1	2
	GCTGAAGTGTCTGGGGTGCT	715.	ACCCCACTGTGGATGAATTGGTACC	1026.	DMSO	1	1	2
	TCGGGGTGACATGGCCATC	716.	TTGCCTCGCAGGGGAAGCAG	1027.	DMSO	0	1	3
	CTCGTGGGAGGCCAACACCT	717.	AGCCACCAACACATACCAAGGCT	1028.	DMSO	2	0	2
	GCATGCCCTTAATCCCGGCT	718.	AGGATTTGAGAGTGATGGGCT	1029.	DMSO	2	1	1
	CGCCAGCCACAAAGTCAT	719.	GCAAAATTTGACCTACTCTAGGCT	1030.	DMSO	1	1	2
	AGCTCACAAGAATTGGAGGTAACAGT	720.	GCAGTCACCTTCACTGCTGT	1031.	DMSO	1	1	2
	AAACTGGGCTGGGCTTCGCG	721.	GGGGTAAGGCATTGTCAAGCC	1032.	DMSO	2	0	2
	GCAGGTAGGCAGTCTGGGCG	722.	TCTCTGCCTCAGCCTCCCA	1033.	betaina a 1M, TD	1	2	1
	GCAGGTAGGCAGTCTGGGCG	723.	TCTCTGCCTCAGCCTCCCA	1034.	betaina a 1M, TD	1	2	1
	GCAGGTAGGCAGTCTGGGCG	724.	TCTCTGCCTCAGCCTCCCA	1035.	betaina a 1M, TD	1	2	1
	GCTCTGGGGTAGAAGGAGGCG	725.	GGCCTGTCAACCAACCAACC	1036.	DMSO	2	2	0
	TGACATGTTGTGTGCTGGGC	726.	AAATCCTGCAGCCTCCCCCT	1037.	DMSO	0	2	2
	TCCTGGTGAGATCGTCCACAGGA	727.	TCCTCCCACTCAGCCTCC	1038.	DMSO	0	3	1
	TCCTAATCCAAGTCCTTTGTTGAGACA	728.	AGGGACCACTCAGCCTCCCTCA	1039.	DMSO	2	2	0
	GGGACACAGTTCCTTCAT	729.	GGGGGAGATTGGAGTTCC	1040.	DMSO	1	0	4
	ACACCACTATCAAGGCAGAGTAGGT	730.	TCTGCCTGGGGTGCTTTCC	1041.	DMSO	1	1	3
	CTGGGAGCGGAGGGAAGTGC	731.	GCCCCGACAGATGAGGCCCT	1042.	DMSO	1	2	2
CAGATTACTGCTGCAGCACCGGG (SEQ ID NO:434)	CGGGTCTCGGAATGCCTCCA	732.	ACCCAGGAATTGCCACCCCT	1043.	DMSO	1	2	3
	TTGCTGTGGTCCCGGTGTG	733.	GCAGACACTAGAGCCCGCC	1044.	DMSO	3	2	0
	GGTGTGGTGACAGGTGGGT	734.	ACCTGCGTCTCTGTGCTGCA	1045.	DMSO	2	3	0
	CTCCAGGACAGTGCTCGGC	735.	CCTGGCCCCATGCTGCCTG	1046.	DMSO	2	2	1
	TGCGTAGGTTTTGCCTCTGTGA	736.	AGGGAATGATTTTTCCACCCCT	1047.	DMSO	2	3	0
	CTCCGACGACACCGTTGGTA	737.	TGCATTGACGTACGATGGCTCA	1048.	DMSO	1	3	1
	ACCTGCAGCATGAACTCTCGCA	738.	ACCTGAGCAACATGACTCACCTGG	1049.	DMSO	2	1	2
ACACAACTTCTGCAGCACCTGG ACACAACTTCTGCAGCACGTGG (SEQ ID NO:435)	TCTCCAGTTTCTTGCTCTCATGG	739.	ACCATTGGTGAACCCAGTCA	1050.	betaina a 1M, TD	3 / 2	3	1
	TGGGGTGGTGGTCTTGAATCCA	740.	TCAGCTATAACCTGGGACTTGTGCT	1051.	DMSO	2	1	3
	AGCAGCCAGTCCAGTGTCTG	741.	CCCTTTTCATCGAGAACCCAGGG	1052.	DMSO	3	1	2
	TGGACGCTGCTGGGAGGAGA	742.	GAGGTCTCGGGCTGCTCGTG	1053.	DMSO	0	3	3
	AGGTTTGCACTCTGTTGCTGG	743.	TGGGGTGATTGGTTGCCAGGT	1054.	DMSO	3	2	1

TABELA B								
Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SE Q ID N O	Iniciador de PCR reverso	SE Q ID N O	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	TCTTCCTTTGCCAGGCAG CACA	744.	TGCAGGAATAGCAGGTAT GAGGAGT	1055	DMSO	4	0	2
	GGACGCCTACTGCCTGG ACC	745.	GCCCTGGCAGCCCATGGT AC	1056	DMSO	3	0	3
	AGGCAGTCATCGCCTTG CTA	746.	GGTCCCACCTTCCCCTAC AA	1057	DMSO	2	3	1
	Não otimizado					3	1	2
	CCCCAGCCCCACCAGT TTC	747.	CAGCCCAGGCCACAGCTT CA	1058	DMSO	1	4	1

Sequências e características de sítios genômicos dentro e fora do alvo para seis RGNs objetivadas a genes endógenos humanos e iniciadores e condições de PCR usados para amplificar estes sítios.

Determinação de frequências de mutação sobre e fora do alvo induzidas por RGN em células humanas

[0117] Para células K562 e U2OS.EGFP, 2×10^5 células foram transfectadas com 250 ng de plasmídeo de expressão de gRNA ou um plasmídeo de promotor U6 vazio (para os controles negativos), 750 ng de plasmídeo de expressão de Cas9 e 30 ng de plasmídeo de expressão de Td-tomate usando o 4D Nucleofector System de acordo com as instruções do fabricante (Lonza). Para células HEK293, $1,65 \times 10^5$ células foram transfectadas com 125 ng de plasmídeo de expressão de gRNA ou um plasmídeo de promotor U6 vazio (para o controle negativo), 375 ng de plasmídeo de expressão de Cas9 e 30 ng de um plasmídeo de expressão de Td-tomate usando reagente Lipofectamine LTX, de acordo com as instruções do fabricante (Life Technologies). O DNA genômico foi coletado a partir das células HEK293, U2OS.EGFP ou K562 transfectadas usando o QIAamp DNA Blood Mini Kit (QIAGEN), de acordo com as instruções do fabricante. Para gerar DNA genômico suficiente para amplificar os sítios candidatos fora do alvo, o DNA de três Nucleofecções (para células U2OS.EGFP), duas Nucleofecções (para células K562) ou duas transfecções com Lipofectamine LTX foi reunido junto antes de realização de T7EI. Isto foi realizado duas vezes para cada condição testada, gerando conjuntos duplicados de DNA genômico que representam um total de quatro ou seis transfecções

individuais. PCR foi, então, realizada usando estes DNAs genômicos como modelos conforme descrito acima e purificados usando esferas Ampure XP (Agencourt) de acordo com as instruções do fabricante. Ensaio de T7EI foram realizados conforme descrito anteriormente (Reyon *et al.*, 2012, *supra*).

Sequenciamento de DNA de mutações indel mediadas por NHEJ

[0118] Produtos de PCR purificados usados para o ensaio de T7EI foram clonados no vetor Zero Blunt TOPO (Life Technologies) e os DNA de plasmídeo foram isolados usando um método miniprep de lise alcalina pelo MGH DNA Automation Core. Os plasmídeos foram sequenciados usando um iniciador direto M13 (5' – GTAAAACGACGGCCAG – 3' (SEQ ID NO:1059) através do método de Sanger (MGH DNA Sequencing Core).

Exemplo 1a. Desemparelhamentos de um Nucleotídeo Individual

[0119] Para começar a definir os determinantes de especificidade de RGNs em células humanas, foi realizado um teste em grande escala para avaliar os efeitos de desemparelhamento sistematicamente em várias posições dentro de múltiplas interfaces de gRNA/DNA alvo. Para fazer isso, um ensaio de ruptura de proteína fluorescente verde intensificada (**E**nhanced **G**reen **F**luorescent **P**rotein - EGFP) com base em células humanas quantitativo previamente descrito (vide métodos acima e Reyon *et al.*, 2012, *supra*), o qual permite a quantificação rápida de atividade de nucleases objetivadas (**Figura 2B**) foi usado. Neste ensaio, a atividade de nucleases objetivadas a um único gene repórter *EGFP* integrado pode ser quantificada ao avaliar a perda de sinal de fluorescência em células U2OS.EGFP humanas causada por mutações de inserção/deleção (indel) de inativação por desvio de quadro introduzidas por meio de reparo de união terminal não homólogo

propenso a erros (**Non-Homologous End-Joining** - NHEJ) de rupturas fita dupla induzidas por nuclease (**Double-Stranded Break** - DSB) (**Figura 2B**). Para os estudos descritos aqui, três gRNAs individuais de ~ 100 nt objetivados a diferentes sequências dentro de *EGFP* foram usados, como segue:

Sítio 1 de EGFP GGGCACGGGCAGCTTGCCGGTGG
(SEQ ID NO:9)

Sítio 2 de EGFP GATGCCGTTCTTCTGCTTGTCGG
(SEQ ID NO:10)

Sítio 3 de EGFP GGTGGTGCAGATGAACTTCAGGG
(SEQ ID NO:11)

[0120] Cada um destes gRNAs pode comandar eficientemente a ruptura mediada por Cas9 de expressão de EGFP (vide **Exemplos 1e e 2a** e **Figuras 3E** (superior) e **3F** (superior)).

[0121] Em experimentos iniciais, os efeitos desemparelhamentos de um único nucleotídeo em 19 de 20 nucleotídeos da região de objetivação complementar de três gRNAs objetivados ao *EGFP* foram testados. Para fazer isso, gRNAs variantes foram gerados para cada um dos três sítios alvo trazendo desemparelhamentos por transversão de Watson-Crick nas posições 1 a 19 (numeradas de 1 a 20 de 3' para 5'; vide **Figura 1**) e as capacidades destes vários gRNAs de comandar a ruptura de *EGFP* mediada por Cas9 em células humanas testadas (gRNAs variantes trazendo uma substituição na posição 20 não foram gerados porque este nucleotídeo é parte da sequência do promotor U6 e, portanto, deve permanecer uma guanina para evitar afetar a expressão).

[0122] Para o sítio #2 de EGFP, desemparelhamentos individuais nas posições 1 - 10 do gRNA têm efeitos dramáticos sobre a atividade de Cas9 associada (**Figura 2C**, painel central), consistente com estudos anteriores que sugerem que desemparelhamentos na extremidade 5' de

gRNAs são melhor tolerados do que aqueles na extremidade 3' (Jiang *et al.*, Nat Biotechnol 31, 233-239 (2013); Cong *et al.*, Science 339, 819-823 (2013); Jinek *et al.*, Science 337, 816-821 (2012)). No entanto, com sítios alvo #1 e #3 de EGFP, desemparelhamentos individuais em geral, exceto quanto a algumas posições no gRNA, pareceram ser bem tolerados, mesmo dentro da extremidade 3' da sequência. Além disso, as posições específicas que eram sensíveis a desemparelhamentos diferiam para estes dois alvos (**Figura 2C**, comparar os painéis superior e inferior). Por exemplo, o sítio alvo #1 era particularmente sensível a um desemparelhamento na posição 2, enquanto que o sítio alvo #3 era mais sensível a desemparelhamentos nas posições 1 e 8.

Exemplo 1b. Desemparelhamentos Múltiplos

[0123] Para testar os efeitos de mais de um desemparelhamento na interface gRNA/DNA, uma série de gRNAs variantes de trazendo desemparelhamentos duplos por transversão de Watson-Crick em posições adjacentes e separadas foram criados e as capacidades destes gRNAs de comandar a atividade de nuclease Cas9 foram testadas em células humanas usando o ensaio de ruptura de EGFP. Todos os três sítios alvo mostraram, em geral, uma maior sensibilidade a alterações duplas nas quais um ou ambos os desemparelhamentos ocorrem dentro da metade 3' da região de objetivação do gRNA. No entanto, a magnitude destes efeitos exibiu variação específica para o sítio, com o sítio alvo #2 mostrando a maior sensibilidade a estes desemparelhamentos duplos e o sítio alvo #1 mostrando geralmente menos. Para testar o número de desemparelhamentos adjacentes que podem ser tolerados, gRNAs variantes foram construídos trazendo números crescentes de posições não pareadas que vão desde as posições 19 a 15 na extremidade 5' da região de objetivação do gRNA (onde desemparelhamentos individuais e duplos pareciam ser melhor tolerados).

[0124] Testagem destes gRNAs com números crescentes de desemparelhamentos revelou que, para todos os três sítios alvo, a introdução de três ou mais desemparelhamentos adjacentes resulta em perda significativa de atividade de RGN. A queda súbita na atividade ocorreu para três diferentes gRNAs objetivados à EGFP a medida que se faz desemparelhamentos progressivos a partir de posição 19 na extremidade 5' e se adiciona mais desemparelhamentos que se movem em direção à extremidade 3'. Especificamente, gRNAs contendo desemparelhamentos nas posições 19 e 19+18 mostram atividade essencialmente completa, enquanto que aqueles com desemparelhamentos nas posições 19+18+17, 19+18+17+16 e 19+18+17+16+15 não mostram essencialmente nenhuma diferença em relação a um controle negativo (**Figura 2F**). (Note que nós não fizemos desemparelhamentos na posição 20 em nestes gRNAs variantes porque esta posição precisa de permanecer como um G, pois faz parte do promotor U6 que comanda a expressão do gRNA).

[0125] Prova adicional de que o encurtamento de complementaridade do gRNA pode levar a RGNs com maiores especificidades foi obtida a partir do seguinte experimento: para quatro gRNAs diferentes que objetivam EGFP (**Figura 2H**), a introdução de um desemparelhamento duplo nas posições 18 e 19 não impactou significativamente a atividade. No entanto, a introdução de outro desemparelhamento duplo nas posições 10 e 11, então, nestes gRNAs, resulta em perda quase completa de atividade. Curiosamente, a introdução apenas dos desemparelhamentos duplos 10/11 não tem, em geral, um impacto tão grande sobre a atividade.

[0126] Tomados em conjunto, estes resultados em células humanas confirmam que as atividades de RGNs podem ser mais sensíveis a desemparelhamentos na metade 3' da sequência de objetivação do gRNA. No entanto, os dados revelam também

claramente que a especificidade de RGNs é complexa e dependente do sítio alvo, com desemparelhamentos individuais e duplos frequentemente bem tolerados, mesmo quando um ou mais desemparelhamentos ocorrem na metade 3' da região de objetivação do gRNA. Além disso, estes dados sugerem também que nem todos os desemparelhamentos na metade 5' da interface gRNA/DNA são necessariamente bem tolerados.

[0127] Além disso, estes resultados sugerem fortemente que gRNAs tendo regiões de complementaridade mais curtas (especificamente ~ 17 nts) serão mais específicos quanto às suas atividades. Notamos que 17 nts de especificidade combinados com os 2 nts de especificidade conferidos pela sequência PAM resulta em especificação de uma sequência de 19 pb, um comprimento suficiente para ser único em grandes genomas complexos, tais como aqueles encontrados em células humanas.

Exemplo 1c. Mutações fora do Alvo

[0128] Para determinar se podiam ser identificadas mutações fora do alvo para RGNs objetivados a genes humanos endógenos, seis gRNAs individuais que objetivam três sítios diferentes no gene *VEGFA*, um no gene *EMX1*, um no gene *RNF2* e um no gene *FANCF* foram usados (**Tabela 1** e **Tabela A**). Estes seis gRNAs comandaram eficientemente indels mediadas por Cas9 em seu respectivo *locus* endógeno em células U2OS.EGFP humanas, conforme detectado pelo ensaio de Endonuclease I T7 (T7EI) (métodos acima e **Tabela 1**). Para cada um destes seis RGNs, então, nós examinamos dezenas de potenciais sítios fora do alvo (que variam quanto ao número de 46 até 64) para evidencia de mutações indel mediadas por NHEJ induzidas por nucleases em células U2OS.EGFP. Os *loci* avaliados incluíam todos os sítios genômicos que diferem em um ou dois nucleotídeos, bem como subconjuntos de sítios genômicos que diferem em três a seis

nucleotídeos e com uma tendência para aqueles que tinham um ou mais destes desemparelhamentos na metade 5' da sequência de objetivação do gRNA (**Tabela B**). Usando o ensaio de T7EI, quatro sítios fora do alvo (de um total de 53 sítios candidatos examinados) para o sítio 1 de *VEGFA*, doze (de 46 analisados) para o sítio 2 de *VEGFA*, sete (de 64 analisados) para o sítio 3 de *VEGFA* e um (de 46 analisados) para o sítio de *EMX1* (**Tabela 1** e **Tabela B**) foram prontamente identificados. Nenhuma mutação fora do alvo foi detectada dentre os 43 e os 50 sítios potenciais examinados para os genes *RNF2* ou *FANCF*, respectivamente (**Tabela B**). As taxas de mutações em sítios fora do alvo verificados foram muito altas, variando de 5,6% a 125% (média de 40%) da taxa observada no sítio alvo pretendido (**Tabela 1**). Estes sítios fora do alvo *bona fide* incluíam sequências com desemparelhamentos na extremidade 3' do sítio alvo e com tanto quanto um total de cinco desemparelhamentos, com a maioria dos sítios fora do alvo ocorrendo dentro de genes que codificam proteínas (**Tabela 1**). Sequenciamento de DNA de um subconjunto de sítios fora do alvo forneceu confirmação molecular adicional de que mutações indel ocorrem no sítio de clivagem de RGN esperado (**Figuras 8A-C**).

Tabela 1. Mutações sobre e fora do alvo induzidas por RGNs concebidos para genes humanos endógenos

Alvo	Nome sítio	Sequência	SEQ ID NO:	Frequência de mutação indel (%) ± SEM			Gene
				U2OS.EGFP	HEK293	K562	
Alvo 1 (Sítio 1 de VEGFA)	T1	GGGTGGGGGAGTTGCTCCTGG	1059.	26,0 ± 2,9	10,5 ± 0,07	3,33 ± 0,42	VEGFA
	OT1-3	GG A TGG A GGGAGTTGCTCCTGG	1060.	25,7 ± 9,1	18,9 ± 0,77	2,93 ± 0,04	IGDCC3
	OT1-4	GGG A GGG I GGAGTTGCTCCTGG	1061.	9,2 ± 0,8	8,32 ± 0,51	N.D.	LOC116437
	OT1-6	C GGG G GA A GGGAGTTGCTCCTGG	1062.	5,3 ± 0,2	3,67 ± 0,09	N.D.	CACNA2D
	OT1-11	GGG G A A GGG A AGTTGCTCCTGG	1063.	17,1 ± 4,7	8,54 ± 0,16	N.D.	
Alvo 2 (Sítio 2 de VEGFA)	T2	GACCCCTCCACCCGCCTCCGG	1064.	50,2 ± 4,9	38,6 ± 1,92	15,0 ± 0,25	VEGFA
	OT2-1	GACCC C CCACCCCGCCCCCGG	1065.	14,4 ± 3,4	33,6 ± 1,17	4,10 ± 0,05	FMN1
	OT2-2	G GGCCCTCCACCCGCCTCTGG	1066.	20,0 ± 6,2	15,6 ± 0,30	3,00 ± 0,06	PAX6
	OT2-6	C TACCCCTCCACCCGCCTCCGG	1067.	8,2 ± 1,4	15,0 ± 0,64	5,24 ± 0,22	PAPD7
	OT2-9	G CC C CC A CCACCCGCCTCTGG	1068.	50,7 ± 5,6	30,7 ± 1,44	7,05 ± 0,48	LAMA3
	OT2-15	I ACCC C CA C ACCCGCCTCTGG	1069.	9,7 ± 4,5	6,97 ± 0,10	1,34 ± 0,15	SPNS3
	OT2-17	A CACCC C CCACCCGCCTCAGG	1070.	14,0 ± 2,8	12,3 ± 0,45	1,80 ± 0,03	
	OT2-19	A TTCC C CCACCCGCCTCAGG	1071.	17,0 ± 3,3	19,4 ± 1,35	N.D.	HDLBP
	OT2-20	C CC C ACCCACCCGCCTCAGG	1072.	6,1 ± 1,3	N.D.	N.D.	ABLM1
	OT2-23	C GGCC I CCACCCGCCTCCGG	1073.	44,4 ± 6,7	28,7 ± 1,15	4,18 ± 0,37	CALY
	OT2-24	C TC C CC A CCACCCGCCTCAGG	1074.	62,8 ± 5,0	29,8 ± 1,08	21,1 ± 1,68	
	OT2-29	I GCC C TC C ACCCGCCTCTGG	1075.	13,8 ± 5,2	N.D.	N.D.	ACLY
	OT2-34	A GGCC C CA C ACCCGCCTCAGG	1076.	2,8 ± 1,5	N.D.	N.D.	
Alvo 3 (Sítio 3 de VEGFA)	T3	GGTGAGTGAGTGTGCGTGTGG	1077.	49,4 ± 3,8	35,7 ± 1,26	27,9 ± 0,52	VEGFA
	OT3-1	GGTGAGTGAGTGTG I GTGAGG	1078.	7,4 ± 3,4	8,97 ± 0,80	N.D.	(abParts)
	OT3-2	A GTGAGTGAGTGTG I GTGGG	1079.	24,3 ± 9,2	23,9 ± 0,08	8,9 ± 0,16	MAX
	OT3-4	G C TGAGTGAGTGT A TGCGTGTGG	1080.	20,9 ± 11,8	11,2 ± 0,23	N.D.	
	OT3-9	GGTGAGTGAGTG C GTGCG G GTGG	1081.	3,2 ± 0,3	2,34 ± 0,21	N.D.	TPCN2
	OT3-17	G I TGAGTGA A TGTGTGCGTGAGG	1082.	2,9 ± 0,2	1,27 ± 0,02	N.D.	SLIT1
	OT3-18	I GTG G GTGAGTGTGTGCGTGAGG	1083.	13,4 ± 4,2	12,1 ± 0,24	2,42 ± 0,07	COMDA
	OT3-20	A GA G AGTGAGTGTGTG C ATGAGG	1084.	16,7 ± 3,5	7,64 ± 0,05	1,18 ± 0,01	
Alvo 4 (EMX1)	T4	GAGTCCGAGCAGAAGAAGAGG	1085.	42,1 ± 0,4	26,0 ± 0,70	10,7 ± 0,50	EMX1
	OT4-1	GAGT I AGAGCAGAAGAAGAGG	1086.	16,8 ± 0,2	8,43 ± 1,32	2,54 ± 0,02	HCN1
Alvo 5 (RNF2)	T5	GTCATCTTAGTCATTACCTGTGG	1087.	26,6 ± 6,0	---	---	RNF2
Alvo 6 (FANCF)	T6	GGAATCCCTTCTGCAGCACCCAGG	1088.	33,2 ± 6,5	---	---	FANCF

"OT" indica sítios fora do alvo (com numeração de sítios conforme na **Tabela E**). Desemparelhamentos sobre o alvo (dentro da região de 20 pb à qual o gRNA hibridiza) são destacados como negrito, sublinhado. Frequências de mutação indel em células U2OS.EGFP, HEK293 e K562 foram determinadas conforme descrito em Métodos. Os genes em sítios que foram localizados (se houver) são mostrados. Todos os sítios listados não mostram qualquer evidência de modificação nas células transfectadas com plasmídeo de expressão de Cas9 e um plasmídeo de promotor U6 de controle que não expressam um gRNA funcional. N.D. = nenhum detectado; --- = Não testado.

Exemplo 1d. Mutações fora do alvo em outros tipos de células

[0129] Tendo estabelecido que RGNs podem induzir a mutações fora do alvo com altas frequências em células U2OS.EGFP, em seguida, nós procuramos determinar se estas nucleases também teriam estes efeitos sobre outros tipos de células humanas. Nós escolhemos células U2OS.EGFP para nossos experimentos iniciais, porque estas células foram usadas anteriormente para avaliar as atividades de TALENs¹⁵, mas as células HEK293 e K562 humanas têm sido mais amplamente

usadas para testar as atividades de nucleases objetivadas. Portanto, nós também avaliamos as atividades das quatro RGNs objetivadas aos sítios 1, 2 e 3 de *VEGFA* e ao sítio *EMX1* em células HEK293 e K562. Descobriu-se que cada uma destas quatro RGNs induziram eficientemente a mutações indel mediadas por NHEJ em seus sítios alvo pretendidos nestas duas linhagens de células humanas adicionais (conforme avaliado pelo ensaio de T7EI) (**Tabela 1**), embora com frequências de mutação um tanto mais baixas do que aquelas observadas em células U2OS.EGFP. Avaliação dos 24 sítios fora do alvo para estas quatro RGNs originalmente identificadas em células U2OS.EGFP revelou que muitos tinham nova mutação em células HEK293 e K562 com frequências similares àquelas em seus sítios sobre o alvo correspondentes (**Tabela 1**). Conforme esperado, sequenciamento de DNA de um subconjunto destes sítios fora do alvo a partir de células HEK293 forneceu evidência molecular adicional de que alterações ocorrem no *locus* genômico esperado (**Figuras 9A-C**). Não sabemos ao certo por que, em células HEK293, quatro e, em células K562, onze dos sítios fora do alvo identificados em células U2OS.EGFP não mostraram mutações detectáveis. No entanto, notamos que muitos desses sítios fora do alvo também mostravam frequências de mutação relativamente mais baixas em células U2OS.EGFP. Portanto, especula-se que as taxas de mutação destes sítios em células HEK293 e K562 pode cair abaixo do limite de detecção confiável de nosso ensaio de T7EI (~ 2-5%) porque RGNs geralmente parecem ter atividades mais baixas em células HEK293 e K562 comparado com células U2OS.EGFP em nossos experimentos. Tomados em conjunto, nossos resultados em células HEK293 e K562 fornecem evidências de que a alta frequência de mutações fora do alvo que observamos com RGNs será um fenômeno geral visto em vários tipos de células humanas.

Exemplo 1e. Titulação de quantidades de plasmídeo que expressam gRNA e Cas9 usados para o ensaio de ruptura de EGFP

[0130] gRNAs individuais foram gerados para três sequências diferentes (sítios 1-3 de EGFP, mostrados acima) localizadas a montante do nucleotídeo 502 de *EGFP*, uma posição na qual a introdução de mutações de desvio de quadro através de união terminal não homóloga pode romper robustamente a expressão de EGFP (Maeder, M.L. *et al.*, Mol Cell 31, 294-301 (2008); Reyon, D. *et al.*, Nat Biotech 30, 460-465 (2012)).

[0131] Para cada um dos três sítios alvo, uma faixa de quantidades de plasmídeos que expressam gRNA (12,5 a 250 ng) foi inicialmente transfectada, juntamente com 750 ng de um plasmídeo que expressa uma versão com códons otimizados da nuclease Cas9 em nossas células repórter U2OS.EGFP trazendo uma única cópia do gene repórter *EGFP-PEST* constitutivamente expresso. Todas as três RGNs romperam eficientemente a expressão de EGFP na concentração mais elevada de plasmídeo de codificação de gRNA (250 ng) (**Figura 3E** (superior)). No entanto, RGNs para os sítios alvo #1 e #3 exibiram níveis equivalentes de ruptura quando quantidades menores de plasmídeo expressando gRNA foram transfectadas, enquanto que a atividade de RGN no sítio alvo #2 caiu imediatamente quando a quantidade de plasmídeo expressando gRNA transfectado foi diminuída (**Figura 3E** (superior)).

[0132] A quantidade de plasmídeo que codifica Cas9 (variando de 50 ng a 750 ng) transfectado em nossas células repórter U2OS.EGFP foi titulada e a ruptura de EGFP ensaiada. Conforme mostrado na **Figura 3F** (superior), o sítio alvo #1 tolerou uma redução de três vezes na quantidade de plasmídeo que codifica Cas9 transfectado sem perda substancial de atividade de ruptura de EGFP. No entanto, as atividades de RGNs que objetivam os sítios alvo #2 e #3 diminuiu imediatamente,

com uma redução de três vezes na quantidade de plasmídeo de Cas9 transfectado (**Figura 3F** (superior)). Com base nestes resultados, 25 ng/250 ng, 250 ng/750 ng e 200 ng/750 ng de plasmídeos que expressam gRNA/Cas9 foram usados para os sítios alvo #1, #2 e #3 de *EGFP*, respectivamente, para os experimentos descritos nos Exemplos 1a -1d.

[0133] As razões pelas quais algumas combinações de gRNA/Cas9 funcionam melhor do que outras na ruptura de expressão de *EGFP* não é compreendida, nem porque algumas destas combinações são mais ou menos sensíveis à quantidade de plasmídeos usados para transfecção. Embora seja possível que a variedade de sítios fora do alvo presentes no genoma para estes três gRNAs possa influenciar cada uma de suas atividades, não foram observadas diferenças no número de sítios genômicos que diferem em um a seis bp para cada um destes sítios alvo particulares (**Tabela C**), que seriam responsáveis pelo comportamento diferencial dos três gRNAs.

Tabela C. Números de sítios fora do alvo no genoma humano para seis RGNs objetivadas a genes humanos endógenos e três RGNs objetivadas ao gene repórter *EGFP*

Sítio alvo	Número de desemparelhamentos sobre sítio sobre o alvo						
	0	1	2	3	4	5	6
Alvo 1 (<i>VEGFA</i> sítio 1)	1	1	4	32	280	2175	13873
Alvo 2 (<i>VEGFA</i> sítio 2)	1	0	2	35	443	3889	17398
Alvo 3 (<i>VEGFA</i> sítio 3)	1	1	17	377	6028	13398	35517
Alvo 4 (<i>EMX</i>)	1	0	1	18	276	2309	15731
Alvo 5 (<i>RNF2</i>)	1	0	0	6	116	976	7443
Alvo 6 (<i>FANCF</i>)	1	0	1	18	271	1467	9551
Sítio alvo #1 de <i>EGFP</i>	0	0	3	10	156	1365	9755
Sítio alvo #2 de <i>EGFP</i>	0	0	0	11	96	974	7353
Sítio alvo #3 de <i>EGFP</i>	0	0	1	14	165	1439	10361

Sítios fora do alvo para cada uma das seis RGNs objetivadas aos genes *VEGFA*, *RNF2*, *FANCF* e *EMX1* e as três RGNs objetivadas aos sítios alvo #1, #2 e #3 de *EGFP* foram

identificados na sequência do genoma humano, construto GRCh37. Desemparelhamentos foram permitidos apenas para a região de 20 nt à qual o gRNA hibridiza e não na sequência PAM.

Exemplo 2: Encurtamento de comprimento complementaridade de gRNA para melhorar a especificidade de clivagem de RGN

[0134] Postula-se que os efeitos fora do alvo de RGNs pode ser minimizado, sem comprometer a atividade sobre o alvo, simplesmente ao diminuir o comprimento da interface gRNA-DNA, uma abordagem que pode parecer, a princípio, contra-intuitiva. gRNAs mais longos podem realmente funcionar com menos eficiência sobre sítios sobre o alvo (vide abaixo e Hwang *et al.*, 2013a; Ran *et al.*, 2013). Em contraste, conforme mostrado acima no Exemplo 1, gRNAs tendo múltiplos desemparelhamentos em suas extremidades 5' ainda podem induzir à clivagem robusta de seus sítios alvo (**Figuras 2A e 2C-2F**), sugerindo que estes nucleotídeos podem não ser necessários para atividade completa sobre o alvo. Portanto, foi especulado de que gRNAs truncados carecendo destes nucleotídeos 5' podem mostrar atividades comparáveis aos gRNAs de comprimento completo (**Figura 2A**). Foi especulado que, se os nucleotídeos 5' dos gRNAs de comprimento completo não são necessários para a atividade sobre o alvo, então, sua presença pode também compensar os desemparelhamentos em outras posições ao longo da interface gRNA/DNA alvo. Se isso fosse verdade, foi levantada a hipótese de que gRNAs podem ter maior sensibilidade a desemparelhamentos e, portanto, também pode induzir a níveis substancialmente mais baixos de mutações fora do alvo mediadas por Cas9 (**Figura 2A**).

Procedimentos Experimentais

[0135] Os procedimentos experimentais a seguir foram usados no Exemplo 2.

Construção de Plasmídeo

[0136] Todos os plasmídeos de expressão de gRNA foram montados através de concepção, síntese, recozimento e clonagem de pares de oligonucleotídeos (IDT) abrigando a região de complementaridade no plasmídeo pMLM3636 (disponível da Addgene), conforme descrito acima (**Exemplo 1**). Os vetores de expressão de gRNA resultantes codificam um ~ 100 nt gRNA cuja expressão é comandada por um promotor U6 humano. As sequências de todos os oligonucleotídeos usados para a construção de vetores de expressão de gRNA são apresentadas na **Tabela D**. O plasmídeo de expressão de nickase D10A de Cas9 (pJDS271) tendo uma mutação no domínio de endonuclease RuvC foi gerado por meio de mutação do plasmídeo pJDS246 usando um kit QuikChange (Agilent Technologies) com as seguintes iniciadores: iniciador senso D10A de Cas9 5'-tgataaaaagtattctattggttagccatcggcactaattccg-3' (SEQ ID NO: 1089); iniciador antissenso D10A de Cas9 5'-cggaattagtgccgatggctaaaccaatagaatacttttatcca-3' (SEQ ID NO: 1090). Todos os plasmídeos de gRNA objetivados e os plasmídeos de nickase Cas9 usados neste estudo estão disponíveis através do serviço de distribuição de plasmídeo sem fins lucrativos Addgene (addgene.org/crispr-cas).

Tabela D. Sequências de oligonucleotídeos usadas para construir plasmídeos de expressão de gRNA																								
Sítio alvo 1 de EGFP																								
20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')	SEQ ID NO :	oligonucleotídeo 2 (5' a 3')	SEQ ID NO:	
		G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGCACGGG CAGTTGCCGGG	1091	AAAACGGGGCAA GCTGCCCGTGCG	1180.	
		G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	c	ACACCGCACGGG CAGTTGCCGGG	1092	AAAACGGGGCAA GCTGCCCGTGCG	1181.	
		G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	c	G	ACACCGCACGGG CAGTTGCCCGG	1093	AAAACGGGGCAA GCTGCCCGTGCG	1182.	
		G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	g	G	G	ACACCGCACGGG CAGTTGCCGGG	1094	AAAACCCGCAA GCTGCCCGTGCG	1183.	
		G	C	A	C	G	G	G	C	A	G	C	T	T	G	g	C	G	G	ACACCGCACGGG CAGTTGCCGGG	1095	AAAACCCGCCAA GCTGCCCGTGCG	1184.	
		G	C	A	C	G	G	G	C	A	G	C	T	T	c	C	C	G	G	ACACCGCACGGG CAGTTCCCGGG	1096	AAAACCCGGGAA GCTGCCCGTGCG	1185.	
		G	C	A	C	G	G	G	C	A	G	C	T	a	G	C	C	G	G	ACACCGCACGGG CAGTAGCCGGG	1097	AAAACCCGGCTA GCTGCCCGTGCG	1186.	
		G	C	A	C	G	G	G	C	A	G	C	a	T	G	C	C	G	G	ACACCGCACGGG CAGCATGCCGGG	1098	AAAACCCGGCAT GCTGCCCGTGCG	1187.	
		G	C	A	C	G	G	G	C	A	G	g	T	T	G	C	C	G	G	ACACCGCACGGG	1099	AAAACCCGGCAA	1188.	

Tabela D. Sequências de oligonucleotídeos usadas para construir plasmídeos de expressão de gRNA																										
																					CAGGTTGCCGGG			CCTGCCGTGCG		
			G	C	A	C	G	G	G	C	A	c	C	T	T	G	C	C	G	G	ACACCGCACGGG CACGTTGCCGGG	1100	AAACCCGGCAA GTTGCCGTGCG	1189.		
			G	C	A	C	G	G	G	C	t	G	C	T	T	G	C	C	G	G	ACACCGCACGGG CTGCTTGCCGGG	1101	AAACCCGGCAA GCAGCCCGTGC G	1190.		
			G	C	A	C	G	G	G	g	A	G	C	T	T	G	C	C	G	G	ACACCGCACGGG GAGCTTGCCGGG	1102	AAACCCGGCAA GCTCCCCGTGCG	1191.		
			G	C	A	C	G	G	c	C	A	G	C	T	T	G	C	C	G	G	ACACCGCACGGC CAGCTTGCCGGG	1103	AAACCCGGCAA GCTGCCGTGCG G	1192.		
			G	C	A	C	G	c	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGCACGCG CAGCTTGCCGGG	1104	AAACCCGGCAA GCTGCCGTGCG G	1193.		
			G	C	A	C	c	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGCACCGG CAGCTTGCCGGG	1105	AAACCCGGCAA GCTGCCGTGCG G	1194.		
			G	C	A	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGCAGGGG CAGCTTGCCGGG	1106	AAACCCGGCAA GCTGCCCTGCG	1195.		
			G	C	t	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGCTCGGG CAGCTTGCCGGG	1107	AAACCCGGCAA GCTGCCGAGC G	1196.		
			G	g	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGGACGGG CAGCTTGCCGGG	1108	AAACCCGGCAA GCTGCCGTCCG	1197.		
			G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	c	c	ACACCGCACGGG CAGCTTGCCCGG	1109	AAACCGGGCAA GCTGCCGTGCG	1198.		
			G	C	A	C	G	G	G	C	A	G	C	T	T	G	g	g	G	G	ACACCGCACGGG CAGCTTGGGGGG	1110	AAACCCCCCAA GCTGCCGTGCG	1199.		
			G	C	A	C	G	G	G	C	A	G	C	T	a	c	C	C	G	G	ACACCGCACGGG CAGTACCCGGG	1111	AAACCCGGGTA GCTGCCGTGCG	1200.		
			G	C	A	C	G	G	G	C	A	G	g	a	T	G	C	C	G	G	ACACCGCACGGG CAGGATGCCGGG	1112	AAACCCGGGAT CTGCCGTGCG	1201.		
			G	C	A	C	G	G	G	C	t	c	C	T	T	G	C	C	G	G	ACACCGCACGGG CTCCTTGCCGGG	1113	AAACCCGGCAA GGAGCCCGTGC G	1202.		
			G	C	A	C	G	G	c	g	A	G	C	T	T	G	C	C	G	G	ACACCGCACGGC GAGCTTGCCGGG	1114	AAACCCGGCAA GCTGCCGTGCG	1203.		
			G	C	A	C	c	c	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGCACCGG CAGCTTGCCGGG	1115	AAACCCGGCAA GCTGCCGGTGC G	1204.		
			G	C	t	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGCTGGGG CAGCTTGCCGGG	1116	AAACCCGGCAA GCTGCCCAAGCG	1205.		
			G	g	t	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGGTCCGG CAGCTTGCCGGG	1117	AAACCCGGCAA GCTGCCGACCG	1206.		
Sitio alvo 2 de EGFP																										
20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')		oligonucleotídeo 2 (5' a 3')				
			G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGCGTTTC TTCTGCTTGTG	1118	AAACACAAGCA GAAGAAGGCG	1207.			
			G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	A	ACACCGCGTTTC TTCTGCTTAGG	1119	AAACTCAAGCA GAAGAAGGCG	1208.			
			G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	c	T	ACACCGCGTTTC TTCTGCTTCTG	1120	AAACCAAGCA GAAGAAGGCG	1209.			
			G	C	C	G	T	T	C	T	T	C	T	G	C	T	a	G	T	ACACCGCGTTTC TTCTGCTAGTG	1121	AAACCACTAGCA GAAGAAGGCG	1210.			
			G	C	C	G	T	T	C	T	T	C	T	G	C	a	T	G	T	ACACCGCGTTTC TTCTGCATGTG	1122	AAACACATGCA GAAGAAGGCG	1211.			
			G	C	C	G	T	T	C	T	T	C	T	G	g	T	T	G	T	ACACCGCGTTTC TTCTGGTTGTG	1123	AAACACAACCA GAAGAAGGCG	1212.			
			G	C	C	G	T	T	C	T	T	C	T	c	C	T	T	G	T	ACACCGCGTTTC TTCTCTTGTG	1124	AAACACAAGGA GAAGAAGGCG	1213.			
			G	C	C	G	T	T	C	T	T	C	a	G	C	T	T	G	T	ACACCGCGTTTC TTCAGCTTGTG	1125	AAACACAAGCT GAAGAAGGCG	1214.			
			G	C	C	G	T	T	C	T	T	g	T	G	C	T	T	G	T	ACACCGCGTTTC TTGTGCTTGTG	1126	AAACACAAGCA CAAGAAGGCG	1215.			
			G	C	C	G	T	T	C	T	a	C	T	G	C	T	T	G	T	ACACCGCGTTTC TACTGCTTGTG	1127	AAACACAAGCA GTAGAAGGCG	1216.			
			G	C	C	G	T	T	C	a	T	C	T	G	C	T	T	G	T	ACACCGCGTTTC ATCTGCTTGTG	1128	AAACACAAGCA GATGAAGGCG	1217.			
			G	C	C	G	T	T	g	T	C	T	G	C	T	G	T	G	T	ACACCGCCGTTG TTCTGCTTGTG	1129	AAACACAAGCA GAACAAGGCG	1218.			
			G	C	C	G	T	a	C	T	T	C	T	G	C	T	T	G	T	ACACCGCGGTAC TTCTGCTTGTG	1130	AAACACAAGCA GAAGTACGCG	1219.			
			G	C	C	G	a	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGCCGATC TTCTGCTTGTG	1131	AAACACAAGCA GAAGATCGCG	1220.			

Tabela D. Sequências de oligonucleotídeos usadas para construir plasmídeos de expressão de gRNA																								
			G	C	C	c	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGCCCTTC TTCTGCTTGTG	1132	AAACACAAGCA GAAGAAGGGCG	1221.	
			G	C	g	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGCGGTTTC TTCTGCTTGTG	1133	AAACACAAGCA GAAGAACC GCG	1222.	
			G	g	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGGGCTTC TTCTGCTTGTG	1134	AAACACAAGCA GAAGAACGCCG	1223.	
			G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	c	a	ACACCGCGCTTC TTCTGCTTCAG	1135	AAACTGAAGCA GAAGAACGGCG	1224.	
			G	C	C	G	T	T	C	T	T	C	T	G	C	a	a	G	T	ACACCGCGCTTC TTCTGCAAGTG	1136	AAACACTTGCAG GAAGAACGGCG	1225.	
			G	C	C	G	T	T	C	T	T	C	T	c	g	T	T	G	T	ACACCGCCGTTTC TTCTGCTTGTG	1137	AAACACAACGA GAAGAACGGCG	1226.	
			G	C	C	G	T	T	C	T	T	g	a	G	C	T	T	G	T	ACACCGCCGTTTC TTGAGCTTGTG	1138	AAACACAAGCT CAAGAACGGCG	1227.	
			G	C	C	G	T	T	C	a	a	C	T	G	C	T	T	G	T	ACACCGCCGTTTC AACTGCTTGTG	1139	AAACACAAGCA GTTGAACGGCG	1228.	
			G	C	C	G	T	a	g	T	T	C	T	G	C	T	T	G	T	ACACCGCGTAG TTCTGCTTGTG	1140	AAACACAAGCA GAACTGCGCG	1229.	
			G	C	C	c	a	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGCCCATC TTCTGCTTGTG	1141	AAACACAAGCA GAAGATGGCG	1230.	
			G	g	g	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGGGGTTTC TTCTGCTTGTG	1142	AAACACAAGCA GAAGAACCCCG	1231.	
Sítio alvo 3 de EGFP																								
2 0	1 9	1 8	1 7	1 6	1 5	1 4	1 3	1 2	1 1	1 0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')		oligonucleotídeo 2 (5' a 3')		
			G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGTGCAG ATGAACCTCAG	1143	AAAACTCTAGTTC ATCTGCACCG	1232.	
			G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	t	ACACCGGTGCAG ATGAACCTCTG	1144	AAAACTCAAGTT CATCTGCACCG	1233.	
			G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	g	A	ACACCGGTGCAG ATGAACCTTGAG	1145	AAAACGTAGATT CATCTGCACCG	1234.	
			G	G	T	G	C	A	G	A	T	G	A	A	C	T	a	C	A	ACACCGGTGCAG ATGAACACAG	1146	AAAACGTAGTGT CATCTGCACCG	1235.	
			G	G	T	G	C	A	G	A	T	G	A	A	C	a	T	C	A	ACACCGGTGCAG ATGAACATCAG	1147	AAAACGAACTT CATCTGCACCG	1236.	
			G	G	T	G	C	A	G	A	T	G	A	A	g	T	T	C	A	ACACCGGTGCAG ATGAAGTTACG	1148	AAAACGGAAGAT CATCTGCACCG	1237.	
			G	G	T	G	C	A	G	A	T	G	A	t	C	T	T	C	A	ACACCGGTGCAG ATGATCTTCAG	1149	AAAACGGAAGTA CATCTGCACCG	1238.	
			G	G	T	G	C	A	G	A	T	G	t	A	C	T	T	C	A	ACACCGGTGCAG ATGATCTTCAG	1150	AAAACGGAAGTT GATCTGCACCG	1239.	
			G	G	T	G	C	A	G	A	T	c	A	A	C	T	T	C	A	ACACCGGTGCAG ATCAACTTCAG	1151	AAAACGGAAGTT CTTCTGCACCG	1240.	
			G	G	T	G	C	A	G	A	a	G	A	A	C	T	T	C	A	ACACCGGTGCAG AAGAACTTCAG	1152	AAAACGGAAGTT CAACTGCACCG	1241.	
			G	G	T	G	C	A	G	t	T	G	A	A	C	T	T	C	A	ACACCGGTGCAG TTGAACCTCAG	1153	AAAACGGAAGTT CATGTGCACCG	1242.	
			G	G	T	G	C	A	c	A	T	G	A	A	C	T	T	C	A	ACACCGGTGCAC ATGAACCTCAG	1154	AAAACGGAAGTT CATCAGCACCG	1243.	
			G	G	T	G	C	t	G	A	T	G	A	A	C	T	T	C	A	ACACCGGTGCTG ATGAACCTCAG	1155	AAAACGGAAGTT CATCTCCACCG	1244.	
			G	G	T	G	g	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGTGGAG ATGAACCTCAG	1156	AAAACGGAAGTT CATCTGGACCG	1245.	
			G	G	T	c	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGTCCAG ATGAACCTCAG	1157	AAAACGGAAGTT CATCTGTCCCG	1246.	
			G	G	a	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGAGCAG ATGAACCTCAG	1158	AAAACGGAAGTT CATCTGCAGCG	1247.	
			G	c	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGTGCAG ATGAACCTCAG	1159	AAAACGGAAGTT CATCTGCAGCG	1248.	
			G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	g	t	ACACCGGTGCAG ATGAACCTTGTG	1160	AAACACAAGTT CATCTGCACCG	1249.	
			G	G	T	G	C	A	G	A	T	G	A	A	C	a	a	C	A	ACACCGGTGCAG ATGAACCAACAG	1161	AAAACGTTGTTTC ATCTGCACCG	1250.	
			G	G	T	G	C	A	G	A	T	G	A	t	g	T	T	C	A	ACACCGGTGCAG ATGATGTTACG	1162	AAAACGAACAT CATCTGCACCG	1251.	
			G	G	T	G	C	A	G	A	T	c	t	A	C	T	T	C	A	ACACCGGTGCAG ATCTACTACCG	1163	AAAACGGAAGTA GATCTGCACCG	1252.	
			G	G	T	G	C	A	G	t	a	G	A	A	C	T	T	C	A	ACACCGGTGCAG TAGAACTTCAG	1164	AAAACGGAAGTT CTACTGCACCG	1253.	
			G	G	T	G	C	t	c	A	T	G	A	A	C	T	T	C	A	ACACCGGTGCTC ATGAACCTCAG	1165	AAAACGGAAGTT CATGAGCACCG	1254.	
			G	G	T	c	g	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGTCCAG ATGAACCTCAG	1166	AAAACGGAAGTT CATCTGCACCG	1255.	
			G	c	a	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCAGCAG ATGAACCTCAG	1167	AAAACGGAAGTT CATCTGCTGCG	1256.	

Tabela D. Sequências de oligonucleotídeos usadas para construir plasmídeos de expressão de gRNA																									
Alvo 1 endógeno (tru-gRNA ao Sítio 1 de VEGFA):																									
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5
			G	T	G	G	G	G	G	A	G	T	T	T	G	C	T	C	C						
Alvo 3 endógeno (tru-gRNA ao Sítio 3 de VEGFA):																									
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5
				G	A	G	T	G	A	G	T	G	T	G	T	G	C	G	T	G					
Alvo 4 endógeno (tru-gRNA ao sítio 1 de EMX1):																									
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5
			G	T	C	C	G	A	G	C	A	G	A	A	G	A	A	G	A	A					
gRNA de CTLA de comprimento completo																									
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5
			G	A	T	G	T	A	G	T	G	T	T	T	C	C	A	C	A						
tru-gRNA de CTLA																									
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5
			G	A	T	G	T	A	G	T	G	T	T	T	C	C	A	C	A						
gRNA de comprimento completo do sítio 4 de VEGFA																									
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5
			T	C	C	C	T	C	T	T	A	G	C	C	A	G	A	G	C	C	G				
gRNA de comprimento completo do sítio 2 de EMX1																									
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5
			G	C	C	G	T	T	T	G	T	A	C	T	T	T	G	T	C	C	T	C			
tru-gRNA do sítio 2 de EMX1																									
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5
			G	T	T	T	G	T	A	C	T	T	T	G	T	C	C	T	C						
gRNA de comprimento completo do sítio 3 de EMX1																									
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5
			G	G	G	A	A	G	A	C	T	G	A	G	G	C	T	A	C	A	T	A			
tru-gRNA do sítio 3 de EMX1																									
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5
			G	A	A	G	A	C	T	G	A	G	G	C	T	A	C	A	T	A					
gRNA de comprimento completo do sítio 4 de EMX1																									
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5
			G	A	G	G	C	C	C	C	A	G	A	G	C	A	G	C	C	A	C				
tru-gRNA do sítio 4 de EMX1																									
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5
			G	C	C	C	C	C	A	G	A	G	C	A	G	C	C	A	C						

Ensaio de ruptura de EGFP com base em células

humanas

[0137] Células U2OS.EGFP que trazem uma cópia única do gene repórter EGFP-PEST integrado foram anteriormente descritas (Reyon

et al., 2012). Estas células foram mantidas em Advanced DMEM (Life Technologies) suplementado com FBS a 10%, GlutaMax a 2 mM (Life Technologies), penicilina/estreptomicina e 400 µg/ml de G418. Para ensaiar a ruptura de expressão de EGFP, 2×10^5 células U2OS.EGFP foram transfectadas em duplicata com plasmídeo de expressão de gRNA ou um plasmídeo de promotor U6 vazio como um controle negativo, plasmídeo de expressão de Cas9 (pJDS246) (**Exemplo 1** e Fu *et al.*, 2013) e 10 ng de plasmídeo de expressão de Td-Tomate (para controlar a eficiência de transfecção), usando um LONZA 4D-Nucleofector™, com uma solução SE e programa DN100 de acordo com as instruções do fabricante. Nós usamos 25 ng/250 ng, 250 ng/750 ng, 200 ng/750 ng e 250 ng/750 ng de plasmídeo de expressão de gRNA / plasmídeo de expressão de Cas9 para os experimentos com os sítios #1, #2, #3 e #4 de de EGFP, respectivamente. Dois dias após transfecção, as células foram tripsinizadas e ressuspensas em meio de Eagle modificado por Dulbecco (DMEM, Invitrogen) suplementado com soro fetal de bovino (**Fetal Bovine Serum - FBS**) a 10% (vol/vol) e analisadas em um citômetro de fluxo BD LSRII. Para cada amostra, transfecções e medições de citometria de fluxo foram realizadas em duplicata.

Transfecção de células humanas e isolamento de DNA genômico

[0138] Para avaliar as mutações indel sobre o alvo e fora do alvo induzidas por RGNs objetivadas aos genes humanos endógenos, plasmídeos foram transfectados em células HEK293 ou U2OS.EGFP usando as seguintes condições: as células U2OS.EGFP foram transfectadas usando as mesmas condições conforme para o ensaio de ruptura de EGFP descrito acima. As células HEK293 foram transfectadas cultivando-as em uma densidade de $1,65 \times 10^5$ células por cavidade em placas com 24 cavidades em Advanced DMEM (Life

Technologies) suplementado com FBS a 10% e GlutaMax a 2 mM (Life Technologies) a 37 °C em uma incubadora de CO₂. Após 22 - 24 horas de incubação, as células foram transfectadas com 125 ng de plasmídeo de expressão de gRNA ou um promotor de plasmídeo U6 vazio (como um controle negativo), 375 ng de plasmídeo de expressão de Cas9 (pJDS246) (**Exemplo 1** e Fu *et al.*, 2013) e 10 ng de um plasmídeo de expressão de Td-Tomate, usando o reagente Lipofectamine LTX de acordo com as instruções do fabricante (Life Technologies). O meio foi trocado 16 horas após transfecção. Para ambos os tipos de células, o DNA genômico foi coletado dois dias após transfecção usando um kit de isolamento de DNA genômico Agencourt DNAdvance (Beckman) de acordo com as instruções do fabricante. Para cada amostra de RGN a ser ensaiada, 12 repetições individuais de transfecção 4D foram realizadas, DNA genômico foi isolado a partir de cada uma destas 12 transfecções e, então, estas amostras foram combinadas para criar dois reservatórios em "duplicata", cada um consistindo em seis amostras de DNA genômico reunidas. Mutações indel foram, então, avaliadas em sítios sobre o alvo e fora do alvo a partir destas amostras em duplicata por meio do ensaio de T7EI, sequenciamento de Sanger e/ou sequenciamento profundo, conforme descrito abaixo.

[0139] Para avaliar as frequências das alterações precisas introduzidas por HDR com modelos doadores ssODN, 2 x 10⁵ células U2OS.EGFP foram transfectadas com 250 ng de plasmídeo de expressão de gRNA ou um plasmídeo de promotor U6 vazio (como um controle negativo), 750 ng de plasmídeo de expressão de Cas9 (pJDS246), 50 pmol de doador ssODN (ou nenhum ssODN para os controles) e 10 ng de plasmídeo de expressão de Td-Tomate (como controle de transfecção). O DNA genômico foi purificado três dias após transfecção usando Agencourt DNAdvance e ensaiado quanto à introdução de um sítio *Bam*HI no sítio de interesse, conforme descrito

abaixo. Todas estas transfecções foram realizadas em duplicata.

[0140] Para experimentos envolvendo nickases Cas9, 2×10^5 células U2OS.EGFP foram transfectadas com 125 ng de cada plasmídeo de expressão de gRNA (se usando gRNAs pareados) ou 250 ng de plasmídeo de expressão de gRNA (se usando um gRNA individual), 750 ng de plasmídeo de expressão de nickase Cas9-D10A (pJDS271), 10 ng de plasmídeo de Td-tomate, e (se realizando HDR), 50 pmol de modelo doador ssODN (que codifica o sítio *Bam*HI). Todas as transfecções foram realizadas em duplicata. O DNA genômico foi coletado dois dias após transfecção (se ensaiando a presença de mutações indel) ou três dias após transfecção (se ensaiando alterações mediadas por HDR/ssODN), usando o kit de isolamento de DNA genômico Agencourt DNAdvance (Beckman).

Ensaio de T7EI para quantificar frequências de mutações indel

[0141] Ensaio de T7EI foram realizados conforme anteriormente descrito (**Exemplo 1** e Fu *et al.*, 2013). Em resumo, reações de PCR para amplificar sítios sobre o alvo ou fora do alvo específicos foram realizadas com DNA polimerase Phusion High-Fidelity (New England Biolabs) usando um dos dois programas a seguir: (1) O programa de PCR Touchdown [(98 °C, 10 s; 72-62 °C, 1 °C/ciclo, 15 s; 72 °C, 30 s) x 10 ciclos, (98 °C, 10 s; 62 °C, 15 s; 72 °C, 30 s) x 25 ciclos] ou (2) o programa de PCR com T_m constante [(98 °C, 10 s; 68 °C ou 72 °C, 15 s; 72 °C, 30 s) x 35 ciclos], com DMSO a 3% ou betaína a 1 M, se necessário. Todos os iniciadores usados para estas amplificações são listados na **Tabela E**. Os produtos de PCR resultantes variaram, quanto ao tamanho, de 300 a 800 pb e foram purificados por esferas Ampure XP (Agencourt) de acordo com as instruções do fabricante. 200 ng de produtos de PCR purificados foram hibridizados em 1x tampão NEB 2 em um volume total de 19 µl e desnaturados para formar heteroduplas

usando as seguintes condições: 95 °C, 5 minutos; 95 a 85 °C e -2 °C/s; 85 até 25 °C, -0.1 °C/s; manter a 4 °C. 1 µL de endonuclease I T7 (New England Biolabs, 10 unidades/µl) foi adicionado para os produtos de PCR hibridizados e incubado a 37 °C durante 15 minutos. A reação de T7EI foi interrompida mediante adição de 2 µl de solução de EDTA a 0,25 M e os produtos da reação foram purificados usando esferas AMPure XP (Agencourt) com eluição em 20 µl de 0,1x tampão EB (Qiagen). Os produtos das reações foram, então, analisados em um sistema de eletroforese capilar QIAxCEL e as frequências das mutações indel foram calculadas usando a mesma fórmula descrita anteriormente (Reyon *et al.*, 2012).

TABELA E										
Pub lica ção ID	Sequências fora do alvo (Esperadas) - HS GRCh37	Desem parelha mentos Em Q sobre o alvo com fora do alvo	Alvo real em células U2OS.EGFP	iniciador de PCR direto	S E Q I Iniciador de PCR reverso	S E Q I Condições de PCR	Tran sver sões de Wat son- Cric k	Tran sver sões de Wat son- Cric k	Tra n si ção s	
Alvo 1	GGGTGGGG GGAGTTTG CTCCTGG	126	0	TCCAGATGGC ACATTGTCAG	127	AGGGAGCAGG AAAGTGAGGT	127	DMSO		
OT1 -1	GGGTGGGG GGAGTTTG CCCCAGG	127	1	GGGGCCCACT CTTCTCCAT	127	ACCCAGACTC CTGGTGTCG	127	No DMSO	0	1
OT1 -2	GCGTGGGG GGTGTGTTG CTCCCCG	127	2	GCTAAGCAGA GATGCCATG CC	127	ACCACCTTTC CCCCAGAAA	127	DMSO	2	0
OT1 -3	GGATGGAG GGAGTTTG CTCCTGG	127	2	ACCCACAGC CAGGTTTCA	127	GAATCACTGCA CCTGGCCATC	128	DMSO	0	2
OT1 -4	GGGAGGGT GGAGTTTG CTCCTGG	128	2	TGCGGCAACT TCAGACAACC	128	TAAAGGGCGT GCTGGAGAG	128	DMSO	1	0
OT1 -5	GGGTGGGT GGAGTTTG CTACTGG	128	2	GCATGTCAGG ATCTGACCCC	128	TGCAGGGCCA TCTTGTGTGT	128	DMSO	0	0
OT1 -6	CGGGGGAG GGAGTTTG CTCCTGG	128	3	CCACCACATG TTCTGGGTGC	128	CTGGGTCTGTT CCCTGTGGG	128	DMSO	1	1
OT1 -7	GAGTGGGT GGAGTTTG CTACAGG	129	3	GGCTCTCCCT GCCCTAGTTT	129	GCAGGTCAAG TTGGAACCCG	129	DMSO	0	1
OT1 -8	GGGAGGGG AGAGTTTGT TCCAGG	129	3	GGGGCTGAGA ACACATGAGA TGCA	129	AGATTTGTGCA CTGCCTGCCT	129	DMSO	1	2
OT1 -9	GGGAGGGG GCAGTTTG CTCCAGG	129	3	CCCGACCTCC GCTCCAAAGC	129	GGACCTCTGC ACACCTGGC	129	DMSO	2	0
OT1 -10	GGGAGGGG GGAGTTGTG TTCCGGG	129	3	TGCAAGGTCTG CATAGTCCCA	130	CAGGAGGGGG AAGTGTGTCC	130	DMSO	1	1

TABELA E										
OT1 -11	GGGGAGGG GAAGTTTG CTCCTGG	130f	3		GCCCCATTCTTT TTGCAAGTGA	130f	GAGAGCAAGT TTGTTCCCCAG G	130f	DMSO	0 1 2
OT1 -12	GGGGCTGG GGACTTTG CTCCAGG	130f	3		GCCCCCAGCC CCTCTGTTTC	130f	GCTGCTGGTA GGGGAGCTGG	130f	DMSO	1 2 0
OT1 -13	GGGTGGGG GGAGTGGG CTCCAGG	130f	3		CGGCTGCCTT CCCTGAGTCC	130f	GGGTGACGCT TGCCATGAGC	131f	72C Recozer, DMSO a 3%	1 2 0
OT1 -14	GGGTGGCT GGAGTTTG CTGCTGG	131f	3		TGACCCTGGA GTACAAAATG TTCCCA	131f	GCTGAGACAA CCAGCCCAGC T	131f	72C Recozer, DMSO a 3%	2 1 0
OT1 -15	GGGTGGGG GGTGCCTG CTCCAGG	131f	3		TGCCTCCACC CTTAGCCCCCT	131f	GCAGCCGATC CACACTGGGG	131f	DMSO	1 0 2
OT1 -16	GGTTGAGG GGAGTCTG CTCCAGG	131f	3		AACTCAGGAC AACACTGCCT GT	131f	CCCAGGAGCA GGGTACAATG C	131f	DMSO	0 1 2
OT1 -17	GTGTGGGT GGCGTTTG CTCCAGG	132f	3		TCCTCCTTGG AGAGGGGCC	132f	CCTTGAAGG GGCCTTGGTG G	132f	DMSO	0 3 0
OT1 -18	AGGTGGTG GGAGCTTG TTCCTGG	132f	4		CCGAGGGCAT GGGCAATCCT	132f	GGCTGCTGCG AGTTGCCAAC	132f	DMSO	0 1 3
OT1 -19	AGTTTGGG GGAGTTTG CCCCAGG	132f	4		TGCTTTGCAT GGGGTCTCAG ACA	132f	GGGTGCTTG CCCTCTGTGT	132f	DMSO	0 2 2
OT1 -20	ATGTGTGG GGAATTTG CTCCAGG	132f	4		AGTCCTTCT CATTCTCTTC TGCTGT	133f	CACAGAAGGA TGTGTGCAGG TT	133f	DMSO	0 2 2
OT1 -21	CAGTGGGG GGAGCTTT CTCCTGG	133f	4		AGCAGACACA GGTGAATGCT GCT	133f	GGTCAGGTGT GCTGTAGGC A	133f	DMSO	1 1 2
OT1 -22	GAGGGGGA GCAGTTTG CTCCAGG	133f	4		CCTGTGGGGC TCTCAGGTGC	133f	ACTGCCTGCC AAAGTGGGTG T	133f	Sem DMSO TD	1 1 2
OT1 -23	GGAGGAGG GGAGTCTG CTCCAGG	133f	4		AGCTGCACGT GGGAATGAGT	133f	TGCCGGGTAA TAGCTGGCTT	134f	DMSO	0 1 3
OT1 -24	GGAGGGGG GGCTTTTGC TCCAGG	134f	4		CCAGCCTGGG CAACAAAGCG	134f	GGGGGCTTCC AGGTCACAGG	134f	72C Recozer, DMSO a 3%, 6% DMSO	0 3 1
OT1 -25	GGGCAAGG GGAGTTTG CTCCTGG	134f	4		TACCCCACT GCCCCATTGC	134f	ACAGGTCCAT GCTTAGCAGA GGG	134f	DMSO	0 1 3
OT1 -26	GGGTGATT GAAGTTTG CTCCAGG	134f	4	GGGTGATTGAAGT TTGCTCCAGG (SEQ ID NO:2225) GGGTGATTGAAGT TTGCTGCAGG (SEQ ID NO:2226)	ACGGATTAC GACGGAGGTG C	134f	CCGAGTCCGT GGCAGAGAGC	134f	DMSO	0 / 1 2 2
OT1 -27	GGGTGTGG GGTCAATTG CTCCAGG	135f	4		TGTGTTGAA GTAGGGGACA GGT	135f	TGGCCCAATTG GAAGTGATTTT GT	135f	DMSO	3 1 0
OT1 -28	GGTGGGGG TGGGTTTG CTCCTGG	135f	4		TGGGATGGCA GAGTCATCAA CGT	135f	GGCCCAATCG GTAGAGGATG CA	135f	DMSO	0 3 1
OT1 -29	GTGGGGGT AGAGTTTG CTCCAGG	135f	4		ATGGGGCGCT CCAGTCTGTG	135f	TGCACCCACA CAGCCAGCAA	135f	DMSO	0 3 1
OT1 -30	TAGTGGAG GGAGCTTG CTCCTGG	135f	4		GGGGAGGGA GGACCAGGGA A	136f	AATTAGCTGGG CGCGTGGT	136f	72C Recozer, DMSO a 3%	0 1 3
OT1 -31	TGCTCGGG GGAGTTTG CACCAGG	136f	4		ATCCCGTGCA GGAAGTCGCC	136f	CAGGCGGCC CTTGAGGAAT	136f	DMSO	3 1 0
OT1 -32	TGGAGAGG GGAGTTGG CTCCTGG	136f	4		CCCCAACCT TTGCTCAGCG	136f	TGAGGAGAAC ACCACAGGCA GA	136f	DMSO	1 2 1
OT1 -33	TGGTGTGG GGAGTCTG CTCCAGG	136f	4		ATCGACGAGG AGGGGGCCTT	136f	CCCCTCACTCA AGCAGGCC	137f	DMSO	0 3 1
OT1 -34	TTGGGGGG GCAGTTTG CTCCTGG	137f	4		TGCTCAAGGG GCCTGTCCA	137f	CAGGGGCAGT GGCAGGAGTC	137f	No DMSO	1 3 0

TABELA E										
OT1 -35	AAGTAAGG GAAGTTTG CTCCTGG	137 ⁺ 5		TGCCTGGCAC GCAGTAGGTG	137 ⁺	GGGAAGGGGG AACAGGTGCA	137 ⁺ DMSO	0	0	5
OT1 -36	AGAAGAGG GGATTTTGC TCCTGG	137 ⁺ 5		Não otimizado				1	1	3
OT1 -37	ATCTGGGG TGATTTTGC TCCTGG	137 ⁺ 5		ACCTGGGCTT GCCACTAGGG	137 ⁺	GCTGCTCGCA GTTAAGCACCA	138 ⁺ DMSO	1	3	1
OT1 -38	CTCTGCTG GGAGTTTG CTCCTGG	138 ⁺ 5		GTGGCCGGG CTACTGCTAC C	138 ⁺	GGTTCACAA GCTGGGGGCA	138 ⁺ DMSO	3	2	0
OT1 -39	CTGGTGGG GGAGCTTG CTCCAGG	138 ⁺ 5		Não otimizado				1	3	1
OT1 -40	CTTTCGGG GGAGTTTG CGCCGGG	138 ⁺ 5		GCAAGAGGCG GAGGAGACCC	138 ⁺	AGAGTCATCCA TTTCCTGGGG GC	138 ⁺ DMSO	2	3	0
OT1 -41	CTTTGGGG TTAGTTTGC TCCTGG	138 ⁺ 5		GGGGTCAGTG GTGATATCCC CCT	138 ⁺	AGGGAATCCTT TTTCCATTGCT TGTTT	139 ⁺ betaina a 1M, TD	1	4	0
OT1 -42	GCTCTGGG GTAGTTTGC TCCAGG	139 ⁺ 5		AGAGAGGCCA CGTGGAGGGT	139 ⁺	GCCTCCCCTC CTCCTTCCCA	139 ⁺ DMSO	1	3	1
OT1 -43	GTCTCTGG GGAGTTTG CTCCGGG	139 ⁺ 5		GACAGTGCCT TGCGATGCAC	139 ⁺	TCTGACCGGTA TGCCCTGACG	139 ⁺ DMSO	3	2	0
OT1 -44	TCCTGAGG GCAGTTTG CTCCAGG	139 ⁺ 5		TGTGTGAACG CAGCCTGGCT	139 ⁺	TGGTCTAGTAC TTCTCTCAGCC TT	139 ⁺ DMSO	3	1	1
OT1 -45	TCTTTGGGA GAGTTTGCT CCAGG	140 ⁺ 5		GGTTCTCCCT TGGCTCCTGT GA	140 ⁺	CCCACTGCTC CTAGCCCTGC	140 ⁺ DMSO	1	3	1
OT1 -46	ACAAGTGG GGAGTTTG CTCCTGG	140 ⁺ 6		TGAAGTCAAC AATCTAAGCTT CCACCT	140 ⁺	AGCTTTGGTAG TTGGAGTCTTT GAAGG	140 ⁺ DMSO	3	1	2
OT1 -47	ACAAGGTG GAAGTTTG CTCCTGG	140 ⁺ 6		TGATTGGGCT GCAGTTCATG TACA	140 ⁺	GCACAGCCTG CCCTTGGGAA	140 ⁺ DMSO	2	1	3
OT1 -48	ACATAGAA GGAGTTTG CTCCAGG	140 ⁺ 6		TCCATGGGCC CCTCTGAAAG A	141 ⁺	AGCGGCTTCT GCTTCTGCGA	141 ⁺ DMSO	1	0	5
OT1 -49	AGACCCAG GGAGTTTG CTCCCGG	141 ⁺ 6		GCGGTTGGTG GGGTTGATGC	141 ⁺	GAGTTCCTCCT CCCGCCAGT	141 ⁺ DMSO	2	0	4
OT1 -50	AGACCCAG GGAGTTTG CTCCCGG	141 ⁺ 6		AGGCAAGATT TTCCAGTTGT CAAGA	141 ⁺	GCTTTTGCCTG GGACTCCGC	141 ⁺ DMSO	2	0	4
OT1 -51	CACGGAGG GGTGTGTTG CTCCTGG	141 ⁺ 6		GCTGCTGGTC GGGCTCTCTG	141 ⁺	GCTCTGTCCCA CTTCCCCTGG	142 ⁺ Sem TD DMSO	3	1	2
OT1 -52	CAGAGCTT GGAGTTTG CTCCAGG	142 ⁺ 6		GCTGCGAGGC TTCCGTGAGA	142 ⁺	CGCCCTAGA GCTAAGGGGG T	142 ⁺ DMSO	3	2	1
OT1 -53	CTATTGATG GAGTTTGCT CCTGG	142 ⁺ 6		CCAGGAGCCT GAGAGCTGCC	142 ⁺	AGGGCTAGGA CTGCAGTGAG C	142 ⁺ DMSO	1	3	2
OT1 -54	CTTTCTAGG GAGTTTGCT CCTGG	142 ⁺ 6		CTGTGCTCAG CCTGGGTGCT	142 ⁺	GCCTGGGGCT GTGAGTAGTTT	142 ⁺ DMSO	2	3	1
OT1 -55	GCCATGCT GGAGTTTG CTCCAGG	143 ⁺ 6		AGCTCGCGCC AGATCTGTGG	143 ⁺	ACTTGGCAGG CTGAGGCAGG	143 ⁺ 72C Recozer, DMSO a 3%	4	2	0
Alvo 2		143 ⁺			143 ⁺		143 ⁺			
	GACCCCT CCACCCCG CCTCCGG	143 ⁺ 0		AGAGAAGTCG AGGAAGAGAG AG	143 ⁺	CAGCAGAAAG TTCATGGTTTC G	143 ⁺ DMSO			
OT2 -1	GACCCCT CCACCCCG CCCCCGG	143 ⁺ 2		TGGACAGCTG CAGTACTCCC TG	144 ⁺	ACTGATCGATG ATGGCCTATG GGT	144 ⁺ DMSO	0	0	2
OT2 -2	GGGCCCC CCACCCCG CCTCTGG	144 ⁺ 2		CAAGATGTGC ACTTGGGCTA	144 ⁺	GCAGCCTATTG TCTCCTGGT	144 ⁺ DMSO	1	0	1
OT2 -3	AACCCAT CCACCCCG CCTCAGG	144 ⁺ 3		GTCCAGTGCC TGACCTGGC	144 ⁺	AGCATCATGCC TCCAGCTTCA	144 ⁺ DMSO	1	1	1

TABELA E									
OT2 -4	CACCCCT CAACACCG CCTCAGG	144f	3	GCTCCCGATC CTCTGCCACC	144f	GCAGCTCCCA CCACCCCTCAG	145f	DMSO	1 2 0
OT2 -5	CACCCCT CCCTCCG CCTCAGG	145f	3	GGGGACAGG CAGGCAAGGA G	145f	GTGCGTGTCC GTTCAACCCCT	145f	DMSO	1 1 1
OT2 -6	CTACCCCT CCACCCCG CCTCCGG	145f	3	AAGGGGCTGC TGGGTAGGAC	145f	CGTGATTGCA GTTCTTGCA	145f	DMSO	2 1 0
OT2 -7	GACCCGCG CCGCCCGG CCTCTGG	145f	3	GACCCCTCAGG AAGCTGGGAG	145f	CTGCGAGATG CCCCAAATCG	145f	betaina a 1M, TD	1 0 2
OT2 -8	GATCGACT CCACCCCG CCTCTGG	146f	3	CCGCGGCGCT CTGCTAGA	146f	TGCTGGGATTA CAGGCGCGA	146f	DMSO	1 1 1
OT2 -9	GGCCCCAC CCACCCCG CCTCTGG	146f	3	CCAGGTGGTG TCAGCGGAGG	146f	TGCTGGGCC TCTCTGAGTCT	146f	DMSO	0 2 1
OT2 -10	GGCCCGCT CCTCCCGG CCTCCGG	146f	3	CGACTCCACG GCGTCTCAGG	146f	CAGCGCAGTC CAGCCCGATG	146f	betaina a 1M, TD	2 1 0
OT2 -11	GGCCCGCT CCACCCAGG CCTCAGG	146f	3	CTTCCCTCCC CCAGCACCCAC	147f	GCTACAGGTT GCACAGTGAG AGGT	147f	DMSO	1 1 1
OT2 -12	GGCCCGCT CCTCTCG CCTCTGG	147f	3	CCCCGGGGA GTCTGTCTG A	147f	CCCAGCCGTT CCAGGTCTTCC	147f	72C Recozer, DMSO a 3%	1 0 2
OT2 -13	GGCGCCCT CCACCCCTG CCTCGGG	147f	3	GAAGCGCGAA AACC CGGCTC	147f	TCCAGGGTCC TTCTCGGCC	147f	DMSO	1 0 2
OT2 -14	GTCTCCA CCACCCCG CCTCTGG	147f	3	AGGGTGGTCA GGAGGCGCTT	147f	CATGGGGCTC GGACCTCGTC	148f	DMSO	2 0 1
OT2 -15	TACCCCC ACACCCCG CCTCTGG	148f	3	GGGAAGAGG CAGGGGTGTC G	148f	TGCCAGGAAG GAAGCTGGCC	148f	72C Recozer, DMSO a 3%	0 2 1
OT2 -16	AACCCATTC CACCTGTC CTCAGG	148f	4	GAGTGACGAT GAGCCCCGG G	148f	CCCTTAGCTGC AGTCGCCCC	148f	68C Recozer, DMSO a 3%	0 1 3
OT2 -17	ACACCCCG CCACCCCG CCTCAGG	148f	4	CCCATGAGGG GTTTGAGTGC	148f	TGAAGATGGG CAGTTTGGGG	148f	DMSO	0 2 2
OT2 -18	AGCCCCCA CCTCCCGG CCTCGGG	149f	4	CACCTGGGGC ATCTGGGTGG	149f	ACTGGGGTTG GGGAGGGGAT	149f	DMSO	2 0 2
OT2 -19	ATTCCCC CCACCCCG CCTCAGG	149f	4	TCATGATCCC CAAAAGGGCT	149f	CCATTGTGCT GATCTGTGGG T	149f	DMSO	1 0 3
OT2 -20	CCCCACCC CCACCCCG CCTCAGG	149f	4	TGGTGCCAG AATAGTGCC A	149f	AGGAAATGTGT TGTGCCAGGG C	149f	DMSO	1 2 1
OT2 -21	CCCCCCCCA CCACCCCG CCCCGGG	149f	4	GCCTCAGACA ACCCTGCCCC	150f	GCCAAGTGTTA CTCATCAAGAA AGTGG	150f	Sem DMSO TD	2 1 1
OT2 -22	CCCCCCCCC CCCCCCCCG CCTCAGG	150f	4	GCCGGGACAA GACTGAGTTG GG	150f	TCCCGAACTCC CGCAAACG	150f	DMSO	1 2 1
OT2 -23	CGCCCTCC CCACCCCG CCTCCGG	150f	4	TGCTGCAGGT GGTTCCGGAG	150f	CTGGAACCGC ATCCTCCGCA	150f	Sem DMSO TD	1 0 3
OT2 -24	CTCCCCAC CCACCCCG CCTCAGG	150f	4	ACACTGGTCC AGGTCCCGTC T	150f	GGCTGTGCCT TCCGATGGAA	151f	DMSO	2 1 1
OT2 -25	CTCTCCCC CCACCCCG CCTCTGG	151f	4	CTCTCCCCCACC CCCCCTCTGG (SEQ ID NO:2227)	151f	ATCGCGCCCA AAGCACAGGT	151f	DMSO	3 0 2
OT2 -26	GCCTCTCT GCACCCCG CCTCAGG	151f	4	Não otimizado					1 1 2
OT2 -27	GTCACCTC CCACCCCG CCTCTGG	151f	4	CCCTCATGGT GGTCTTACGG CA	151f	AGCCACACATC TTTCTGGTAGG G	151f	DMSO	1 1 2
OT2 -28	TGCCCCCT CCCCCAG CCTCTGG	151f	4	TGCGTCGCTC ATGCTGGGAG	151f	AGGGTGGGGT GTA CTGGCTCA	152f	DMSO	0 3 1

TABELA E									
OT2 -29	TGCCCCCTC CCACCCCG CCTCTGG	152 ^t 4		GAGCTGAGAC GGCACCACTG	152 ^t	TGGCCTTGAAC TCTTGGGCT	152 ^t	betaina a 1M, TD	0 1 3
OT2 -30	TTCCCCCTC CACCCAGC CTCTGG	152 ^t 4		Não otimizado					1 2 1
OT2 -31	TTTCCCTC CTCCCCGC CTCGGG	152 ^t 4		AGTGAGAGTG GCACGAACCA	152 ^t	CAGTAGGTGG TCCCTTCCGC	152 ^t	DMSO	2 1 1
OT2 -32	ACCCCTGCG CCACCCCG CCTCAGG	152 ^t 5		Não otimizado					1 1 3
OT2 -33	AGCCAACC CCACCCCG CCTCTGG	152 ^t 5		GGGAGAACCT TGTCACGCCT	153 ^t	AAGCCGAAAA GCTGGGCAAA	153 ^t	DMSO	0 2 3
OT2 -34	AGGCCCCC ACACCCCG CCTCAGG	153 ^t 5		CTTCCAGTG TGGCCGTCC	153 ^t	ACACAGTCAGA GCTCCGCCG	153 ^t	DMSO	1 1 3
OT2 -35	AGGCCCCC CCGCCCG CCTCAGG	153 ^t 5		Não otimizado					1 0 4
OT2 -36	ATCTGCCA CCACCCCG CCTCCGG	153 ^t 5		CTGAGAGGGG GAGGGGGAG G	153 ^t	TCGACTGGTCT TGTCCTCCA	153 ^t	68C Recozer, DMSO a 3%	3 0 2
OT2 -37	CATCTTCCC CACCCCGC CTCTGG	153 ^t 5		CAGCCTGCTG CATCGAAAA	154 ^t	TGCAGCCAAG AGAAAAAGCCT	154 ^t	betaina a 1M, TD	1 0 4
OT2 -38	CTTTCCTC CACCCAGC CTCTGG	154 ^t 5		TCCCTCTGAC CCGGAACCCA	154 ^t	ACCCGACTTCC TCCCCATTGC	154 ^t	DMSO	2 1 2
OT2 -39	GTGAGGT CCACCCCG CCTCAGG	154 ^t 5		TGGGGGTTGC GTGCTGTCA	154 ^t	GCCAGGAGGA CACCAGGACC	154 ^t	DMSO	4 1 0
OT2 -40	GTGAGGT CCACCCCG CCTCAGG	154 ^t 5		ATCAGGTGCC AGGAGGACAC	154 ^t	GGCCTGAGAG TGGAGAGTGG	155 ^t	DMSO	4 1 0
OT2 -41	TCAGACCT CCACCCCG CCTCAGG	155 ^t 5		Não otimizado					1 4 0
OT2 -42	TGCAACCT CCTCCCG CCTCGGG	155 ^t 5		TGAGCCACAT GAATCAAGGC CTCC	155 ^t	ACCTCTCCAAG TCTCAGTAACT CTCT	155 ^t	DMSO	1 3 1
OT2 -43	ACCACTCT GCACCCCG CCTCTGG	155 ^t 6		GGTCCCTCTG TGCACTGGAA	155 ^t	CTTTGGTGGAC CTGCACAGC	155 ^t	DMSO	2 2 2
OT2 -44	ACTACCCA CCTCCCG CCTCAGG	155 ^t 6		GCGAGGCTGC TGACTTCCCT	155 ^t	GCTGGGACTA CAGACATGTG CCA	156 ^t	DMSO	2 2 2
OT2 -45	ATTTCCTCC CCCCCGC CTCAGG	156 ^t 6	ATTTCCTCCCTCC CC-CCTCAGG (SEQ ID NO:2228)	ATTGCAGGCG TGTCAGGCA	156 ^t	AAATCCTGCAT GGTGATGGGA GT	156 ^t	DMSO	1 1 5
OT2 -46	CCACCATC CCACCCCG CCTCTGG	156 ^t 6		TGCTCTGCCA TTTATGTCCTA TGAAT	156 ^t	ACAGCCTCTC TCCATGACTGA GC	156 ^t	DMSO	1 3 2
OT2 -47	CCCAAGCC CCACCCCG CCTCGGG	156 ^t 6		TCCGCCCAAA CAGGAGGCAG	156 ^t	GCGGTGGGGA AGCCATTGAG	156 ^t	DMSO	2 3 1
OT2 -48	CCGCGCTT CCGCCCG CCTCTGG	157 ^t 6		GGGGGTCTGG CTCACCTGGA	157 ^t	CCTGTGCGGA GAGTGCCTGC	157 ^t	DMSO	3 1 2
OT2 -49	CCTGCCAT GCACCCCG CCTCAGG	157 ^t 6		TCCTGTTCA TTTGCTAGAA CTCTGGA	157 ^t	ACTCCAGATGC AACCAGGGCT	157 ^t	DMSO	3 2 1
OT2 -50	CTGCTCCT CACCCCGC CTCAGG	157 ^t 6		CGTGTGGTGA GCCTGAGTCT	157 ^t	GCTTACCCTGA GAGGCTGCT	157 ^t	DMSO	3 0 3
OT2 -51	TCCTCTTT CACCCCGC CTCAGG	157 ^t 6		AGGCCCTGAT AATTCTATGCTA CCAA	158 ^t	TCAGTGACAAC CTTTGTATTG GGCA	158 ^t	DMSO	0 2 4
OT2 -52	TTGACCCC CCGCCCG CCTCAGG	158 ^t 6		Não otimizado					2 2 2
Alvo 3	GGTGAGTG AGTGTGTG CGTGTGG	158 ^t 0		TCCAGATGGC ACATTGTCAG	158 ^t	AGGGAGCAGG AAAGTGAGGT	158 ^t	DMSO	

TABELA E									
OT3 -1	GGTGAGTG AGTGTGTGT GTGAGG	158f 1	GCAGGCAAGC TGCAAGGGT	158f	CACCGACACA CCCCTCACC	158f	DMSO	0	0 1
OT3 -2	AGTGAGTG AGTGTGTGT GTGGGG	158f 2	GAGGGGGAA GTCACCGACA A	159f	TACCCGGGCC GTCTGTTAGA	159f	DMSO	0	0 2
OT3 -3	AGTGAGTG AGTGTGTG CGTGTGG	159f 2	GACACCCAC ACACTCTCAT GC	159f	TGAATCCCTTC ACCCCAAG	159f	DMSO	1	0 1
OT3 -4	GCTGAGTG AGTGTATG CGTGTGG	159f 2	TCCTTTGAGG TTCATCCCC	159f	CCAATCCAGG ATGATTCCGC	159f	DMSO	1	0 1
OT3 -5	GGTGAGTG AGTGTGTG AGTGAGG	159f 2	CAGGGCCAGG AACACAGGAA	159f	GGGAGGTATG TGCGGGAGTG	160f	DMSO	1	1 0
OT3 -6	GGTGAGTG AGAGTGAGT TGTGTGG	160f 2	TGCAGCCTGA GTGAGCAAGT GT	160f	GCCCAGGTGC TAAGCCCCC	160f	DMSO	1	0 1
OT3 -7	GGTGAGTG AGTGAGTG AGTGAGG	160f 2	TACAGCCTGG GTGATGGAGC	160f	TGTGTCATGGA CTTTCCCATTT T	160f	betaina a 1M, TD	1	1 0
OT3 -8	GGTGAGTG AGTGAGTG AGTGAGG	160f 2	GGCAGGCATT AAACTCATCA GGTCC	160f	TCTCCCCAAG GTATCAGAGA GCT	160f	DMSO	1	1 0
OT3 -9	GGTGAGTG AGTGCGTG CGGGTGG	161f 2	GGGCCTCCCT GCTGGTTCTC	161f	GCTGCCGTCC GAACCCAAGA	161f	DMSO	0	1 1
OT3 -10	GGTGAGTG TGTGTGTG AGTGTGG	161f 2	ACAAACGCAG GTGGACCGAA	161f	ACTCCGAAAAT GCCCCGCAGT	161f	DMSO	1	1 0
OT3 -11	GGTGAGTG TGTGTGTG CATGTGG	161f 2	AGGGGAGGG GACATTGCCT	161f	TTGAGAGGGTT CAGTGGTTGC	161f	DMSO	1	0 1
OT3 -12	GGTGAGTG AGTGTGTGT GTGTGG	161f 2	CTAATGCTTAC GGCTGCGGG	162f	AGCCAACGGC AGATGCAAAT	162f	DMSO	1	0 1
OT3 -13	GGTGAGTG TGTGTGTG CGTGCCG	162f 2	GAGCGAAGTT AACCACCGC	162f	CACACATGCAC ATGCCCTG	162f	68C, DMSO a 3%	2	0 0
OT3 -14	GGTGAGTG TGTGTGTG CGTGTGG	162f 2	GCAATGTGTCT AATGAGGAC AATAGCA	162f	TCCCCCATATC AACACACACA	162f	DMSO	2	0 0
OT3 -15	GGTGAGTG TGTGTGTG CGTGTGG	162f 2	GCCCCCTCCG CCTTTGTGT	162f	TGGGCAAAAG ACATGAAACAG ACA	163f	DMSO	2	0 0
OT3 -16	GGTGAGTG TGTGTGTG CGTGTGG	163f 2	GCCTCAGCTC TGCTCTTAAG CCC	163f	ACGAACAGATC ATTTTTCATGG CTTCC	163f	DMSO	2	0 0
OT3 -17	GTTGAGTG AATGTGTG CGTGAGG	163f 2	CTCCAGAGCC TGCCCTACCA	163f	CCCTCTCCGG AAGTGCCTTG	163f	DMSO	0	1 1
OT3 -18	TGTGGGTG AGTGTGTG CGTGAGG	163f 2	TCTGTACCA CACAGTTACC ACC	163f	GTTGCCTGGG GATGGGGTAT	163f	DMSO	0	1 1
OT3 -19	ACTGTGTG AGTGTGTG CGTGAGG	164f 3	GGGGACCCCTC AAGAGGCACT	164f	GGGCATCAA GGATGGGGAT	164f	DMSO	2	0 1
OT3 -20	AGAGAGTG AGTGTGTG CATGAGG	164f 3	TGTGGAGGGT GGGACCTGGT	164f	ACAGTGAGGT GCGGTCTTTG GG	164f	DMSO	1	0 2
OT3 -21	AGCGAGTG GGTGTGTG CGTGGGG	164f 3	CGGGGTGGCA GTGACGTCAA	164f	GGTGAGTCC AAGAGCCCC	164f	DMSO	0	0 3
OT3 -22	AGGGAGTG ACTGTGTG CGTGTGG	164f 3	AGCTGAGGCA GAGTCCCCGA	165f	GGGAGACAGA GCAGCGCCTC	165f	DMSO	1	1 1
OT3 -23	AGTGAGTG AGTGAGTG AGTGAGG	165f 3	ACCACAGAC CCCACCTCCA	165f	AGGACGACTT GTGCCCCATT A	165f	72C Recozer, DMSO a 3%	1	1 1
OT3 -24	CATGAGTG AGTGTGTG GGTGGGG	165f 3	GGGTGAGGAC GCAGGTGAGA	165f	TCCACCCACC CACCATCCT	165f	72C Recozer, DMSO a 3%	2	0 1
OT3 -25	CGTGAGTG TGTGTATGC GTGTGG	165f 3	ACACTCTGGG CTAGGTGCTG GA	165f	GCCCCCTCAC CACATGATGCT	166f	DMSO	2	0 1

TABELA E									
OT3 -26	GGACTGTG AGTGTGTG CGTGAGG	166° 3	GGGGCCATTG CTGTGCTGCA	166°	TGGGGATCCTT GCTCATGGC	166°	DMSO	3	0 0
OT3 -27	GGTGTGTG CCTGTGTG CGTGTTG	166° 3	ACACACTGGC TCGCATTAC CA	166°	CCTGCACGAG GCCAGGTGTT	166°	DMSO	2	1 0
OT3 -28	GTITCATGA GTGTGTGC GTGGGG	166° 3	TGGGCACGTA GTAACTGCA CCA	166°	CTCGCCGCCG TGACTGTAGG	166°	DMSO	0	3 1
OT3 -29	TGAGTGTG AGTGTGTG CGTGGGG	167° 3	TCAGCTGGTC CTGGGCTTGG	167°	AGAGCACTGG GTAGCAGTCA GT	167°	DMSO	2	1 0
OT3 -30	TGCCAGTG AGTGTGTG CGTGTTG	167° 3	AGACACAGCC AGGGCCTCAG	167°	GGTGGGCGTG TGTGTGTACC	167°	68°C, DMSO a 3%	1	1 1
OT3 -31	TGGGTGTG AGTGTGTG CGTGTTG	167° 3	ACACTCTCAC ACACGCACCA A	167°	GAGAAGTCAG GGCTGGCGGG	167°	72°C Recozer, DMSO a 3%	1	2 0
OT3 -32	TGTATGTGA GTGTGTGC GTGTGG	167° 3	ACTGCCTGCA TTTCCCCGGT	168°	TGGTGAGGGC TTCAGGGAGC	168°	DMSO	1	1 1
OT3 -33	TGTGAGAG AGAGTGTG CGTGTTG	168° 3	GCCAGGTTCA TTGACTGCC	168°	TCCTTCTACAC ATCGGCGGC	168°	DMSO	2	1 0
OT3 -34	TGTGCCTG AGTGTGTG CGTGTTG	168° 3	CGAGGGAGCC GAGTTCGTAA	168°	CTGACCTGGG GCTCTGGTAC	168°	DMSO	1	2 0
OT3 -35	TGTGTGTGT GTGTGTGC GTGTGG	168° 3	TCCTCGGGAA GTCATGGCTT CA	168°	GCACTGAGCA ACCAGGAGCA C	169°	DMSO	2	1 0
OT3 -36	AGCGTGTG AGTGTATG CGTGGGG	169° 4	Não otimizado					1	0 3
OT3 -37	ATTGAGTGT GTGAGTGC GTGGGG	169° 4	TAAACCGTTG CCCCCGCCTC	169°	GCTCCCTGTC CAGGTGAACC	169°	DMSO	2	1 1
OT3 -38	CATGTGTG GGTGTGTG CGTGTTG	169° 4	CCTGCTGAGA CTCCAGGTCC	169°	CTGCGGAGTG GCTGGCTATA	169°	DMSO	2	0 2
OT3 -39	CCCGAGTG TGTGTGTG CGTGTTG	169° 4	CTCGGGGACT GACAAGCCGG	169°	GGAGCAGCTC TTCCAGGGCC	170°	DMSO	3	0 1
OT3 -40	CTGGAGTG AGTGTGTGT GTGTGG	170° 4	CCCCGACCAA AGCAGGAGCA	170°	CTGGCAGCCT CTGGATGGGG	170°	DMSO	1	2 1
OT3 -41	GTITCATGA GTGTGTGC GTGGGG	170° 4	Não otimizado					0	3 1
OT3 -42	TATGTGTGC GTGTGTGC GTGTGG	170° 4	ATTTCAGAGC CCCGGGGAAA	170°	AGGCCGCGGT GTTATGGTTA	170°	DMSO	1	2 1
OT3 -43	TATGTGTGT GTGTGTGC GTGGGG	170° 4	GCCAGTGGCT TAGTGTCTTTG TGT	170°	TGACATATTTT CCTGGGCCAT GGGT	171°	DMSO	2	1 1
OT3 -44	TCTGTGTGT GTGTGTGC GTGGGG	171° 4	TGCCAGAAGA ACATGGGCCA GA	171°	CCATGCTGACA TCATATACTGG GAAGC	171°	DMSO	3	1 0
OT3 -45	TCTGTGTGT GTGTGTGC GTGTGG	171° 4	GCGTGTCTCT GTGTGCGTGC	171°	CCAGGCTGGG CACACAGGTT	171°	DMSO	3	1 0
OT3 -46	TGAGCGTG AGTGTGAG CGTGTTG	171° 4	Não otimizado					2	2 0
OT3 -47	TGTCTTTGA GTGTGTGC GTGTGG	171° 4	TGCCCAGTCC AATATTTACG AGCT	171°	AGGATGAGTTC ATGTCCTTTGT GGGG	172°	DMSO	2	2 0
OT3 -48	TTTGTGTGT GTGTGTGC GTGTGG	172° 4	GGGTGAAAT TTGGTACTGTT AGCTGT	172°	AATGACTCATT CCCTGGGTAT CTCCCA	172°	DMSO	2	2 0
OT3 -49	AAGCGCTG TGTGTGTG CGTGTTG	172° 5	TGCCCCATCA ATCACCTCGG C	172°	CAAGGTCGGC AGGGCAGTGA	172°	DMSO	1	2 2
OT3 -50	AATTCGTGT GTGTGTGC GTGGGG	172° 5	GCCTCCTCTG CCGCTGGTAA	172°	TGAGAGTTCTT GTTGCTCCACA CT	172°	DMSO	1	2 2

TABELA E										
OT3 -51	ATGGTGTG TGTGTGTG CGTGTGG	173	5	Não otimizado				2	2	1
OT3 -52	CACGTGTG TGTGTGTG CGTGTGG	173	5	GCCACCAAAA TAGCCAGCGT	173	ACATGCATCTG TGTGTGCGT	173	DMSO	3	0 2
OT3 -53	GAAATTTGA GTGTGTGC GTGTGG	173	5	ACAGACTGAC CCTTGAAAAAT ACCACT	173	TGTATCTTTCT TGCCAATGGTT TTCCC	173	DMSO	2	1 2
OT3 -54	TAAGTGTGT GTGTGTGC GTGTGG	173	5	AGCCAAATTT CTCAACAGCA GCACT	173	TCCTGGAGAG CAGGCATTTT GT	173	DMSO	3	1 1
OT3 -55	TATATGTGT GTGTGTGC GTGGGG	174	5	ACCTCCTTGT GCTGCCTGGC	174	GGCGGGAAGG TAACCTGGG	174	DMSO	2	1 2
OT3 -56	TATCTGTGT GTGTGTGC GTGTGG	174	5	CACAAAGCTC TACCTTTCCA GTAGTGT	174	TGATCCGATG GTTGTTACAG CT	174	DMSO	3	1 1
OT3 -57	TTTATGTGT GTGTGTGC GTGTGG	174	5	TGTGGGGATT ACCTGCCTGG C	174	ACGCACAAAA TGCCCTGTCA	174	DMSO	2	2 1
OT3 -58	TTTTGTGT GTGTGTGC GTGGGG	174	5	TGAGGCAGAC CAGTCATCCA GC	175	GCCCAGGCAC AGTGTAGGGC	175	DMSO	2	3 0
OT3 -59	AAAAATTGT GTGTGTGC GTGGGG	175	6	ATTAGCTGGG CGTGGCGGAG	175	ACTGCATCTCA TCTCAGGCAG CT	175	DMSO	2	1 3
OT3 -60	ACAATGTG TGTGTGTG CGTGTGG	175	6	TGAAGCAGAA GGAGTGGAGA AGGA	175	TCAGCTTCACA TCTGTTTCAGT TCAGT	175	DMSO	4	0 2
OT3 -61	ATGGGGTG TGTGTGTG CGTGTGG	175	6	TGGTGAGTG TGTGTGTGGT	175	AGAGCAGAAA GAGAGTGCCC A	176	DMSO	1	3 2
OT3 -62	CAAAATTGT GTGTGTGC GTGTGG	176	6	GCCCCTGTAC GTCCTGACAG C	176	TGCACAAGCC ACTTAGCCTCT CT	176	DMSO	3	1 2
OT3 -63	CCCTGGTG TGTGTGTG CGTGTGG	176	6	AGCGCAGGTA AACAGGCCCA	176	TCTCTCGCCCC GTTTCCTTGT	176	DMSO	3	1 2
OT3 -64	TCCGCTTGT GTGTGTGC GTGGGG	176	6	ATGGGTGCCA GGTACCACGC	176	ACAGCAGGAA GGAGCCGCAG	176	DMSO	2	3 1
OT3 -65	TCCTCGTGT GTGTGTGC GTGTGG	177	6	CGGGCGGGT GGACAGATGA G	177	AGGAGGTCTC GAGCCAGGGG	177	DMSO	2	3 1
OT3 -66	TTAAGGTG GGTGTGTG CGTGGGG	177	6	TCAACCTAGT GAACACAGAC CACTGA	177	GTCTATATACA GCCCCACAACC TCATGT	177	DMSO	1	2 3
OT3 -67	TTATATTGT GTGTGTGC GTGGGG	177	6	GCCAGGGCCA GTGGATTGCT	177	TGTCATTCTCT AGTATGTCAGC CGGA	177	DMSO	2	4 0
OT3 -68	TTGAGGAG AGTGTGTG CGTGAGG	177	6	GAGCCCCACC GGTTCAGTCC	178	GCCAGAGCTA CCCACTCGCC	178	DMSO	1	3 2
		178			178		178			
Alvo 4	GAGTCCGA GCAGAAGA AGAAGGG	178	0	GGAGCAGCTG GTCAGAGGGG	178	GGGAAGGGGG ACACTGGGGA	178	DMSO		
OT4 -1	GAGTTAGA GCAGAAGA AGAAAGG	178	2	TCTCTCTTCA ACTCATGACC AGCT	178	ATCTGCACATG TATGTACAGGA GTCAT	179	DMSO	0	1 1
OT4 -2	AAGTCAGA GGAGAAGA AGAAGGG	179	3	AAGACAGAGGAG AAGAAGAAGG (SEQ ID NO:2229)	179	TGGGGAATCT CCAAAGAACC CCC	179	DMSO	2	1 1
OT4 -3	AAGTCCGA GGAGAGGA AGAAAGG	179	3	GATGGCCCCA CTGAGCACGT	179	ACTTCGTAGAG CCTTAAACATG TGCC	179	DMSO	1	0 2
OT4 -4	AAGTCTGA GCACAAGA AGAATGG	179	3	AGGATTAATG TTTAAAGTCAC TGGTGG	179	TCAACAAGGT GCAGATACAG CA	179	betaina a 1M, TD	1	0 2
OT4 -5	ACGTCTGA GCAGAAGA AGAATGG	180	3	TCCAAGCCAC TGGTTTCTCA GTCA	180	TGCTCTGTGGA TCATATTTTGG GGGA	180	DMSO	0	1 2
OT4 -6	GACTCCTA GCAAAAGA AGAATGG	180	3	ACTTTCAGAG CTTGGGGCAG GT	180	CCCACGCTGA AGTGCAATGG C	180	DMSO	1	1 1

TABELA E										
OT4 -7	GAGACTGA GAAGAAGA AGAAAGG	180f 3	CAAAGCATGC CTTTCCAGCG	180f	GGCTCTTCGAT TTGGCACCT	180f	betaina a 1M, TD	1	1	1
OT4 -8	GAGCCGGA GCAGAAGA AGGAGGG	180f 3	Não otimizado					1	0	2
OT4 -9	GAGCCTGA GCAGAAGG AGAAGGG	181f 3	GGAAGCATGC CAGCTCCAGC	181f	AGGAACACAG GCCAGGCTGG	181f	72C Recozer, 6% DMSO	0	0	3
OT4 -10	GAGGCCGA GCAGAAGA AAGACGG	181f 3	CCCTTTAGGC ACCTTCCCA	181f	CCGACCTTCAT CCCTCCTGG	181f	DMSO	0	1	2
OT4 -11	GAGTAAGA GAAGAAGA AGAAGGG	181f 3	TGATTCTGCC TTAGAGTCCC AGGT	181f	TGGGCTCTGT GTCCCTACCCA	181f	DMSO	0	3	0
OT4 -12	GAGTAGGA GGAGAAGA AGAAAGG	181f 3	Não otimizado					2	1	0
OT4 -13	GAGTCCGG GAAGGAGA AGAAAGG	182f 3	AGGCAGGAGA GCAAGCAGGT	182f	ACCCTGACTAC TGACTGACCG CT	182f	DMSO	0	1	2
OT4 -14	GATTCCTAC CAGAAGAA GAATGG	182f 3	CTCCCCATTG CGACCCGAGG	182f	AGAGGCATTG ACTTGGAGCA CCT	182f	DMSO	1	2	0
OT4 -15	GCGACAGA GCAGAAGA AGAAGGG	182f 3	CTGGAGCCCA GCAGGAAGGC	182f	CCTCAGGGAG GGGCGCTGAT	182f	DMSO	1	2	0
OT4 -16	AAATCCAA CCAGAAGA AGAAAGG	182f 4	ACTGTGGCG TTGTCCCCAC	183f	AGGTCCGGTGC AGGGTTTAAG GA	183f	DMSO	1	0	3
OT4 -17	AAGTCTGA GGACAAGA AGAATGG	183f 4	GGCGTCCCT TTTTCCCTTG T	183f	CGTCAACCATC GTCTCGTGA	183f	DMSO	2	0	2
OT4 -18	AAGTTGGA GCAGGAGA AGAAGGG	183f 4	TGCCATCTATA GCAGCCCCCT	183f	GCATCTTGCTA ACCGTACTTCT TCTGA	183f	DMSO	1	0	3
OT4 -19	AATACAGA GCAGAAGA AGAATGG	183f 4	GTGGAGACGC TAAACCTGTG AGGT	183f	GCTCCTGGCC TCTTCCTACAG C	184f	DMSO	1	2	1
OT4 -20	AGGTAATA GCAGAAGA AGAAAGG	184f 4	CCGAATTCT GCTGAGCTTG ATGC	184f	CCAAGTCAATG GGCAACAAGG GA	184f	DMSO	0	2	2
OT4 -21	AGGTGCTA GCAGAAGA AGAAGGG	184f 4	Não otimizado					1	1	2
OT4 -22	AGGTGGGA GCAGAAGA AGAAGGG	184f 4	TGCCCCAAG ACCTTTCTCC	184f	ATGGCAGGCA GAGGAGGAAG	184f	DMSO	2	0	2
OT4 -23	CAAACGGA GCAGAAGA AGAAAGG	184f 4	GGGTGGGGC CATTGTGGGT T	184f	CTGGGGCCAG GGTTTCTGCC	185f	DMSO	3	0	1
OT4 -24	CACTCTGA GGAGAAGA AGAAAGG	185f 4	TGGAGAACAT GAGAGGCTTG CAA	185f	TCCTTCTGTAG GCAATGGGAA CAA	185f	DMSO	3	0	1
OT4 -25	CAGTCAATG GCAGAAGA AGAAAGG	185f 4	GCCACATGGT AGAAGTCGGC	185f	GGCAGATTTCC CCCATGCTG	185f	betaina a 1M, TD	1	2	1
OT4 -26	CCGTCCCA GCAGTAGA AGAATGG	185f 4	TGTACACCCC AAGTCCTCCC	185f	AAGGGGAGTG TGCAAGCCTC	185f	DMSO	3	1	0
OT4 -27	GTCTGCGA TCAGAAGAA GAAAGG	186f 4	AGGTCTGGCT AGAGATGCAG CA	186f	AGTCCAACACT CAGGTGAGAC CCT	186f	DMSO	3	1	0
OT4 -28	TAATCCAAT CAGAAGAA GAAAGG	186f 4	CCAAGAGGAC CCAGCTGTTG GA	186f	GGGTATGGAA TTCTGGATTAG CAGAGC	186f	DMSO	0	2	2
OT4 -29	TATACGGA GCAGAAGA AGAATGG	186f 4	ACCATCTCTTC ATTGATGAGT CCCA	186f	ACACTGTGAGT ATGCTTGGCGT	186f	DMSO	2	2	0
OT4 -30	ACTTCCCTG CAGAAGAA GAAAGG	186f 5	GGCTGCGGG GAGATGAGCT C	187f	TCGGATGCTTT TCCACAGGGC T	187f	DMSO	2	2	1
OT4 -31	AGGACTGG GCAGAAGA AGAAGGG	187f 5	TCTTCCAGGA GGGCAGCTCC	187f	CCAATCCTGAG CTCTACAAGG CT	187f	DMSO	1	0	4

TABELA E									
OT4 -32	AGGTTGGA GAAGAAGA AGAAGGG	187: 5	GAGCTGCACT GGATGGCACT	187: 5	TGCTGGTTAAG GGGTGTTTTG GA	187: 5	DMSO	1	1 3
OT4 -33	AGTTTCAGA GCAGGAGA AGAATGG	187: 5	TCTGGGAAGG TGAGGAGGCC A	187: 5	TGGGGGACAA TGGAAAAGCAA TGA	188: 5	DMSO	0	2 3
OT4 -34	ATGACACA GCAGAAGA AGAAGGG	188: 5	CTTGCTCCCA GCCTGACCCC	188: 5	AGCCCTTGCC ATGCAGGACC	188: 5	DMSO	3	1 1
OT4 -35	ATGACAGA GAAGAAGA AGAAAGG	188: 5	GGGATTTTAT CTGTTGGGTG CGAA	188: 5	AACCACAGATG TACCCTCAAAG CT	188: 5	DMSO	2	2 1
OT4 -36	CCGCCCT GCAGAAGA AGAACGG	188: 5	ACCCATCAGG ACCGCAGCAC	188: 5	TCTGGAACCTG GGAGGCGGA	188: 5	72C Recozer, DMSO a 3%	3	1 1
OT4 -37	GCAGGAGA GCAGAAGA AGAAAGG	189: 5	CGTCCCTCAC AGCCAGCCTC	189: 5	CCTCCTTGGG CCTGGGGTTC	189: 5	DMSO	1	3 1
OT4 -38	GTTCAAGA GCAGAAGA AGAATGG	189: 5	CCCTCTGCAA GGTGGAGTCT CC	189: 5	AGATGTTCTGT CCCCAGGCCT	189: 5	DMSO	1	3 1
OT4 -39	GTTTGAAG CAGAAGAA GAAAGG	189: 5	GGCTTCCACT GCTGAAGGCC T	189: 5	TGCCGCTCCA CATACCCTCC	189: 5	DMSO	2	1 2
OT4 -40	TATGGCAA GCAGAAGA AGAAAGG	189: 5	AGCATTGCCT GTCGGGTGAT GT	190: 5	AGCACCTATTG GACACTGGTTC TCT	190: 5	DMSO	1	3 1
OT4 -41	TGGTGGGA TCAGAAGAA GAAAGG	190: 5	TCTAGAGCAG GGGCACAATG C	190: 5	TGGAGATGGA GCCTGGTGGG A	190: 5	DMSO	2	2 1
OT4 -42	ACCCACGG GCAGAAGA AGAAGGG	190: 6	GGTCTCAGAA AATGGAGAGA AAGCACG	190: 6	CCCACAGAAA CCTGGGCCCT	190: 6	DMSO	1	2 3
OT4 -43	ACTCCTGAT CAGAAGAA GAAGGG	190: 6	GGTTGCTGAT ACCAAAACGT TTGCCT	190: 6	TGGGTCTCTC CACCTCTGCA	191: 6	DMSO	0	3 3
OT4 -44	ACTGATGA GCAGAAGA AGAAAGG	191: 6	ACTCTCCTTAA GTAATGATAT GGCTGT	191: 6	CAGAATCTTGC TCTGTTGCCCA	191: 6	DMSO	0	4 2
OT4 -45	ATTTTAGTG CAGAAGAA GAAAGG	191: 6	Não otimizado					2	2 2
OT4 -46	ATTTTAGTG CAGAAGAA GAAAGG	191: 6	Não otimizado					2	2 2
OT4 -47	CCATGGCA GCAGAAGA AGAAGGG	191: 6	CAATGCCTGC AGTCTCAGG A	191: 6	TCCCAAGAGAA AACTCTGTCTC GACA	191: 6	DMSO	4	1 1
OT4 -48	CCATTACA GCAGAAGA AGAAGGG	191: 6	GCATTGGCTG CCCAGGAAAA	192: 6	TGGCTGTGCT GGGCTGTGTT	192: 6	DMSO	2	2 2
OT4 -49	CGAGGCGG GCAGAAGA AGAAAGG	192: 6	CCACAAGCCT CAGCCTACCC G	192: 6	ACAGGTGCCA AAACACTGCCT	192: 6	DMSO	2	1 3
OT4 -50	TCATTGCAG CAGAAGAA GAAAGG	192: 6	TCATTGCAGCAGA AGAAGAAAGG TCATTGTAGCAGA AGAAGAAAGG (SEQ ID NO:2230)	192: 6	GCCTCTTGCA AATGAGACTC CTTTT	192: 6	CGATCAGTCC CCTGGCGTCC	192: 6	DMSO 2 / 1 2 / 3 2
OT4 -51	TCTCCAGG GCAGAAGA AGAAAGG	192: 6	TCCCAGAAATC TGCTCCGCA	192: 6	AGGGGTTTCC AGGCACATGG G	193: 6	DMSO	0	4 2
Alvo 5	GTCATCTTA GTCATTACC TGAGG	193: 0	TCCTAAAAATC AGTTTTGAGAT TTACTTCC	193: 0	AAAGTGTTAGC CAACATACAGA AGTCAGGA	193: 0	DMSO		
OT5 -1	GGTATCTAA GTCATTACC TGTGG	193: 3	GGTATCTAAGTCA TTACCTGTGG (SEQ ID NO:2231) GGTATCTAAGTCA ATACCTGTGG (SEQ ID NO:2232)	193: 3	ACATCTGGGG AAAGCAAAAG TCAACA	193: 3	TGTCTGAGTAT CTAGGCTAAAA GTGGT	193: 3	DMSO 1 / 2 1 1
OT5 -2	GTAATATTA GTCATTACC GGTGG	194: 3	ACGATCTTGC TTCATTTCCT GTACA	194: 3	AGTGCTTTGTG AACTGAAAAGC AAACA	194: 3	DMSO	0	3 0

TABELA E									
OT5 -3	GTAATCTGA GTCATTCC TGGGG	194f 3	GCACCTTGGT GCTGCTAAAT GCC	194f	GGGCAACTGA ACAGGCATGA ATGG	194f DMSO	1	2	0
OT5 -4	GTCATCCTA GTCATTAC TGGGG	194f 3	AACTGTCTCG CATCCCCGCC	194f	GGTGCACCTG GATCCACCCA	194f DMSO	1	1	1
OT5 -5	GTCATCCTA GTGCTTACC TGAGG	194f 3	Não otimizado	195f		195f	1	1	1
OT5 -6	GTCATCTGA GGCATTAA CTGGGG	195f 3	CATCACCTC CACCAGGCC	195f	ACCACTGCTG CAGGCTCCAG	195f 72C Recozer, DMSO a 3%	0	3	0
OT5 -7	AATATGTTA GTCATTACC TGAGG	195f 4	Não otimizado				2	0	2
OT5 -8	ATAACGTA GTCATTACC TGGGG	195f 4	CCTGACCCGT GGTCCCCGAC	195f	TGGTGCGTGG TGTGTGTGGT	195f 72C Recozer, DMSO a 3%	1	2	1
OT5 -9	ATCATCATC GTCATTATC TGGGG	195f 4	TGGGAACATT GGAGAAGTTT CCTGA	196f	CCATGTGACTA CTGGGCTGCC C	196f DMSO	1	1	2
OT5 -10	ATCATTTTA CTCATTACT TGTGG	196f 4	AGCCTTGGCA AGCAACTCCC T	196f	GGTTCTCTCTC TCAGAAAAGAA AGAGG	196f DMSO	1	0	3
OT5 -11	ATCATTTTA GTCATCTCC TGTGG	196f 4	GGCAGCGGAC TTCAGAGCCA	196f	GCCAGAGGCT CTCAGCAGTG C	196f DMSO	1	0	3
OT5 -12	CACAGCTTA GTCATCACC TGGGG	196f 4	CCAGCCTGGT CAATATGGCA	196f	ACTGTGCCA GCCCATATT	197f DMSO	2	1	1
OT5 -13	CCCAGCTTA GTCATTAGC TGTGG	197f 4	ATGCCAACAC TCGAGGGGCC	197f	CGGGTTGTGG CACCGGTTA	197f DMSO	2	1	1
OT5 -14	CTCACCTTT GTCATTCC TGAGG	197f 4	TTGCTCTAGT GGGGAGGGG G	197f	AGAGTTCAGG CATGAAAAGAA GCAACA	197f DMSO	3	0	1
OT5 -15	CTCATTTTA TTCATTGCC TGGGG	197f 4	AGCTGAAGAT AGCAGTGT AAGCCT	197f	TGCAATTTGAG GGGCTCTCTC A	197f DMSO	1	1	2
OT5 -16	CTCTCCTTA GTCACTACC TGAGG	198f 4	AGTCACTGGA GTAAGCCTGC CT	198f	TGCCAGCCAA AAGTTGTTAGT GTGT	198f DMSO	2	0	2
OT5 -17	CTTATCTCT GTCATTACC TGGGG	198f 4	GGGTCTCCCT CAGTGCCCTG	198f	TGTGTGGTAG GGAGCAAAAC GACA	198f DMSO	2	0	2
OT5 -18	GACAGCTC CGTCATTAC CTGGGG	198f 4	TGGGGGCTGT TAAGAGGCAC A	198f	TGACCACACAC ACCCCCACG	198f DMSO	1	2	1
OT5 -19	GCCACCTC AGTCATTAG CTGGGG	198f 4	TCAAAACAGA TTGACCAAGG CCAAAT	199f	TGTGTTTTTAA GCTGCACCCC AGG	199f DMSO	1	0	3
OT5 -20	GGAATCTTA CTCATTACT TGGGG	199f 4	TCTGGCACCA GGAAGTATTG TACA	199f	GCACGCAGCT GACTCCCAGA	199f DMSO	1	2	1
OT5 -21	GTGGCCTC AGTCATTAC CTGCGG	199f 4	Não otimizado				1	0	3
OT5 -22	GTGTTTTTA GTGATTACC TGAGG	199f 4	AGCATCTGTG ATACCCTACC TGTCT	199f	ACCAGGGCTG CCACAGAGTC	199f DMSO	1	0	3
OT5 -23	TACATCTTA GTCCTCAC CTGTGG	199f 4	TAGTCTTGTG CCCAGGCTG	200f	CTCGGCCCT GAGAGTTCAT	200f DMSO	1	2	1
OT5 -24	TCCATCTCA CTCATTACC TGAGG	200f 4	TCCATCTCACTCA TTACCTGAGG (SEQ ID NO:2233) TCCATCTCACTCA TTACCTGATG (SEQ ID NO:2234)	200f	GAGCAGCAGC AAAGCCACCG	200f DMSO	1	1	2
OT5 -25	TTCATCCTA GTCAACAC CTGGGG	200f 4	GCCTGGAGAG CAAGCCTGGG	200f	AGCCGAGACA ATCTGCCCGG	200f DMSO	1	1	2
OT5 -26	TTTATATTA GTGATTACC TGTGG	200f 4	TTTATATTAGTGAT TACCTGCGG (SEQ ID NO:2235)	200f	AGTGAAACAA ACAAGCAGCA GTCTGA	201f Sem TD DMSO	1	2	1

TABELA E									
OT5 -27	AACGTGTA AGTCATTAC CTGAGG	201 ¹	5	AGGCTCAGAG AGGTAAGCAA TGGA	201 ²	TGAGTAGACA GAAATGTTACC GGTGTT	201 ³	DMSO	3 0 2
OT5 -28	AAGATCAC AGTCATTAC CTGGGG	201 ⁴	5	TCAGAGATGT TAAAGCCTTG GTGGG	201 ⁵	AGTGAACCAA GGGAATGGGG GA	201 ⁶	DMSO	3 0 2
OT5 -29	AGAATATTA GTCCTTACC TGGGG	201 ⁷	5	TGTGCTTTCT GGGGTAGTGG CA	201 ⁸	CACCTCAGCC CTGTAGTCCTG G	201 ⁹	DMSO	0 4 1
OT5 -30	AGCAGATT AGTGATTAC CTGGGG	202 ¹	5	CCATTGGGTG ACTGAATGCA CA	202 ²	GCCACTGTCC CCAGCCTATT	202 ³	betaina a 1M, TD	1 3 1
OT5 -31	AGTAGCTTA GTGATTACC TGGGG	202 ⁴	5	ACCAAGAAAG TGAAAAGGAA ACCC	202 ⁵	TGAGATGGCAT ACGATTACCC A	202 ⁶	DMSO	1 2 2
OT5 -32	CACGGCTT ACTCATTAC CTGGGG	202 ⁷	5	AGGGTGGGGA CTGAAAGGAG CT	202 ⁸	TGGCATCACTC AGAGATTGGAA CACA	202 ⁹	DMSO	3 1 1
OT5 -33	CATATGTTA GGCATTAC CTGGGG	202 ¹⁰	5	ACCAGTGCTG TGTGACCTTG GA	203 ¹	TCCTATGGGA GGGGAGGCTT CT	203 ²	DMSO	3 1 1
OT5 -34	CATTCTTA GTCATTCC TGAGG	203 ³	5	CCAGGTGTGG TGGTTCATGA C	203 ⁴	GCATACGGCA GTAGAATGAG CC	203 ⁵	68C, DMSO a 3%	4 0 1
OT5 -35	TGCAGCTA ACTCATTAC CTGCGG	203 ⁶	5	CAGGCGCTGG GTTCTTAGCC T	203 ⁷	CCTTCCTGGG CCCCATGGTG	203 ⁸	DMSO	2 3 0
OT5 -36	TTGCTTTTA GTTATTACC TGGGG	203 ⁹	5	TGGGGTCCAA GATGTCCCTT	203 ¹⁰	TGAAACTGCTT GATGAGGTGT GGA	204 ¹	DMSO	1 2 2
OT5 -37	AACTTGAA AGTCATTAC CTGTGG	204 ²	6	GCTGGGCTTG GTGGTATATG C	204 ³	ACTTGCAAAGC TGATAACTGAC TGA	204 ⁴	DMSO	5 0 1
OT5 -38	AAGGTCAC AGTCATTAC CTGGGG	204 ⁵	6	AGTTGGTGTG ACTGACAATG GGA	204 ⁶	CGCAGCGCAC GAGTTCATCA	204 ⁷	DMSO	3 0 3
OT5 -39	AATGTCTTC ATCATTACC TGAGG	204 ⁸	6	AGAGGAGGCA CAATTC AAC CCT	204 ⁹	GGCTGGGGAG GCCTCACAAT	204 ¹⁰	DMSO	1 1 4
OT5 -40	AGATGCTT GGTCATTAC CTGTGG	205 ¹	6	GGGAAAGTTT GGGAAAGTCA GCA	205 ²	AGGACAAGCT ACCCACACCC	205 ³	DMSO	1 3 2
OT5 -41	AGTAGATTA GTTATTACC TGGGG	205 ⁴	6	TGGTGCATCA AAGGGTTGCT TCT	205 ⁵	TCATTCCAGCA CGCCGGGAG	205 ⁶	DMSO	0 3 3
OT5 -42	AGTAGGTTA GTAATTACC TGGGG	205 ⁷	6	CCCAGGCTGC CCATCACACT	205 ⁸	TGGAGTAAGTA TACCTTGGGG ACCT	205 ⁹	DMSO	1 3 2
OT5 -43	CAAATGAG AGTCATTAC CTGAGG	205 ¹⁰	6	TCAGTGCCCC TGGGTCTCA	206 ¹	TGTGCAAATAC CTAGCACGGT GC	206 ²	DMSO	4 2 0
OT5 -44	CATGTCTGA ATCATTACC TGAGG	206 ³	6	AGCACTCCCT TTTGAATTTG GTGCT	206 ⁴	ACTGAAAGTCCA GCCTCTCCAT TTCA	206 ⁵	DMSO	2 1 3
OT5 -45	CCTGACTT GGTCATTAC CTGTGG	206 ⁶	6	GAAACCGGTC CCTGGTGCCA	206 ⁷	GGGGAGTAGA GGGTAGTGTT GCC	206 ⁸	DMSO	2 0 4
OT5 -46	CGTGCAATTA GTCATTACC TGAGG	206 ⁹	6	TTGCGGGTCC CTGTGGAGTC	206 ¹⁰	AGGTGCCGTG TTGTGCCCAA	207 ¹	DMSO	1 2 3
Alvo 6	GGAATCCC TTCTGCAGC ACCTGG	207 ²	0	GCCCTACATC TGCTCTCCCT CCA	207 ³	GGGCCGGGAA AGAGTTGCTG	207 ⁴	DMSO	
OT6 -1	GGAACCCC GTCTGCAG CACCAGG	207 ⁵	2	TTGGAGTGTG GCCCCGGTTG	207 ⁶	ACCTCTCTTTC TCTGCCTCACT GT	207 ⁷	DMSO	0 1 1
OT6 -2	GGAACACC TTCTGCAGC TCCAGG	207 ⁸	3	CACACCATGC TGATCCAGGC	207 ⁹	GCAGTACGGA AGCACGAAGC	207 ¹⁰	DMSO	1 1 1
OT6 -3	GGAAGCTC TGCTGCAG CACCTGG	208 ¹	3	CTCCAGGGCT CGGTGTCCAC	208 ²	CTGGGCTCTG CTGGTTCCCC	208 ³	DMSO	0 2 1
OT6 -4	GGAATATCT TCTGCAGC CCCAGG	208 ⁴	3	CTGTGGTAGC CGTGGCCAGG	208 ⁵	CCCCATACCAC CTCTCCGGGA	208 ⁶	DMSO	0 2 1

TABELA E										
OT6 -5	GGAATCAC TTTACAGC ACCAAGG	208f	3		GGTGGCGGG ACTTGAATGA G	208f	CCAGCGTGT CCAAGGGAT	208f	betaina a 1M, TD	0 1 2
OT6 -6	GGAATCCC CTCTCCAG CCCCTGG	208f	3	GGAATCCCCTCTC CAGCCCCTGG (SEQ ID NO:2236) GGAATCCCCTCTC CAGCCTCTGG(SEQ ID NO:2237)	CCAGAGGTGG GGCCCTGTGA	209f	TTTCCCACTC AGTTCTGCAG GA	209f	DMSO	1 1 1 / 2
OT6 -7	GGAATCTCT TCTTCAGCA TCTGG	209f	3	GGAATCTCTCTCT TGGCATCTGG(SEQ ID NO:2238)	TGTGACTGGT TGTCTGTCTT CCT	209f	GCAGTGTCTT TGGTGATGGG CA	209f	betaina a 1M, TD	0 1 5
OT6 -8	GGAATGCT TCTGCAGC GCCAGG	209f	3		CTGGCCAAGG GGTGAGTGGG	209f	TGGGACCCCA GCAGCCAATG	209f	DMSO	1 0 2
OT6 -9	GGACTCCC CTCTGCAG CAGCTGG	209f	3		ACGGTGTGCT GGCTGCTCTT	209f	ACAGTGTGTA CCGTGCTGGG	210f	DMSO	1 1 1
OT6 -10	GGAGTCCC TCCTACAG CACCAGG	210f	3		TGGTTTGGGC CTCAGGGATG G	210f	TGCCTCCACA AAAAATGTCTAC CT	210f	DMSO	0 0 3
OT6 -11	GGAGTCCC TCCTACAG CACCAGG	210f	3		TGGTTTGGGC CTCAGGGATG G	210f	ACCCCTTATCC CAGAACCCAT GA	210f	DMSO	0 0 3
OT6 -12	GGCATCCA TTCTGCAGC CCCTGG	210f	3		TCCAAGTCAG CGATGAGGGC T	210f	TGGGAGCTGT TCCTTTTTGGC CA	210f	DMSO	0 3 0
OT6 -13	GGCTTCCC TTCTGCAGC CCCAGG	211f	3		CACCCCTCTC AGCTTCCCAA	211f	GCTAGAGGGT CTGCTGCCTT	211f	DMSO	1 2 0
OT6 -14	TGAATCCCA TCTCCAGCA CCAGG	211f	3		AGACCCCTTG GCCAAGCACA	211f	CTTGCTCTCAC CCCGCCTCC	211f	DMSO	2 1 0
OT6 -15	AAAATACCT TCTGCAGTA CCAGG	211f	4		ACATGTGGGA GGCGGACAGA	211f	TCTCACTTTGC TGTTACCGATG TCG	211f	DMSO	0 1 3
OT6 -16	AAAATCCCT TCTTCAACA CCTGG	211f	4		GACGACTGT GCCTGGGACA	212f	AGTGCCCA GTGTTGTAAC GCT	212f	72C Recozer, DMSO a 3%	0 1 3
OT6 -17	ACACTCCCT CCTGCAGC ACCTGG	212f	4		GGAGAGTCTA GCGCCAGGTC	212f	CAGCGTGGCC CGTGGGAATA	212f	DMSO	1 1 2
OT6 -18	ACCATCCCT CCTGCAGC ACCAGG	212f	4		GCTGAAGTGC TCTGGGGTGC T	212f	ACCCCACTGT GGATGAATTG GTACC	212f	DMSO	1 1 2
OT6 -19	AGAGGCC CTCTGCAG CACCAGG	212f	4		TCGGGGTGCA CATGGCCATC	212f	TTGCCTCGCA GGGGAAGCAG	213f	DMSO	0 1 3
OT6 -20	AGGATCCC TTGTGCAG CTCCTGG	213f	4		CTCGTGGGAG GCCAACACCT	213f	AGCCACCAAC ACATACCAGG CT	213f	DMSO	2 0 2
OT6 -21	CCACTCCTT TCTGCAGC ACCCGG	213f	4		GCATGCCCTT AATCCCGGCT	213f	AGGATTTTACA GTGATGGGGC T	213f	DMSO	2 1 1
OT6 -22	GAAGGCC TTCAGCAG CACCTGG	213f	4		CGCCAGCCCA CAAAGTGCAT	213f	GCAAATTTCTG CACCTACTCTA GGCCT	213f	DMSO	1 1 2
OT6 -23	GATATCCCT TCTGTATCA CCTGG	214f	4		AGCTCACAAG AATTGGAGGT AACAGT	214f	GCAGTCACCC TTCACTGCCTG T	214f	DMSO	1 1 2
OT6 -24	GGGTCCGC TTCTGCAGC ACCTGG	214f	4		AAACTGGGCT GGGCTTCCGG	214f	GGGGCTAAGG CATTGTGAGAC CC	214f	DMSO	2 0 2
OT6 -25	GTCTCCCTT TCTGCAGC ACCAGG	214f	4		GCAGGTAGGC AGTCTGGGGC	214f	TCTCTGCCTC AGCCTCCCA	214f	betaina a 1M, TD	1 2 1
OT6 -26	GTCTCCCTT TCTGCAGC ACCAGG	214f	4		GCAGGTAGGC AGTCTGGGGC	215f	TCTCTGCCTC AGCCTCCCA	215f	betaina a 1M, TD	1 2 1
OT6 -27	GTCTCCCTT TCTGCAGC ACCAGG	215f	4		GCAGGTAGGC AGTCTGGGGC	215f	TCTCTGCCTC AGCCTCCCA	215f	betaina a 1M, TD	1 2 1
OT6 -28	TCATTCCCG TCTGCAGC ACCCGG	215f	4		GCTCTGGGGT AGAAGGAGGC	215f	GGCCTGTCAA CCAACCAACC	215f	DMSO	2 2 0

TABELA E										
OT6 -29	TGCACCCC TCCTGCAG CACCAGG	215f	4		TGACATGTTG TGTGCTGGGC	215f	AAATCCTGCAG CCTCCCCCTT	216f	DMSO	0 2 2
OT6 -30	TGCATACC CTCTGCAG CACCAGG	216f	4		TCCTGGTGAG ATCGTCCACA GGA	216f	TCCTCCCCACT CAGCCTCCC	216f	DMSO	0 3 1
OT6 -31	TGCATGGC TTCTGCAGC ACCAGG	216f	4		TCCTAATCCAA GTCCTTTGTTC AGACA	216f	AGGGACCAGC CACTACCCCTC A	216f	DMSO	2 2 0
OT6 -32	AATATTCCC TCTGCAGC ACCAGG	216f	5		GGGACACCAG TTCCTTCCAT	216f	GGGGGAGATT GGAGTTCCCC	216f	DMSO	1 0 4
OT6 -33	ACCATTCT TCTGCAGC ACCTGG	217f	5		ACACCACTAT CAAGGCAGAG TAGGT	217f	TCTGCCTGGG GTGCTTTCCC	217f	DMSO	1 1 3
OT6 -34	AGCTCCCA TTCTGCAGC ACCCGG	217f	5		CTGGGAGCGG AGGGAAGTGC	217f	GCCCCGACAG ATGAGGCCTC	217f	DMSO	1 2 2
OT6 -35	CAGATTCT GCTGCAGC ACCGGG	217f	5	CAGATTACTGCTG CAGCACCGGG (SEQ ID NO:2239)	CGGGTCTCGG AATGCCTCCA	217f	ACCCAGGAATT GCCACCCCC	217f	DMSO	1 2 3
OT6 -36	CCAAGAGC TTCTGCAGC ACCTGG	217f	5		TTGCTGTGGT CCCGTGGTG	218f	GCAGACACTA GAGCCCGCCC	218f	DMSO	3 2 0
OT6 -37	CCCAGCCC TGCTGCAG CACCCGG	218f	5		GGTGTGGTGA CAGGTGGGGT	218f	ACCTGCGTCTC TGTGCTGCA	218f	DMSO	2 3 0
OT6 -38	CCCTCCCC TCCTGCAG CACCGGG	218f	5		CTCCCAGGAC AGTGCTCGGC	218f	CCTGGCCCCA TGCTGCCTG	218f	DMSO	2 2 1
OT6 -39	CTACTGACT TCTGCAGC ACCTGG	218f	5		TGCGTAGGTT TTGCCTCTGT GA	218f	AGGGAATGAT GTTTTCCACCC CCT	219f	DMSO	2 3 0
OT6 -40	CTCCTCCCT CCTGCAGC ACCTGG	219f	5		CTCCGACGCC ACCGTTGGTA	219f	TGCATTGACGT ACGATGGCTC A	219f	DMSO	1 3 1
OT6 -41	TCTGTCCCT CCTGCAGC ACCTGG	219f	5		ACCTGCAGCA TGAACCTCTG CA	219f	ACCTGAGCAA CATGACTCACC TGG	219f	DMSO	2 1 2
OT6 -42	ACACAAAC TTCTGCAGC ACCTGG	219f	6	ACACAAACTTCTG CAGCACCTGG ACACAAACTTCTG CAGCACGTGG(SEQ ID NO:2240)	TCTCCAGTTTC TTGCTCTCAT GG	219f	ACCATTGGTGA ACCCAGTCA	219f	betaina a 1M, TD	3 / 2 3 1
OT6 -43	ACTGTCATT TCTGCAGC ACCTGG	220f	6		TGGGGTGGTG GTCTTGAATC CA	220f	TCAGCTATAAC CTGGGACTTGT GCT	220f	DMSO	2 1 3
OT6 -44	ACTTTATCT TCTGCAGC ACCTGG	220f	6		AGCAGCCAGT CCAGTGTCTC G	220f	CCCTTTTCATCG AGAACCCAG GG	220f	DMSO	3 1 2
OT6 -45	ATCCTTTCT TCTGCAGC ACCTGG	220f	6		TGGACGCTGC TGGGAGGAGA	220f	GAGGTCTCGG GCTGCTCGTG	220f	DMSO	0 3 3
OT6 -46	CACCACCG TTCTGCAGC ACCAGG	220f	6		AGGTTTGAC TCTGTTGCCT GG	221f	TGGGGTGATT GGTTGCCAGG T	221f	DMSO	3 2 1
OT6 -47	CATGTGGC TTCTGCAGC ACCTGG	221f	6		TCTTCCTTTGC CAGGCAGCAC A	221f	TGCAGGAATA GCAGGTATGA GGAGT	221f	DMSO	4 0 2
OT6 -48	CAITTTCTT TCTGCAGC ACCTGG	221f	6		GGACGCCTAC TGCTTGGACC	221f	GCCCTGGCAG CCCATGGTAC	221f	DMSO	3 0 3
OT6 -49	CTCTGTCCT TCTGCAGC ACCTGG	221f	6		AGGCAGTCAT CGCCTTGCTA	221f	GGTCCACCTT CCCCTACAA	222f	DMSO	2 3 1
OT6 -50	CTGTACCCT CCTGCAGC ACCAGG	222f	6		Não otimizado					3 1 2
OT6 -51	TTGAGGCC GTCTGCAG CACCGGG	222f	6		CCCCAGCCCC CACCAGTTTC	222f	CAGCCCAGGC CACAGCTTCA	222f	DMSO	1 4 1

Sequenciamento de Sanger para quantificar frequências de mutações indel

[0142] Produtos de PCR purificados usados para ensaio de T7EI foram ligados em um vector Zero Blunt TOPO (Life Technologies) e transformados em células bacterianas Top 10 quimicamente competentes. Os DNAs de plasmídeo foram isolados e sequenciados pelo Massachusetts General Hospital (MGH) DNA Automation Core, usando um iniciador direto M13 (5'-GTAAAACGACGGCCAG-3') (SEQ ID NO: 1059).

Ensaio de digestão de restrição para quantificação de alterações específicas induzidas por HDR com ssODNs

[0143] As reações de PCR de sítios sobre o alvo específicos foram realizadas usando DNA polimerase Phusion High-Fidelity (New England Biolabs). Os *loci* de *VEGF* e *EMX1* foram amplificados usando um programa de PCR Touchdown ((98 °C, 10 s; 72-62 °C, 1 °C/ciclo, 15 s; 72 °C, 30 s) x 10 ciclos, (98 °C, 10 s; 62 °C, 15 s; 72 °C, 30 s) x 25 ciclos), com DMSO a 3%. Os iniciadores usados para estas reações de PCR estão listados na **Tabela E**. Produtos de PCR foram purificados por esferas Ampure XP (Agencourt) de acordo com as instruções do fabricante. Para detecção do sítio de restrição *Bam*HI codificado pelo modelo doador ssODN, 200 ng de produtos de PCR purificados foram digeridos com *Bam*HI a 37 °C durante 45 minutos. Os produtos digeridos foram purificados por esferas Ampure XP (Agencourt), diluídos em 0,1x 20 µl de tampão EB e analisados e quantificados usando um sistema de eletroforese capilar QIAXCEL.

Geração de biblioteca TruSeq e análise dos dados de sequenciamento

[0144] Iniciadores específicos para um sítio foram concebidos para flanquear sobre o alvo e sítios fora do alvo potenciais e verificados para produzir produtos de PCR ~ 300 pb a 400 bps de comprimento. DNAs genômicos das amostras em duplicata reunidas descritas acima foram usados como modelos para PCR. Todos os produtos de PCR foram

purificados por esferas Ampure XP (Agencourt) segundo as instruções do fabricante. Os produtos de PCR purificados foram quantificados por um sistema de eletroforese capilar QIAxCEL. Os produtos de PCR para cada *locus* foram amplificados a partir de cada uma das amostras reunidas em duplicata (descritos acima), as quais são purificados, quantificados e, então, reunidos em conjunto em quantidades iguais para sequenciamento profundo. Amplicons reunidos foram ligados com adaptadores de Illumina TruSeq duplamente indexados, conforme descrito anteriormente (Fisher *et al.*, 2011). As bibliotecas foram purificadas e passadas em um sistema de eletroforese capilar QIAxCEL para verificar a alteração no tamanho do adaptador após ligação. As bibliotecas do adaptador ligado foram quantificadas por qPCR e depois sequenciadas usando leituras de extremidade pareada de 250 pb Illumina MiSeq realizada pelo Dana-Farber Cancer Institute Molecular Biology Core Facilities. Foram analisados entre 75.000 e 1.270.000 (média de ~ 422 mil) leituras para cada amostra. As leituras TruSeq foram analisadas quanto às taxas de mutagênese indel, conforme descrito anteriormente (Sander *et al.*, 2013). As proporções de especificidade foram calculadas como a proporção de mutagênese observada em um *locus* sobre o alvo para aquela de um determinado *locus* fora do alvo, conforme determinado por sequenciamento de profundidade. Vezes de aprimoramento na especificidade com tru-RGNs para sítios fora do alvo individuais foram calculadas como a proporção de especificidade observada com tru-gRNAs para a proporção de especificidade para o mesmo alvo com o gRNA de comprimento completo correspondente. Conforme mencionado no texto, para alguns dos sítios fora do alvo, nenhuma mutação indel foi detectada com tru-gRNAs. Nestes casos, foi usada uma calculadora de Poisson para determinar, com uma confiança de 95%, que este limite máximo do número real de sequências com mutação seriam um número

de três. Em seguida, usamos este limite máximo para estimar as vezes de aprimoramento mínimo de especificidade para estes sítios fora do alvo.

Exemplo 2a. gRNAs truncados podem comandar eficientemente edição de genoma mediada por Cas9 em células humanas

[0145] Para testar a hipótese de que gRNAs truncados em sua extremidade 5' podem funcionar tão eficientemente quanto suas contrapartes de comprimento total, uma série de gRNAs progressivamente mais curtos foram inicialmente construídos conforme descrito acima para um único sítio alvo no gene repórter *EGFP*, com a seguinte sequência: 5'-**GGCGAG**GGCGATGCCACCTAcGG-3' (SEQ ID NO: 2241). Este sítio de *EGFP* em particular foi escolhido porque era possível fazer gRNAs ao mesmo com 15, 17, 19, e 20 nts de complementaridade os quais têm, cada um, um G em sua extremidade 5' (necessário para expressão eficiente a partir do promotor U6 usado nestes experimentos). Usando um ensaio de repórter com base em células humanas no qual a frequência de indels induzidas por RGN pode ser quantificada por meio de avaliação de ruptura de um único gene de proteína fluorescente integrado e intensificada constitutivamente expresso (*EGFP*) (**Exemplo 1** e Fu *et al.*, 2013; Reyon *et al.*, 2012) (Figura 2B), as habilidades destes gRNAs de comprimento variável de induzir a indels mediadas por Cas9 no sítio alvo foram medidas.

[0146] Como observado acima, gRNAs tendo comprimentos mais longos de complementaridade (21, 23, e 25 nts) exibem diminuição da atividade em relação ao gRNA convencional de comprimento completo contendo 20 nts de sequência complementar (**Figura 2H**), um resultado que corresponde àqueles recentemente reportados por outros (Ran *et al.*, Cell 2013). No entanto, gRNAs trazendo 17 ou 19 nt de complementaridade ao alvo mostraram atividades comparáveis a ou superiores ao gRNA de comprimento completo, enquanto que um gRNA

mais curto tendo apenas 15 nts de complementaridade não demonstrou atividade significativa (**Figura 2H**).

[0147] Para testar a generalidade destes resultados iniciais, gRNAs de comprimento completo e gRNAs correspondentes trazendo 18, 17 e/ou 16 nts de complementaridade para quatro sítios de genes repórter *EGFP* adicionais (sítios #1, #2, #3 e #4 de *EGFP*; **Figura 3A**) foram ensaiados. Em todos os quatro sítios alvo, gRNAs trazendo 17 e/ou 18 nt de complementaridade funcionaram de forma tão eficiente quanto (ou, em um caso, de forma mais eficiente do que) seus gRNAs de comprimento completo correspondentes ao induzir à ruptura de expressão de EGFP mediada por Cas9 (**Figura 3A**). No entanto, gRNAs com apenas 16 nts de complementaridade mostraram atividades significativamente diminuídas ou não detectáveis nos dois sítios nos quais eles poderiam ser feitas (**Figura 3A**). Para cada um dos diferentes sítios testados, nós transfectadas as mesmas quantidades de um plasmídeo de expressão de Cas9 e gRNA de comprimento completo encurtado. Experimentos de controle em que quantidades variadas de plasmídeos de expressão de Cas9 e gRNA truncado transfectados para sítios #1, #2 e #3 de EGFP sugeriram que gRNAs encurtados funcionam de forma equivalente às suas contrapartes de comprimento total (**Figuras 3E (inferior) e 3F (inferior)**) e que, portanto, poderiam usar as mesmas quantidades de plasmídeos ao fazer comparações a qualquer sítio alvo. Tomados em conjunto, estes resultados fornecem evidências de que gRNAs encurtados tendo 17 ou 18 nts de complementaridade geralmente podem funcionar de forma eficiente quanto gRNAs de comprimento completo e, daqui em diante, os gRNAs truncados com estes comprimentos de complementaridade são ditos como "tru-gRNAs" e RGNs usando estes tru -gRNAs como "tru-RGNs".

[0148] Em seguida, foi testado se tru-RGNs poderiam induzir eficientemente a indels sobre genes endógenos cromatinizados alvo.

tru-gRNAs foram construídos para sete sítios em três genes endógenos humanos (*VEGFA*, *EMX1* e *CTLA*), incluindo quatro sítios que tinham sido anteriormente objetivados com gRNAs de comprimento completo em três genes humanos endógenos: sítio 1 de *VEGFA*, sítio 3 de *VEGFA*, *EMX1* e *CTLA* (**Exemplo 1** e Fu *et al.*, 2013; Hsu *et al.*, 2013; Pattanayak *et al.*, 2013) (**Figura 3B**). (Não foi possível testar um tru-gRNA para o sítio 2 de *VEGFA* do Exemplo 1, porque esta sequência alvo não tem G em qualquer posição 17 ou 18 da região de complementaridade necessário para expressão gRNA a partir de um promotor U6). Usando um ensaio de genotipagem de T7 endonuclease I (T7EI) bem estabelecido (Reyon *et al.*, 2012) conforme descrito acima, as frequências de mutação indel mediadas por Cas9 induzidas por cada um dos vários gRNAs em seus respectivos sítios alvo foram quantificadas em células humanas U2OS.EGFP. Para todos os cinco dos sete sítios, tru-RGNs induziram robustamente a mutações indel com eficiências comparáveis àquelas mediadas por RGNs convencionais correspondentes (**Figura 3B**). Para os dois sítios sobre os quais tru-RGNs mostraram atividades mais baixas do que suas contrapartes de comprimento completo, notamos que as taxas absolutas de mutagênese ainda eram elevadas (médias de 13,3% e 16,6%) em níveis que seriam úteis para a maioria das aplicações. Sequenciamento de Sanger para três destes sítios alvo (sítios 1 e 3 de *VEGFA* e *EMX1*) confirmou que indels induzidas por tru-RGNs se originam no sítio de clivagem esperado e que estas mutações são essencialmente indistinguíveis daquelas que foram induzidas com RGNs convencionais (**Figura 3C e Figuras 7A-D**).

[0149] Nós também descobrimos que tru-gRNAs tendo um G 5' não pareado e uma região de complementaridade de 18 nt poderia induzir a indels mediado por Cas9 eficientemente, enquanto aqueles tendo um G 5' não pareado e uma região de complementaridade 17 nt

mostraram atividades menores ou indetectáveis comparado com gRNAs de comprimento completo correspondentes (**Figura 7e**), consistente com nossas descobertas de que um mínimo de 17 nts de complementaridade é necessário para atividade de RGN eficiente.

[0150] Para avaliar ainda mais as capacidades de edição de genoma de tru-RGNs, suas capacidades de induzir alterações precisas de sequência via HDR com modelos doadores ssODN foram testadas. Estudos anteriores demonstraram que as rupturas induzidas por Cas9 podem estimular a introdução de sequência a partir de um doador ssODN homólogo para um *locus* endógeno em células humanas (Cong *et al.*, 2013; Mali *et al.*, 2013c; Ran *et al.*, 2013; Yang *et al.*, 2013). Portanto, as capacidades RGNs de comprimento completo e tru-gRNAs correspondentes objetivados ao sítio 1 de *VEGFA* e ao sítio 1 de *EMX1* de introduzir um sítio de restrição *Bam*HI codificado em ssODNs homólogos endógenos foram comparadas para estes genes. Em ambos os sítios, introdução de um sítio *Bam*HI mediada por tru-RGNs com eficiências comparáveis àquelas observadas com RGNs convencionais que trazem suas contrapartes de gRNA de comprimento completo (**Figura 3D**). Tomados em conjunto, estes dados demonstram que tru-RGNs podem funcionar tão eficientemente quanto RGNs convencionais para comandar tanto indels quanto eventos de edição de genoma mediada por HDR precisos em células humanas.

Exemplo 2b. tru-RGNs exibem sensibilidades aumentadas a desemparelhamentos na interface gRNA/DNA

[0151] Tendo estabelecido que tru-RGNs pode funcionar de forma eficiente para induzir alterações de edição de genoma sobre o alvo, foi testado se estas nucleases mostrariam maior sensibilidade a desemparelhamentos na interface gRNA/DNA. Para avaliar isto, uma série sistemática de variantes foi construída para os tru-gRNAs que foram previamente testados nos sítios #1, #2 e #3 de *EGFP* (**Figura 3A**

acima). Os gRNAs variantes trazem substituições de Watson-Crick individuais em cada posição dentro da região de complementaridade (com exceção do G 5' necessário para expressão a partir do promotor U6) (**Figura 5A**). O ensaio de ruptura de EGFP com base em células humanas foi usado para avaliar as capacidades relativas destes tru-gRNAs variante e um conjunto análogo de gRNAs variantes de comprimento completo correspondente feito nos mesmos três sítios, conforme descrito no Exemplo 1 para indels mediados por Cas9 diretos. Os resultados mostram que, para todos os três sítios alvo de *EGFP*, tru-RGNs geralmente mostraram maiores sensibilidades aos desemparelhamentos individuais do que RGNs convencionais trazendo gRNAs de comprimento completo correspondentes (comparar painéis inferior e superior da **Figura 5A**). A magnitude de sensibilidade variava de acordo com o sítio, com as maiores diferenças observadas para os sítios #2 e #3, cujos tru-gRNAs traziam 17 nts de complementaridade.

[0152] Incentivado pelo aumento da sensibilidade de tru-RGNs a desemparelhamentos de um único nucleotídeo, em seguida procuramos examinar os efeitos do desemparelhamento sistematicamente em duas posições adjacentes na interface gRNA-DNA. Portanto, fizemos variantes dos tru-gRNAs objetivadas aos sítios alvo #1, #2 e #3 de *EGFP*, cada um trazendo substituições por transversões de Watson-Crick em duas posições de nucleotídeos adjacentes (**Figura 5B**). Conforme avaliado pelo ensaio de ruptura de EGFP, os efeitos de desemparelhamentos duplos adjacentes sobre a atividade de RGN foram novamente substancialmente maior para tru-gRNAs do que para as variantes análogas feitas no Exemplo 1 para gRNAs de comprimento completo correspondentes objetivadas a todos os três sítios alvo de EGFP (comparar painéis inferior com superior na **Figura 5B**). Estes efeitos parecem ser dependentes do sítio com quase todos os tru-gRNAs com desemparelhamentos duplos para os sítios #2 e #3 de

EGFP não mostram um aumento das atividades de ruptura de *EGFP* em relação a um gRNA de controle carecendo de uma região de complementaridade e com apenas três dos tru-gRNA variantes com desemparelhamentos para o sítio #1 de *EGFP* que mostram quaisquer atividades residuais (**Figura 5B**). Além disso, embora as mutações duplas mostrassem, em geral, maiores efeitos sobre a extremidade 5' com gRNAs de comprimento total, este efeito não foi observado com tru-gRNAs. Tomados em conjunto, nossos dados sugerem que tru-gRNAs exibem uma maior sensibilidade do que gRNAs de comprimento completo a desemparelhamentos por transversões de Watson-Crick individuais e duplas na interface gRNA-DNA.

Exemplo 2c. tru-RGNs objetivados a genes endógenos mostram especificidades aprimoradas em células humanas

[0153] Os próximos experimentos foram realizados para determinar se tru-RGNs podem mostrar efeitos reduzidos fora do alvo genômico em células humanas em relação às RGNs convencionais trazendo contrapartes de gRNA de comprimento completo. Foram examinados tru-gRNAs de comprimento completo e tru-gRNAs objetivados ao sítio 1 de *VEGFA*, sítio 3 de *VEGFA* e sítio 1 de *EMX1* porque estudos anteriores (vide **Exemplo 1** e Fu *et al.*, 2013 (descrito na **Figura 3B** acima); Hsu *et al.*, 2013) tinham definido 13 sítios de fora do alvo *bona fide* para os gRNAs- de comprimento completo objetivados a estes sítios. (Não fomos capazes de testar um tru-gRNA para o sítio 2 de *VEGFA* a partir de nosso estudo original⁶, porque esta sequência alvo não tem G em uma ou outra das posições 17 ou 18 da região de complementaridade necessária para expressão eficiente de gRNA a partir de um promotor U6). Surpreendentemente, descobrimos que tru-RGNs mostraram atividade de mutagênese substancialmente reduzida em células U2OS. *EGFP* humanas em relação às RGNs convencionais correspondentes em todos os 13 destes sítios fora do alvo *bona fide*,

conforme julgado pelo ensaio de T7EI (**Tabela 3A**); para 11 dos 13 sítios fora do alvo, a frequência de mutação com tru-RGNs caiu abaixo do limite de detecção confiável do ensaio de T7EI (2-5%) (**Tabela 3A**). Foram observados resultados similares quando estes pares correspondentes de RGNs convencionais e tru-RGNs foram testados com os mesmos 13 sítios fora do alvo em outra linhagem de células humana (células FT-HEK293) (**Tabela 3A**).

[0154] Para quantificar a magnitude de aprimoramento de especificidade observado com tru-RGNs, nós medimos as frequências de mutação fora do alvo usando sequenciamento de elevado rendimento, o qual fornece um método mais sensível para detecção e quantificação de mutações de baixa frequência do que o ensaio de T7EI. Nós avaliamos um subconjunto de 12 dos 13 sítios de fora do alvo *bona fide* para o qual tínhamos observado taxas de mutação reduzidas com tru-gRNAs por meio do ensaio de T7EI (por razões técnicas, não fomos capazes de amplificar o amplicon mais curto necessário para um dos sítios) e também analisamos um sítio fora do alvo adicional para o sítio 1 de *EMX1* que tinha sido identificado por outro grupo⁷ (**Figura 6A**). Para todos os 13 sítios fora do alvo testados, tru-RGNs mostraram frequências absolutas de mutagênese substancialmente diminuídas em relação às RGNs convencionais pareadas (**Figura 6A e Tabela 3B**) e obtivemos aprimoramentos na especificidade de tanto quanto ~ 5000 vezes ou mais em relação às suas RGNs homólogas convencionais (**Figura 6B**). Para dois sítios fora do alvo (OT1-4 e OT1-11), foi difícil quantificar as proporções sobre o alvo para fora do alvo para tru-RGNs porque o número absoluto e a frequência de mutações indel induzidas por tru-RGNs caiu para o nível de base ou próximo do nível de base. Assim, a proporção de taxas sobre o alvo para as taxas fora do alvo seria calculada como sendo infinita nestes casos. Para resolver isso, nós identificamos, em vez disso, a frequência de indels provável máxima

com um nível de confiança de 95% para esses sítios e, então, usamos esta estimativa conservada para calcular a magnitude provável mínima de aprimoramento de especificidade para tru-RGNs em relação às RGNs convencionais para estes sítios fora do alvo. Estes cálculos sugerem que tru-RGNs produzem aprimoramentos de ~ 10.000 vezes ou mais nesses sítios (**Figura 6B**).

[0155] Para explorar ainda mais a especificidade de tru-RGNs, nós analisámos sua capacidade de induzir a mutações fora do alvo em sítios adicionais estreitamente relacionados no genoma humano. Para os tru-gRNAs ao sítio 1 de *VEGFA* e *EMX1*, os quais possuem cada um 18 nts de complementaridade ao sítio alvo, nós identificamos computacionalmente todos os sítios adicionais no genoma humano com desemparelhamentos em uma ou duas posições dentro da região de complementaridade (ainda não examinados acima na **Tabela 3A**) e um subconjunto de todos os sítios com desemparelhamentos em três posições que favoreciam desemparelhamentos na extremidade 5' do sítio, conforme descrito no Exemplo 1. Para o tru-gRNA para o sítio 3 de *VEGFA*, o qual possui 17 nt de complementaridade ao sítio alvo, nós identificamos todos os sítios com desemparelhamentos em uma posição e um subconjunto de todos os sítios com desemparelhamentos em duas posições que favoreciam desemparelhamentos de extremidade 5' (ainda não examinados na **Tabela 3A**). Esta análise computacional produziu um total de 30, 30 e 34 sítios adicionais fora do alvo potenciais para as tru-RGNs objetivadas ao sítio 1 de *VEGFA*, o sítio 3 de *VEFGA* e o sítio de *EMX1*, respectivamente, os quais, em seguida, nós avaliamos quanto a mutações usando o ensaio de T7EI em células U2OS.EGFP e células HEK293 humanas nas quais as RGNs tinha sido expressas.

[0156] Surpreendentemente, os três tru-RGNs ao sítio 1 de *VEGFA*, sítio 3 de *VEFGA* e *EMX1* não induziram a mutações indel

detectáveis mediadas por Cas9 em 93 dos 94 sítios fora do alvo potenciais em células U2OS.EGFP humanas examinadas ou em qualquer um dos 94 sítios fora do alvo potenciais em células HEK293 humanas (**Tabela 3C**). Para o sítio no qual foram observadas mutações fora do alvo, se foi examinado a RGN padrão com um gRNA de comprimento total objetivado ao sítio 1 de *VEGFA* também poderia acarretar mutação neste mesmo sítio fora do alvo; ela induziu a mutações detectáveis, embora com uma frequência ligeiramente menor (**Figura 6C**). A falta de aprimoramento observada com encurtamento do gRNA neste sítio fora do alvo pode ser compreendida através ao comparar as sequências de 20 e 18 nt para os gRNAs de comprimento completo e tru-gRNAs, a qual mostra que as duas bases adicionais no alvo de 20 nt de comprimento completo têm ambas desemparelhamentos (**Figura 6C**). Em resumo, com base nesse exame de 94 sítios fora do alvo potenciais adicionais, encurtamento do gRNA não parece induzir a novas mutações fora do alvo de alta frequência.

[0157] Sequenciamento profundo de um subconjunto dos 30 sítios fora do alvo potenciais mais estreitamente alinhados a partir deste conjunto de 94 sítios (ou seja, aqueles com um ou dois desemparelhamentos) mostrou taxas indetectáveis ou muito baixas de mutações indel (**Tabela 3D**) comparáveis àsquelas observadas em outros sítios fora do alvo anteriormente identificados (**Tabela 3B**). Concluímos que tru-RGNs geralmente parecem induzir a níveis muito baixos ou indetectáveis de mutações em sítios que diferem em um ou dois desemparelhamentos a partir dos sítios sobre o alvo. Isto contrasta com RGNs convencionais para as quais foi relativamente fácil encontrar mutações fora do alvo de alta frequência em sítios que diferem em até cinco desemparelhamentos (vide **Exemplo 1**).

TABELA 3A									
Frequências de mutação sobre e fora do alvo de tru-RGNs correspondentes e RGNs convencionais objetivadas a genes endógenos em células U2OS.EGFP e HEK293 humanas									
Alvo ID	20mer Alvo	SEQ ID NO:	Frequência de mutação indel (%) \pm s.e.m.		Alvo truncado	SEQ ID NO:	Frequência de mutação indel (%) \pm s.e.m.		Gene
			U2OS.EGFP	HEK293			U2OS.EGFP	HEK293	
T1	GGGTGGGGGAGTTTGCTCCiGG	2242.	23,69 \pm 1,99	6,98 \pm 1,33	GTGGGGGAGTTTGCTCCiGG	2243.	23,93 \pm 4,37	8,34 \pm 0,01	VEGFA
OT1-3	GGATGGAGGGAGTTTGCTCCiGG	2244.	17,25 \pm 2,97	7,26 \pm 0,62	ATGGAGGGAGTTTGCTCCiGG	2245.	N.D.	N.D.	IGDCC3
OT1-4	GGGAGGGTGGAGTTTGCTCCiGG	2246.	6,23 \pm 0,20	2,66 \pm 0,30	GAGGGTGGAGTTTGCTCCiGG	2247.	N.D.	N.D.	LOC116437
OT1-6	CGGCGGAGGGAGTTTGCTCCiGG	2248.	3,73 \pm 0,23	1,41 \pm 0,07	GCGGAGGGAGTTTGCTCCiGG	2249.	N.D.	N.D.	CACNA2D
OT1-11	GGGCGAGGGGAGTTTGCTCCiGG	2250.	10,4 \pm 0,7	3,61 \pm 0,02	GCGAGGGGAGTTTGCTCCiGG	2251.	N.D.	N.D.	
T3	GGTGAGTGAGTGTGCGTGIGG	2252.	54,08 \pm 1,02	22,97 \pm 0,17	GAGTGAGTGAGTGTGCGTGIGG	2253.	50,49 \pm 1,25	20,05 \pm 0,01	VEGFA
OT3-1	GGTGAGTGAGTGTGTGTGaGG	2254.	6,16 \pm 0,98	6,02 \pm 0,11	GAGTGAGTGAGTGTGTGTGaGG	2255.	N.D.	N.D.	(abParts)
OT3-2	AGTGAGTGAGTGTGTGTGG	2256.	19,64 \pm 1,06	11,29 \pm 0,27	GAGTGAGTGAGTGTGTGTGG	2257.	5,52 \pm 0,25	3,41 \pm 0,07	MAX
OT3-4	GCTGAGTGAGTGTATGCGTGIGG	2258.	7,95 \pm 0,11	4,50 \pm 0,02	GAGTGAGTGTATGCGTGIGG	2259.	1,69 \pm 0,26	1,27 \pm 0,10	
OT3-9	GGTGAGTGAGTGCGTGGIGG	2260.	N.D.	1,09 \pm 0,17	GAGTGAGTGCGTGGIGG	2261.	N.D.	N.D.	TPCN2
OT3-17	GTTGAGTGAATGTGTGCGTGaGG	2262.	1,85 \pm 0,08	N.D.	GAGTGAATGTGTGCGTGaGG	2263.	N.D.	N.D.	SLIT1
OT3-18	IGTGCGTGAGTGTGTGCGTGaGG	2264.	6,16 \pm 0,56	6,27 \pm 0,09	GCGTGAGTGTGTGCGTGaGG	2265.	N.D.	N.D.	COMDA
OT3-20	AGAGAGTGAGTGTGTGCAATGaGG	2266.	10,47 \pm 1,08	4,38 \pm 0,58	GAGTGAGTGTGTGCAATGaGG	2267.	N.D.	N.D.	
T4	GAGTCCGAGCAGAAGAAGaGG	2268.	41,56 \pm 0,20	12,65 \pm 0,31	GTCCGAGCAGAAGAAGaGG	2269.	43,01 \pm 0,87	17,25 \pm 0,64	EMX1
OT4-1	GAGTTAGCAGAGAAGAAGaGG	2270.	19,26 \pm 0,73	4,14 \pm 0,66	GTTAGCAGAGAAGAAGaGG	2271.	N.D.	N.D.	HCN1
OT-4_Hsu31	GAGTCTAGCAGAGAAGAAGaGG	2272.	4,37 \pm 0,58	N.D.	GTC TAGCAGAGAAGAAGaGG	2273.	N.D.	N.D.	MFAP1

Frequências de mutação foram medidas pelo ensaio de T7EI. Médias de medições em duplicata são mostradas com erros da média. *Sítio fora do alvo OT4_53 é o mesmo conforme o alvo 3 de EMX1 OT31 de Hsu *et al.*, 2013.

Tabela 3B									
Números de leituras de sequenciamento de tipo selvagem (WT) mutação indel a partir de experimentos de sequenciamento profundo									
Sítio	Controle			tru-RGN			RGN convencional		
	Indel	WT	Freq.	Indel	WT	Freq.	Indel	WT	Freq.
Sítio 1 de VEGFA	45	140169	0,03%	122858	242127	33,66%	150652	410479	26,85%
OT1-3	0	132152	0,00%	1595	205878	0,77%	50973	144895	26,02%
OT1-4	0	133508	0,00%	0	223881	0,00%	22385	240873	8,50%
OT1-6	3	213642	0,00%	339	393124	0,09%	24332	424458	5,21%
OT1-11	1	930894	0,00%	0	274779	0,00%	43738	212212	17,09%
Sítio 3 de VEGFA	5	212571	0,00%	303913	292413	50,96%	183626	174740	51,24%

OT3-2	1169	162545	0,71%	9415	277616	3,28%	26545	222482	10,66%
OT3-4	7	383006	0,00%	15551	113567 3	1,35%	42699	546203	7,25%
OT3-9	73	145367	0,05%	113	227874	0,05%	1923	168293	1,13%
OT3-17	8	460498	0,00%	31	127127 6	0,00%	16760	675708	2,42%
OT3-18	7	373571	0,00%	284	127598 2	0,02%	72354	599030	10,78%
OT3-20	5	140848	0,00%	593	325162	0,18%	30486	202733	13,07%
Sítio 1 de EMX1	1	158838	0,00%	49104	102805	32,32%	128307	307584	29,44%
OT4-1	10	169476	0,01%	13	234039	0,01%	47426	125683	27,40%
OT4-52	2	75156	0,00%	10	231090	0,00%	429	340201	0,13%
OT4-53	0	234069	0,00%	6	367811	0,00%	17421	351667	4,72%

Freq. = frequência de mutações indel = número de sequências indel/número de sequências de tipo selvagem. gRNA de controle = gRNA carecendo de uma região de complementaridade

Tabela 3C					
Frequências de mutação indel em sítios fora do alvo potenciais of tru-RGNs					
alvoed to endogenous genes in human cells					
Alvo ID	Sítio Alvo + PAM	SEQ ID NO:	Número de desemparelhamentos	Frequência de mutação Indel (%)	
				± s.e.m.	
Sítio 1 de VEGFA	GTGGGGGGAGTTTGCTCCtGG	2274.	0 (sobre o alvo)	23,93 ± 4,37	8,34 ± 0,01
	GTGGGGGGAGTTTGCCCaGG	2275.	1	Não detectado	Não detectado
	GTGGGGGGtGTTTGCTCCcGG	2276.	1	Não detectado	Não detectado
	GTGGGtGGAGTTTGCTACtGG	2277.	2	Não detectado	Não detectado
	GTGGGGGGAGCtTTCTCCtGG	2278.	2	Não detectado	Não detectado
	GTGGGtGGCgTTTGCTCCaGG	2279.	2	Não detectado	Não detectado
	GTGGAGGGAGCtTGCTCCtGG	2280.	2	6,88 ± 0,19	Não detectado
	GTGGGtGGAGTTTGCTACaGG	2281.	2	Não detectado	Não detectado
	GtGGGGGGAGTTTGCTCCtGG	2282.	2	Não detectado	Não detectado
	GTGtGGGGAAITTTGCTCCaGG	2283.	2	Não detectado	Não detectado
	CtGCTGGGAGTTTGCTCCtGG	2284.	3	Não detectado	Não detectado
	ITtGGGAGAGTTTGCTCCaGG	2285.	3	Não detectado	Não detectado
	CtGAGGGCAGTTTGCTCCaGG	2286.	3	Não detectado	Não detectado
	GTAAGGGAGAGTTTGCTCCtGG	2287.	3	Não detectado	Não detectado
	GtGGGtAGAGTTTGCTCCaGG	2288.	3	Não detectado	Não detectado
	GtGtGGGGACtTTTGCTCCaGG	2289.	3	Não detectado	Não detectado
	GtGGGAGCAGTTTGCTCCaGG	2290.	3	Não detectado	Não detectado
	ITGGGGtTAGTTTGCTCCtGG	2291.	3	Não detectado	Não detectado
	ITGAGGGGAGTCTGCTCCaGG	2292.	3	Não detectado	Não detectado
	CtGGGGtGATTTTGCTCCtGG	2293.	3	Não detectado	Não detectado
	GAGAGGGGAGTTGCTCCtGG	2294.	3	Não detectado	Não detectado
	ITtGGGGGAGTTTGCTCCaGG	2295.	3	Não detectado	Não detectado
	ITCtGGGGAGTTTGCTCCgGG	2296.	3	Não detectado	Não detectado
	CtCtGGGGAGTTTGCTCCaGG	2297.	3	Não detectado	Não detectado
	GTGtGGGAGTCTGCTCCaGG	2298.	3	Não detectado	Não detectado
	GAGGGGGCAGTTTGCTCCaGG	2299.	3	Não detectado	Não detectado
	GAGGGGAGAGTTTGCTCCaGG	2300.	3	Não detectado	Não detectado
	GTGGCtGGAGTTTGCTCtGG	2301.	3	Não detectado	Não detectado
	GTcGGGGGAGTGGCTCCaGG	2302.	3	Não detectado	Não detectado
	GAGGGGGGAGTGTGtTCCgGG	2303.	3	Não detectado	Não detectado
	GTGGtGGGAGCtTGCTCCtGG	2304.	3	Não detectado	Não detectado
	GTGGGGGtGcCTGCTCCaGG	2305.	3	Não detectado	Não detectado

Tabela 3C					
Frequências de mutação indel em sítios fora do alvo potenciais of tru-RGNs					
alvoed to endogenous genes in human cells					
Alvo ID	Sítio Alvo + PAM	SEQ ID NO:	Número de desemparelhamentos	Frequência de mutação Indel (%)	
				± s.e.m.	
				Células U2OS.EGFP	Células HEK293
Sítio 3 de <i>VEGFA</i>	GAGTGAGTGTGTGCGTGTGG	2306.	0 (sobre o alvo)	50,49 ± 1,25	20,05 ± 0,01
	CAGTGAGTGTGTGCGTGTGG	2307.	1	Não detectado	Não detectado
	G ^U TGTGAGTGTGTGCGTGgGG	2308.	1	Não detectado	Não detectado
	G ^U TGTGAGTGTGTGCGTGaGG	2309.	1	Não detectado	Não detectado
	G ^U TGTGAGTGTGTGCGTGTGG	2310.	1	Não detectado	Não detectado
	GAGTG ^T GTGTGTGCGTGTGG	2311.	1	Não detectado	Não detectado
	GAGTG ^G GTGTGTGCGTGgGG	2312.	1	Não detectado	Não detectado
	GAGTGA ^C TGTGTGCGTGTGG	2313.	1	Não detectado	Não detectado
	GAGTGAGTGTGTG ^C GTGgGG	2314.	1	Não detectado	Não detectado
	GAGTGAGTGTGTG ^T GTGTGG	2315.	1	Não detectado	Não detectado
	GAGTGAGTGTGTG ^T GTGTGG	2316.	1	Não detectado	Não detectado
	GAGTGAGTGTGTG ^T GTGgGG	2317.	1	Não detectado	Não detectado
	GAGTGAGTGTGTG ^T GTGTGG	2318.	1	Não detectado	Não detectado
	GAGTGAGTGTGTG ^C GTGgGG	2319.	1	Não detectado	Não detectado
	C ^T TGTGAGTGTGTGCGTGaGG	2320.	2	Não detectado	Não detectado
	A ^T TGTGAGTGTGTGCGTGTGG	2321.	2	Não detectado	Não detectado
	G ^C C ^T GAGTGTGTGCGTGTGG	2322.	2	Não detectado	Não detectado
	G ^T TGTGTGTGTGTGCGTGTGG	2323.	2	Não detectado	Não detectado
	G ^T TGTG ^G GTGTGTGCGTGTGG	2324.	2	Não detectado	Não detectado
	G ^C C ^T G ^T TGTGTGTGCGTGTGG	2325.	2	Não detectado	Não detectado
	G ^T TGTGTGTGTGTGCGTGgGG	2326.	2	Não detectado	Não detectado
	G ^T TGTG ^C GTGTGTGCGTGTGG	2327.	2	Não detectado	Não detectado
	G ^T TGTGTGTGTGTGCGTGcGG	2328.	2	Não detectado	Não detectado
	GAG ^A GAG ^A GTGTGCGTGTGG	2329.		Não detectado	Não detectado
	GAGTG ^T TGTG ^A GTGCGTGgGG	2330.	2	Não detectado	Não detectado
	G ^T TGTGAGTGTGTGTGTGTGG	2331.	2	Não detectado	Não detectado
	GAGTG ^T GTGT ^A TGCGTGTGG	2332.	2	Não detectado	Não detectado
	GAGT ^C AGTGTGTG ^A GTGaGG	2333.	2	Não detectado	Não detectado
	GAGTG ^T GTGTGTG ^A GTGTGG	2334.	2	Não detectado	Não detectado
	GAGTG ^T GTGTGTG ^C A ^T GTGG	2335.	2	Não detectado	Não detectado
	GAGTGAG ^A GTGTGTGTGTGG	2336.	2	Não detectado	Não detectado
	GAGTGAGTG ^A GTG ^A GTGaGG	2337.	2	Não detectado	Não detectado
Sítio 1 de <i>EMX1</i>	GTCCGAGCAGAAGAAGAAgGG	2338.	0 (sobre o alvo)	43,01 ± 0,87	17,25 ± 0,64
	GT ^C TGAGCAGAAGAAGAA ^t GG	2339.	1	Não detectado	Não detectado
	GTCC ^C AGCAG ^T AGAAGA ^t GG	2340.	2	Não detectado	Não detectado
	GTCCGAG ^G AGAG ^G GAAGAAaGG	2341.	2	Não detectado	Não detectado
	GT ^C A ^G AG ^G AGAAGAAGAAgGG	2342.	2	Não detectado	Não detectado
	GAC ^A GAGCAGAAGAAGAAgGG	2343.	2	Não detectado	Não detectado
	GT ^G GAGCAGAAGAAGAAgGG	2344.	2	Não detectado	Não detectado
	GT ^A C ^T AGCAGAAGAAGAAaGG	2345.	2	Não detectado	Não detectado
	GT ^C TGAGCAG ^A AGAAGA ^t GG	2346.	2	Não detectado	Não detectado
	GTG ^C TAGCAGAAGAAGAAgGG	2347.	2	Não detectado	Não detectado
	TAC ^A GAGCAGAAGAAGAA ^t GG	2348.	3	Não detectado	Não detectado
	TAC ^G GAGCAGAAGAAGAA ^t GG	2349.	3	Não detectado	Não detectado
	AA ^C GAGCAGAAGAAGAAaGG	2350.	3	Não detectado	Não detectado
	GAC ^A CAGCAGAAGAAGAAgGG	2351.	3	Não detectado	Não detectado
	C ^T TGCGA ^T CAGAAGAAGAAaGG	2352.	3	Não detectado	Não detectado
	GACTG ^G GCAGAAGAAGAAgGG	2353.	3	Não detectado	Não detectado
	ITCC ^C TGCAGAAGAAGAAaGG	2354.	3	Não detectado	Não detectado
	ITCC ^T AC ^C CAGAAGAAGAA ^t GG	2355.	3	Não detectado	Não detectado
	C ^T CTGAG ^G AGAAGAAGAAaGG	2356.	3	Não detectado	Não detectado
	ATCCA ^T TCAGAAGAAGAAgGG	2357.	3	Não detectado	Não detectado
	GCCC ^C TGCAGAAGAAGAAcGG	2358.	3	Não detectado	Não detectado
	ATCCA ^A CCAGAAGAAGAAaGG	2359.	3	Não detectado	Não detectado
	GACTGAG ^A AGAAGAAGAAaGG	2360.	3	Não detectado	Não detectado
	GTGGGATCAGAAGAAGAAaGG	2361.	3	Não detectado	Não detectado
	GACAGAG ^A AGAAGAAGAAaGG	2362.	3	Não detectado	Não detectado

Tabela 3C					
Frequências de mutação indel em sítios fora do alvo potenciais of tru-RGNs alvoed to endogenous genes in human cells					
Alvo ID	Sítio Alvo + PAM	SEQ ID NO:	Número de desemparelhamentos	Frequência de mutação Indel (%) ± s.e.m.	
				Células U2OS.EGFP	Células HEK293
	GTCTGGCAGAGAAGAAGAGG	2363.	3	Não detectado	Não detectado
	GTTGGAGAGAAGAAGAGG	2364.	3	Não detectado	Não detectado
	GTAGAGAGAAGAAGAGG	2365.	3	Não detectado	Não detectado
	CTCCTAGCAAGAAGAAGAGG	2366.	3	Não detectado	Não detectado
	TTCAAGCAGAGAAGAAGAGG	2367.	3	Não detectado	Não detectado
	GTTGGAGCAGAGAAGAAGAGG	2368.	3	Não detectado	Não detectado
	GCTGAGCAGAAGAGAAGAGG	2369.	3	Não detectado	Não detectado
	GTCTGAGGAGAAGAAGAGG	2370.	3	Não detectado	Não detectado
	GTCCGGAGAGAAGAAGAGG	2371.	3	Não detectado	Não detectado
	GCCGAGCAGAAGAAGAGG	2372.	3	Não detectado	Não detectado
	GTCTAGCAGAGAAGAAGAGG	2373.	3	Não detectado	Não detectado

Tabela 3D: Frequências de mutações indel induzidas por tru-RGN em sítios fora do alvo potenciais em células U2OS.EGFP humanas conforme determinado por sequenciamento profundo

Sítio sobre o alvo	Sequência de sítio fora do alvo	S#	tru-RGN			Controle		
			Indel	WT	Freq.	Indel	WT	Freq.
Sítio 1 de VEGFA	GTGGGGGGAGTTTGCCCaGG	2374.	1500	225640	0,66%	3	135451	0,00%
	GTGGGGGGTGTTCCTCCgGG	2375.	1552	152386	1,01%	0	86206	0,00%
	GTGGGTGGAGTTTGCTCtGG	2376.	1	471818	0,00%	0	199581	0,00%
	GTGGGTGGAGTTTGCTCaGG	2377.	0	337298	0,00%	1	211547	0,00%
	GTGGGTGGCTTTGCTCCaGG	2378.	2	210174	0,00%	1	105531	0,00%
	GTGTGGGAATTGCTCCaGG	2379.	673	715547	0,09%	1	387097	0,00%
	GTGGGGGAGCTTCTCCtGG	2380.	5	107757	0,00%	1	58735	0,00%
	GGGGGGGAGTTTGCTCCtGG	2381.	1914	566548	0,34%	3	297083	0,00%
Sítio 3 de VEGFA	GTTGTGAGTGTGCGTGtGG	2382.	58	324881	0,02%	9	122216	0,01%
	GTTGTGAGTGTGCGTGaGG	2383.	532	194914	0,27%	11	73644	0,01%
	GAGTGGGTGTGTCGTGgGG	2384.	70	237029	0,03%	10	178258	0,01%
	GAGTGACTGTGTCGTGtGG	2385.	6	391894	0,00%	0	239460	0,00%
	GAGTGAGTGTGTCGTGgGG	2386.	15	160140	0,01%	10	123324	0,01%
	GTTGTGAGTGTGCGTGgGG	2387.	19	138687	0,01%	1	196271	0,00%
	CAGTGAGTGTGTCGTGtGG	2388.	78	546865	0,01%	41	355953	0,01%
	GTTGTGAGTGTGCGTGtGG	2389.	128	377451	0,03%	56	133978	0,04%
	GAGTGTTGTGTGCGTGtGG	2390.	913	263028	0,35%	78	178979	0,04%
	GAGTGAGTGTGTGTGtGG	2391.	40	106933	0,04%	36	58812	0,06%
	GAGTGAGTGTGTGTGtGG	2392.	681	762999	0,09%	63	222451	0,03%
	GAGTGAGTGTGTGTGgGG	2393.	331	220289	0,15%	100	113911	0,09%
	GAGTGAGTGTGTGTGtGG	2394.	0	35725	0,00%	8	186495	0,00%
	GAGTGAGTGTGTGCGCGgGG	2395.	94	246893	0,04%	16	107623	0,01%
Sítio 1 de EMX1	GTCAGAGAGAAGAAGAGG	2396.	0	201483	0,00%	4	148416	0,00%
	GTCAGAGAGAAGAAGAGG	2397.	10	545662	0,00%	5	390884	0,00%
	GTCAGAGCAAGAAGAAGAGG	2398.	2	274212	0,00%	0	193837	0,00%
	GTCAGAGCAGAAGAAGAAGAGG	2399.	440	375646	0,12%	10	256181	0,00%
	GACAGAGCAGAAGAAGAAGAGG	2400.	2	212472	0,00%	1	158860	0,00%
	GTAAGTACAGAGAAGAAGAAGAGG	2401.	152	229209	0,07%	103	157717	0,07%
	GAGGAGCAGAAGAAGAAGAGG	2402.	50	207401	0,02%	36	111183	0,03%
	GTCCAGCAGTAGAAGAAGAGG	2403.	0	226477	0,00%	1	278948	0,00%

Exemplo 2d. tru-gRNAs podem ser usados com nickases Cas9 duplas para induzir de forma eficiente à edição de genoma em células humanas

[0158] tru-gRNAs foram testados com a abordagem de nickase

dupla Cas9 recentemente descrita para induzir a mutações indel. Para fazer isso, a nickase Cas9-D10A juntamente com dois gRNAs de comprimento completo objetivados a sítios no gene de *VEGFA* humano (sítio 1 de *VEGFA* e uma sequência adicional que nós referimos como sítio 4 de *VEGFA*) foram coexpressos em células U2OS.EGFP (**Figura 4A**). Conforme descrito anteriormente (Ran *et al.*, 2013), este par de nickases funcionou cooperativamente para induzir a altas taxas de mutações indel no *locus* de *VEGFA* (**Figura 4B**). Curiosamente, a nickase Cas9-D10A coexpressa apenas com o gRNA objetivado ao sítio 4 de *VEGFA* também induziu a mutações indel em uma elevada frequência, embora em uma taxa ligeiramente menor do que aquela observada com os gRNAs de comprimento completo pareados (**Figura 4B**). Mais importante, o uso de um tru-gRNA ao sítio 1 de *VEGFA*, em lugar de um gRNA de comprimento completo, não afetou a eficácia da abordagem de nickase dupla para induzir a mutações indel (**Figura 4B**).

[0159] A estratégia de nickase dupla também foi usada para estimular a introdução de alterações específicas para sequência usando ssODNs (Mali *et al.*, 2013a; Ran *et al.*, 2013) e, assim, se tru-gRNAs poderiam ser usados para este tipo de alteração foi também testado. gRNAs de comprimento completo pareados para os sítios 1 e 4 de *VEGFA* em conjunto com nickase Cas9-D10A reforçou cooperativamente a introdução eficiente de uma inserção curta de um doador ssODN (**Figura 3A**) no *locus* de *VEGFA* em células U2OS.EGFP humanas, conforme esperado (**Figura 3C**). Mais uma vez, a eficiência de alteração de sequência mediada por ssODN por nicking duplo permaneceu igualmente elevada com o uso de um tru-gRNA em lugar do gRNA de comprimento completo objetivado ao sítio 1 de *VEGFA* (**Figura 3C**). Tomados em conjunto, estes resultados demonstram que tru-gRNAs podem ser usados como parte de uma estratégia de nickase Cas9 dupla para induzir tanto a mutações indel quanto alterações de

sequência mediadas por ssODN, sem comprometer a eficácia de edição de genoma com esta abordagem.

[0160] Tendo estabelecido que o uso de um tru-gRNA não diminui as atividades de edição de genoma sobre o alvo de nickases pareadas, em seguida, nós usamos sequenciamento profundo para examinar frequências de mutação em quatro sítios de fora do alvo *bona fide* previamente identificados do gRNA para o sítio 1 de *VEGFA*. Esta análise revelou que as taxas de mutação caíram para níveis não detectáveis essencialmente em todos os quatro destes sítios fora do alvo quando se usa nickases pareadas com um tru-gRNA (**Tabela 4**). Em contraste, nem um tru-RGN (**Tabela 3B**) nem as nickases pareadas com gRNAs de comprimento completo (**Tabela 4**) foram capazes de eliminar completamente mutações fora do alvo em um destes quatro sítios fora do alvo (OT1-3). Estes resultados demonstram que o uso de tru-gRNAs pode reduzir ainda mais os efeitos fora do alvo de nickases Cas9 pareadas (e vice-versa) sem comprometer a eficiência de edição de genoma sobre o alvo.

Tabela 4. Frequências de mutações indel induzidas por nickase pareada em sítios sobre e fora do alvo do sítio 1 de <i>VEGFA</i> usando gRNAs de comprimento completo e tru-gRNAs									
Sítio	gRNAs de comprimento completo pareados			tru-gRNA/gRNA de comprimento completo			Controle		
	Indel	WT	Freq.	Indel	WT	Freq.	Indel	WT	Freq.
Sítio 1 de <i>VEGFA</i>	78905	345696	18,583%	65754	280720	18,978%	170	308478	0,055%
OT1-3	184	85151	0,216%	0	78658	0,000%	2	107850	0,002%
OT1-4	0	89209	0,000%	1	97010	0,001%	0	102135	0,000%
OT1-6	2	226575	0,001%	0	208218	0,000%	0	254580	0,000%
OT1-11	0	124729	0,000%	0	121581	0,000%	0	155173	0,000%

REFERÊNCIAS

Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B., and Jaenisch, R. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res* 23, 1163-1171. (2013).

Cho, S.W., Kim, S., Kim, J.M. & Kim, J.S. Targeted genome

engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* 31, 230-232 (2013).

Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-823 (2013).

Cradick, T.J., Fine, E.J., Antico, C.J., and Bao, G. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* (2013).

Dicarlo, J.E. et al. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res* (2013).

Ding, Q., Regan, S.N., Xia, Y., Oostrom, L.A., Cowan, C.A., and Musunuru, K. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell* 12, 393-394. (2013).

Fisher, S., Barry, A., Abreu, J., Minie, B., Nolan, J., Delorey, T.M., Young, G., Fennell, T.J., Allen, A., Ambrogio, L., et al. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biol* 12, R1. (2011).

Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiacovo, M.P., Church, G.M., and Calarco, J.A. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat Methods* 10, 741-743. (2013).

Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* 31, 822-826. (2013).

Gabriel, R. et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol* 29, 816-823 (2011).

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell* 154, 442-451.

Gratz, S.J. et al. Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* (2013).

Hockemeyer, D. et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* 29, 731-734 (2011).

Horvath, P. & Barrangou, R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327, 167-170 (2010).

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* 31, 827-832. (2013).

Hwang, W.Y. et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 31, 227-229 (2013).

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Kaini, P., Sander, J.D., Joung, J.K., Peterson, R.T., and Yeh, J.R. Heritable and Precise Zebrafish Genome Editing Using a CRISPR-Cas System. *PLoS One* 8, e68708. (2013a).

Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L.A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 31, 233-239 (2013).

Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821 (2012).

Jinek, M. et al. RNA-programmed genome editing in human cells. *Elife* 2, e00471 (2013).

Li, D., Qiu, Z., Shao, Y., Chen, Y., Guan, Y., Liu, M., Li, Y., Gao, N., Wang, L., Lu, X., et al. Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat Biotechnol* 31, 681-683. (2013a).

Li, W., Teng, F., Li, T., and Zhou, Q. Simultaneous generation

and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nat Biotechnol* 31, 684-686. (2013b).

Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., and Joung, J.K. CRISPR RNA-guided activation of endogenous human genes. *Nat Methods* 10, 977-979. (2013).

Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* 31, 833-838. (2013a).

Mali, P., Esvelt, K.M., and Church, G.M. Cas9 as a versatile tool for engineering biology. *Nat Methods* 10, 957-963. (2013b).

Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* 339, 823-826 (2013c).

Pattanayak, V., Lin, S., Guilinger, J.P., Ma, E., Doudna, J.A., and Liu, D.R. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* 31, 839-843. (2013).

Pattanayak, V., Ramirez, C.L., Joung, J.K. & Liu, D.R. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods* 8, 765-770 (2011).

Perez, E.E. et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol* 26, 808-816 (2008).

Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods* 10, 973-976. (2013).

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.

Cell 152, 1173-1183. (2013).

Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380-1389. (2013).

Reyon, D. et al. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotech* 30, 460-465 (2012).

Sander, J.D., Maeder, M.L., Reyon, D., Voytas, D.F., Joung, J.K., and Dobbs, D. ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Res* 38, W462-468. (2010).

Sander, J.D., Ramirez, C.L., Linder, S.J., Pattanayak, V., Shores, N., Ku, M., Foden, J.A., Reyon, D., Bernstein, B.E., Liu, D.R., et al. In silico abstraction of zinc finger nuclease cleavage profiles reveals an expanded landscape of off-target sites. *Nucleic Acids Res.* (2013).

Sander, J.D., Zaback, P., Joung, J.K., Voytas, D.F., and Dobbs, D. Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool. *Nucleic Acids Res* 35, W599-605. (2007).

Shen, B. et al. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res* (2013).

Sugimoto, N. et al. Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry* 34, 11211-11216 (1995).

Terns, M.P. & Terns, R.M. CRISPR-based adaptive immune systems. *Curr Opin Microbiol* 14, 321-327 (2011).

Wang, H. et al. One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell* 153, 910-918 (2013).

Wiedenheft, B., Sternberg, S.H. & Doudna, J.A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331-338

(2012).

Yang, L., Guell, M., Byrne, S., Yang, J.L., De Los Angeles, A., Mali, P., Aach, J., Kim-Kiselak, C., Briggs, A.W., Rios, X., et al. (2013). Optimization of scarless human stem cell genome editing. Nucleic Acids Res 41, 9049-9061.

OUTRAS MODALIDADES

[0161] Deverá ser entendido que, embora a invenção tenha sido descrita em conjunto com a descrição detalhada da mesma, a descrição precedente se destina a ilustrar, e não limitar, o âmbito da invenção, o qual é definido pelo âmbito das reivindicações anexas. Outros aspectos, vantagens e modificações estão dentro do âmbito das reivindicações a seguir.

REIVINDICAÇÕES

1. Método para aumentar a especificidade de edição de genoma orientado por RNA em uma célula, caracterizado pelo fato de que compreende o contato da célula com um RNA guia de Cas9 que inclui uma região de complementaridade que consiste em 17 a 18 nucleotídeos que são complementares a 17 a 18 nucleotídeos consecutivos da fita complementar de uma sequência genômica alvo selecionada, em que a sequência alvo está imediatamente a 5' de um motivo adjacente ao protoespaçador (PAM).

2. Método de indução de uma ruptura em uma região alvo de uma molécula de DNA de fita dupla em uma célula e de aumento de especificidade da edição do genoma guiado por RNA em uma célula, caracterizado pelo fato de que compreende expressão ou introdução, na célula, de:

uma nuclease Cas9 ou uma nickase Cas9; e

um RNA guia de Cas9 que inclui uma região de complementaridade que consiste em 17 a 18 nucleotídeos que são complementares a 17 a 18 nucleotídeos consecutivos da fita complementar da molécula de DNA de fita dupla, em que a região alvo está imediatamente a 5' de um motivo adjacente ao protoespaçador (PAM).

3. Método de modificação de uma região alvo de uma molécula de DNA de fita dupla em uma célula e de aumento de especificidade da edição do genoma guiado por RNA em uma célula, caracterizado pelo fato de compreender expressão ou introdução, na célula, de:

Uma proteína de fusão do domínio funcional de dCas9heteróloga (dCas9-HFD) em que a dCas9 possui mutações catalíticas inativadoras em um ou nos dois domínios de nuclease;

em que a proteína de fusão do domínio funcional dCas9

heteróloga compreende um domínio funcional heterólogo (HFD) selecionado do grupo que consiste em: um HFD que modifica a expressão gênica, histonas ou DNA; um domínio de ativação transcricional; uma enzima que catalisa a desmetilação do DNA; uma enzima que catalisa a modificação de histonas ou um domínio de silenciamento de transcrição; e

um RNA guia de Cas9 que inclui uma região de complementaridade que consiste em 17 a 18 nucleotídeos que são complementares a 17 a 18 nucleotídeos consecutivos da fita complementar de uma sequência genômica alvo selecionada, em que a sequência alvo está imediatamente a 5' de um motivo adjacente ao protoespaçador (PAM).

4. Método, de acordo com as reivindicações 1 a 3, caracterizado pelo fato de que o RNA guia de Cas9 é:

(i) um RNA guia simples que inclui uma região de complementaridade que consiste em 17 a 18 nucleotídeos que são complementares a 17 a 18 nucleotídeos consecutivos da fita complementar de uma sequência genômica alvo selecionada, ou

(ii) um crRNA que inclui uma região de complementaridade que consiste em 17 a 18 nucleotídeos que são complementares a 17 a 18 nucleotídeos consecutivos da fita complementar de uma sequência genômica alvo selecionada, e um tracrRNA.

5. Método, de acordo com as reivindicações 1 a 3, caracterizado pelo fato de que o RNA guia de Cas9 é ou compreende um ácido ribonucleico selecionado do grupo que consiste em:

(X₁₇₋₁₈) GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 2407); ou

(X₁₇₋₁₈) GUUUUAGAGCUAGAAUAGCAAGUUAAAAUAAG GCUAGUCCG (SEQ ID NO: 1),

(X₁₇₋₁₈) GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUU

AAAAUAAGGCUAGUCCGUUAUC (SEQ ID NO: 2),

(X₁₇₋₁₈)GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACA
GCAUAGCAAGUUA AAAUAAGGCUAGUCCGUUAUC (SEQ ID NO:
3),

(X₁₇₋₁₈)GUUUUAGAGCUAGAAAUAGCAAGUUA AAAUAAG
GCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC
(SEQ ID NO: 4),

em que X₁₇₋₁₈ é uma região de complementaridade que é complementar a 17 a 18 nucleotídeos consecutivos da fita complementar de uma sequência genômica alvo selecionada.

6. Método, de acordo com a reivindicação 5, caracterizado pelo fato de que o ácido ribonucleico inclui um ou mais Us na extremidade 3' da molécula.

7. Método, de acordo com a reivindicação 5, caracterizado pelo fato de que o ácido ribonucleico inclui um ou mais nucleotídeos adicionais na extremidade 5' da molécula de RNA que não são complementares à sequência alvo.

8. Método, de acordo com a reivindicação 5, caracterizado pelo fato de que o ácido ribonucleico inclui um, dois ou três nucleotídeos adicionais na extremidade 5' da molécula de RNA que não são complementares à sequência alvo.

9. Método, de acordo com qualquer uma das reivindicações 1 a 5, caracterizado pelo fato de que a região de complementaridade é complementar a 17 nucleotídeos consecutivos da fita complementar de uma sequência alvo selecionada.

10. Método, de acordo com qualquer uma das reivindicações 1 a 5, caracterizado pelo fato de que a região de complementaridade é complementar a 18 nucleotídeos consecutivos da fita complementar de uma sequência alvo selecionada.

11. Método, de acordo com a reivindicação 4, caracterizado

pelo fato de que o tracrRNA consiste na sequência GGAACCAUUCAAAACAGCAUAGCAAGUUAUAAAUAAGGCUAGUCC GUUAUCAACUUGAAAAAGUGGGCACCGAGUCGGUGC (SEQ ID NO: 8), ou uma porção ativa da mesma; UAGCAAGUUAUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC (SEQ ID NO: 2405) ou uma porção ativa da mesma;

AGCAUAGCAAGUUAUAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGC (SEQ ID NO: 2407) ou uma porção ativa da mesma; CAAAACAGCAUAGCAAGUUAUAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGGCACCGAGUCGGUGC (SEQ ID NO: 2409) ou uma porção ativa da mesma; UAGCAAGUUAUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGU G (SEQ ID NO: 2410) ou uma porção ativa da mesma; UAGCAAGUUAUAAAUAAGGCUAGUCCGUUAUCA (SEQ ID NO: 2411) ou uma porção ativa da mesma; ou UAGCAAGUUAUAAAUAAGGCUAGUCCG (SEQ ID NO: 2412) ou uma porção ativa da mesma.

12. Método, de acordo com a reivindicação 4, caracterizado pelo fato de que o cRNA é (X₁₇₋₁₈)GUUUUAGAGCUAUGCUGUUUUG (SEQ ID NO: 2407) e o tracrRNA é GGAACCAUUCAAAACAGCAUAGCAAGUUAUAAAUAAGGCUAGUCC GUUAUCAACUUGAAAAAGUGGGCACCGAGUCGGUGC (SEQ ID NO: 8); o cRNA é (X₁₇₋₁₈)GUUUUAGAGCUA (SEQ ID NO: 2404) e o tracrRNA é UAGCAAGUUAUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC (SEQ ID NO: 2405); ou o cRNA é (X₁₇₋₁₈) GUUUUAGAGCUAUGCU (SEQ ID NO: 2408) e o tracrRNA é AGCAUAGCAAGUUAUAAAUAAGGCUAGUCCGUUAUCAACUUGAAA

AAGUGGCACCGAGUCGGUGC (SEQ ID NO: 2406).

13. Método, de acordo com a reivindicação 3, caracterizado pelo fato de que o domínio de ativação de transcrição é de VP64 ou p65 de NF-kB.

14. Método, de acordo com a reivindicação 3, caracterizado pelo fato de que a enzima que catalisa a modificação de histona é LSD1, uma histona metiltransferase (HNMT), histona acetiltransferase (HAT), histona desacetilase (HDAC) ou histona demetilase.

15. Método, de acordo com a reivindicação 3, caracterizado pelo fato de que o domínio de silenciamento de transcrição é proteína 1 de heterocromatina (HP1), por exemplo, HP1 α ou HP1 β .

16. Método, de acordo com qualquer uma das reivindicações 1 a 5, caracterizado pelo fato de resultar em uma mutação indel ou alteração de sequência na sequência genômica alvo selecionada.

17. Método, de acordo com qualquer uma das reivindicações 1 a 5, caracterizado pelo fato de que a célula é uma célula eucariótica.

18. Método, de acordo com a reivindicação 17, caracterizado pelo fato de que a célula é uma célula de mamífero.

CRISPR/Cas9 - RGN

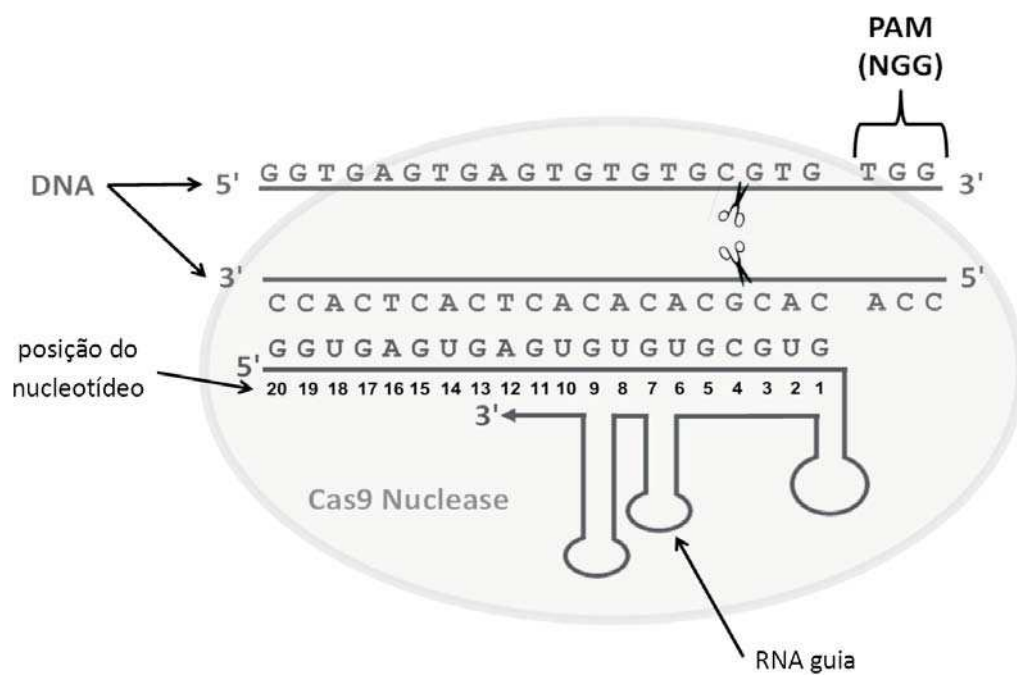


Figura 1

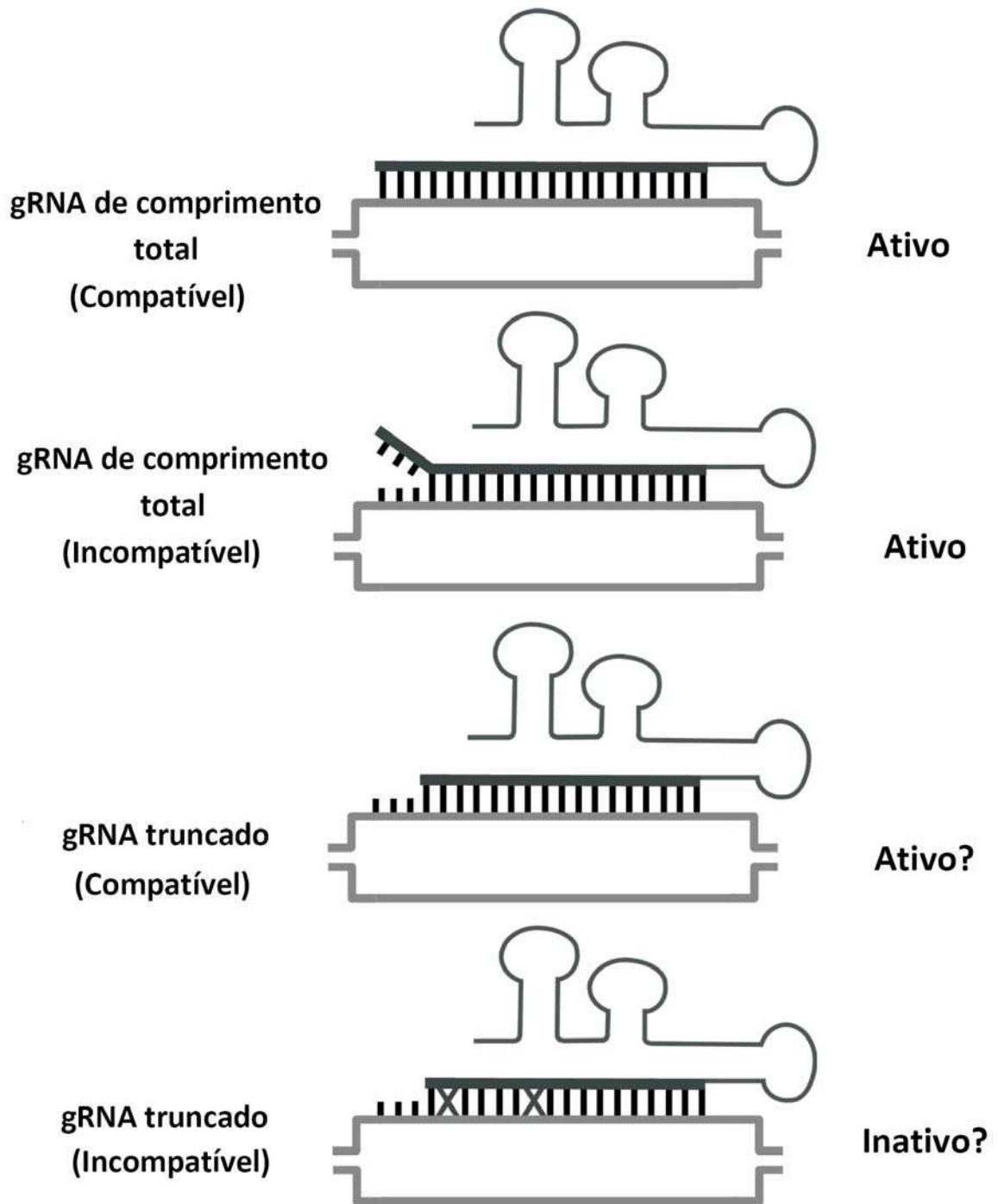


Figura 2A

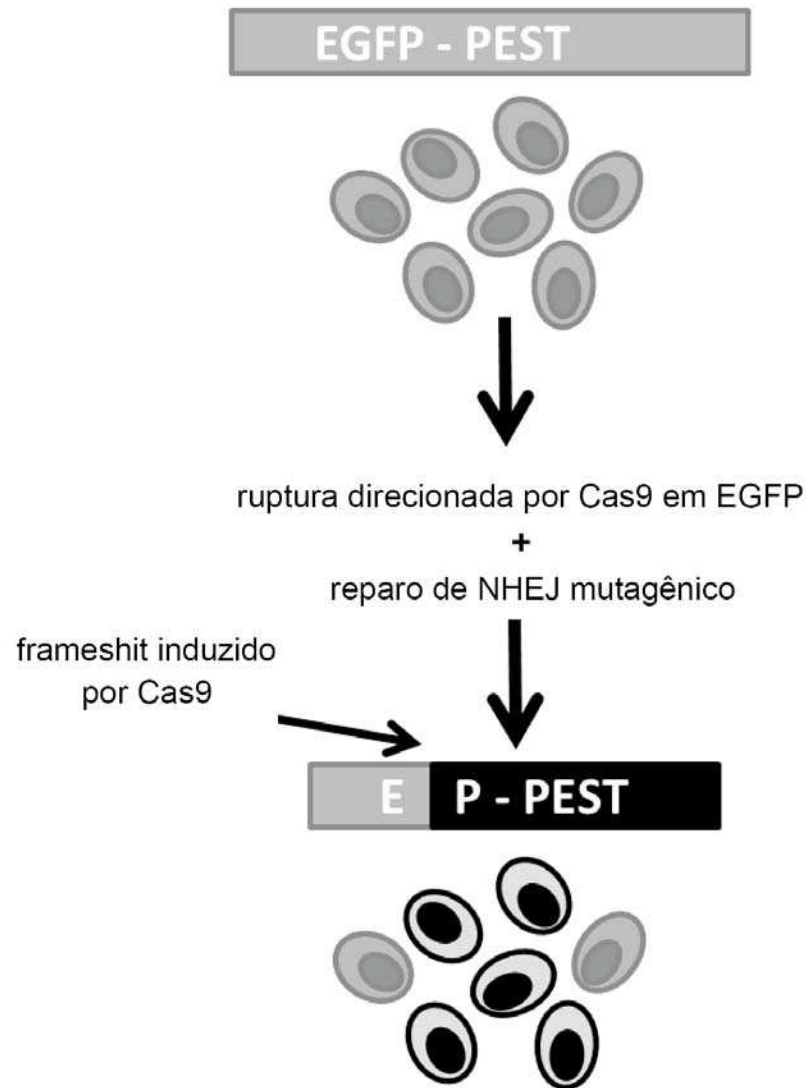


Figura 2B

Interrupção média de EGFP (relativa a gRNA perfeitamente compatível)

3'
↑
5'

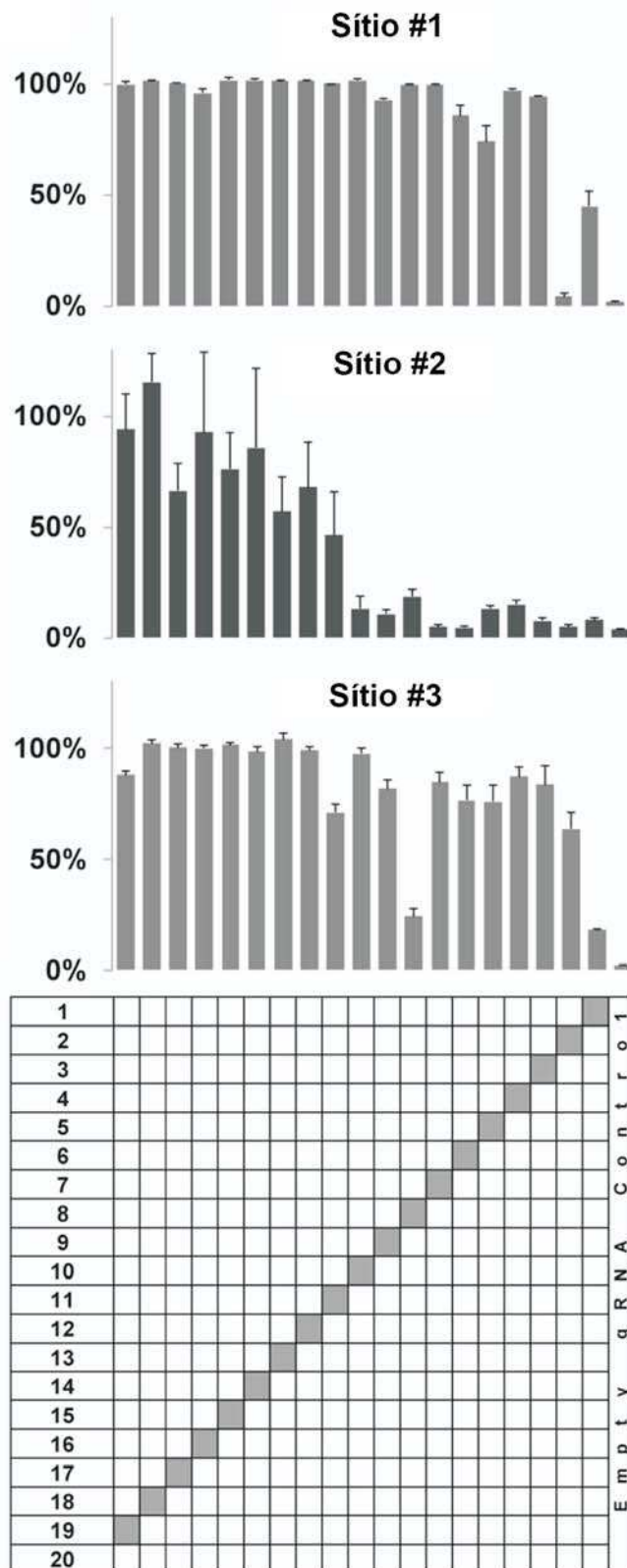


FIG. 2C

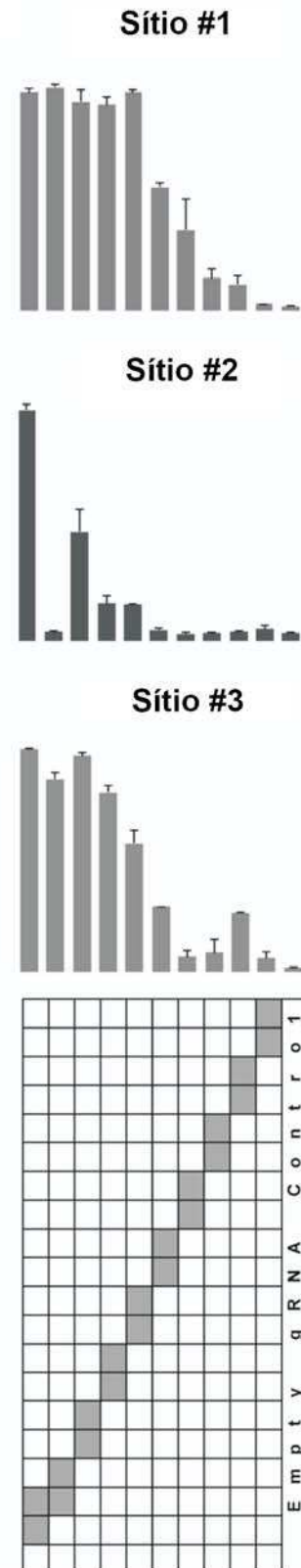


FIG. 2D

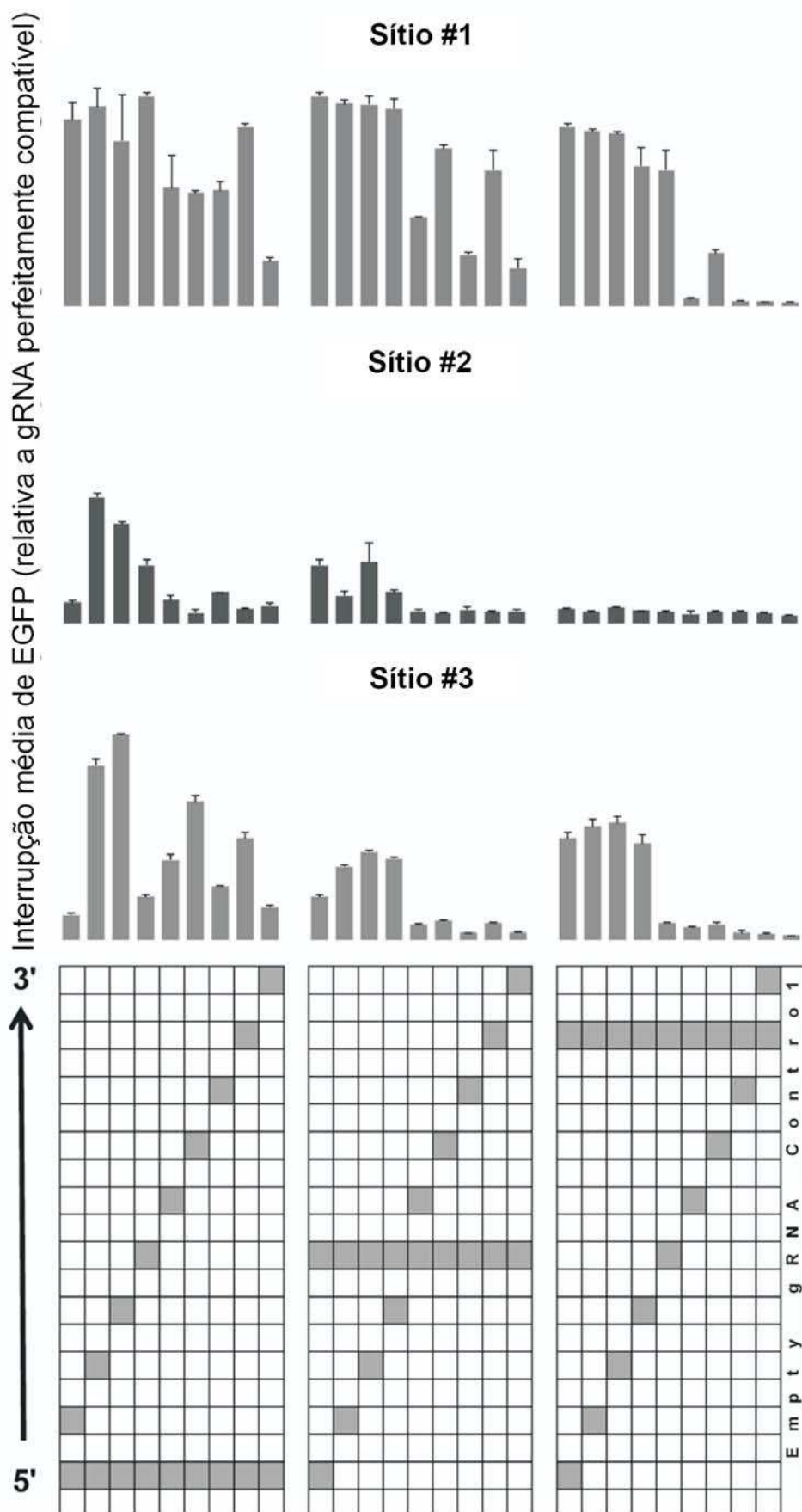


FIG. 2E

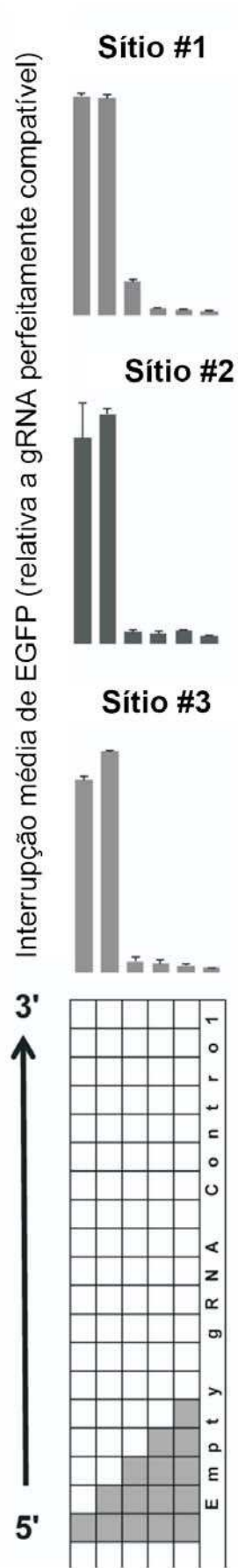


FIG. 2F

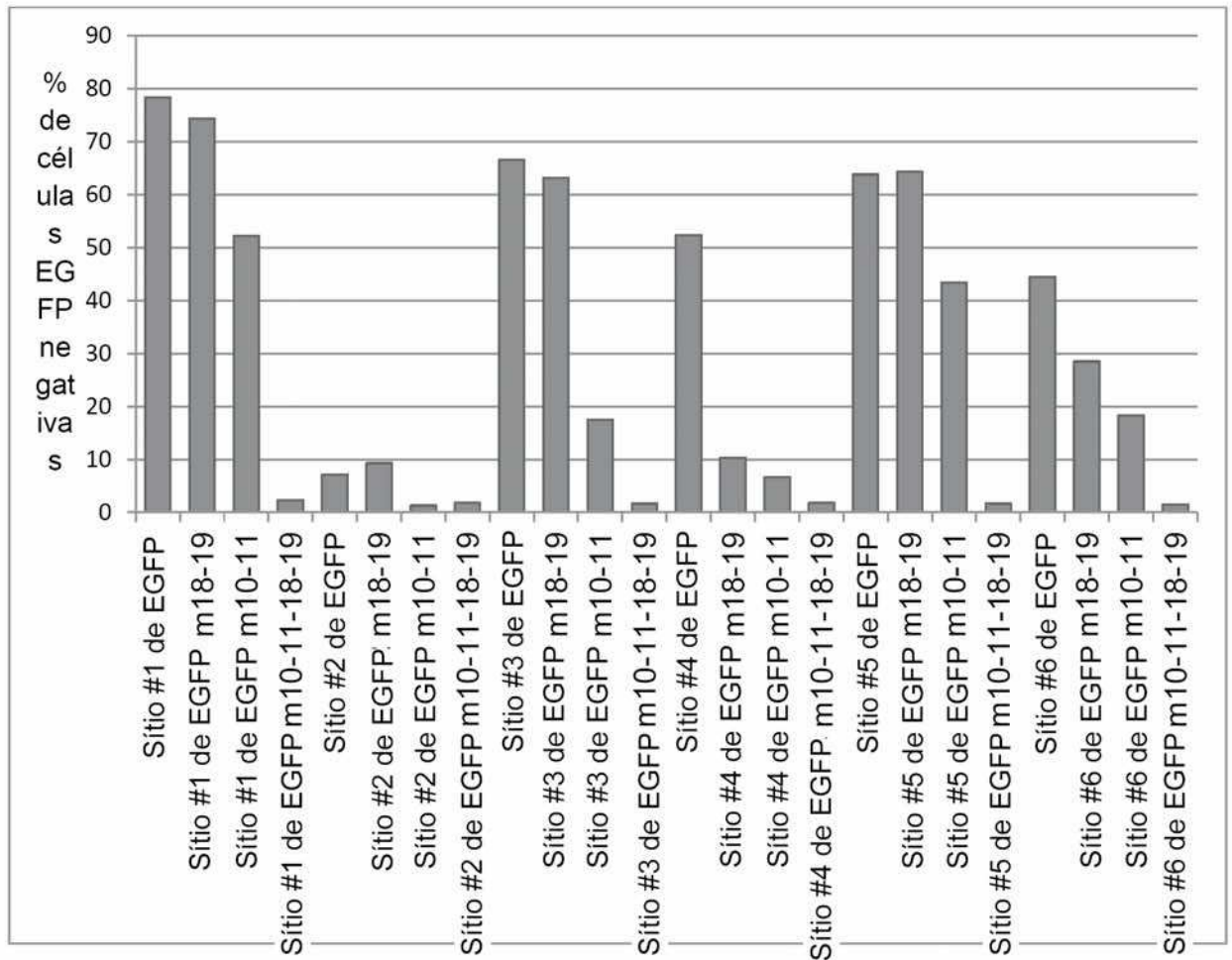


FIG. 2G

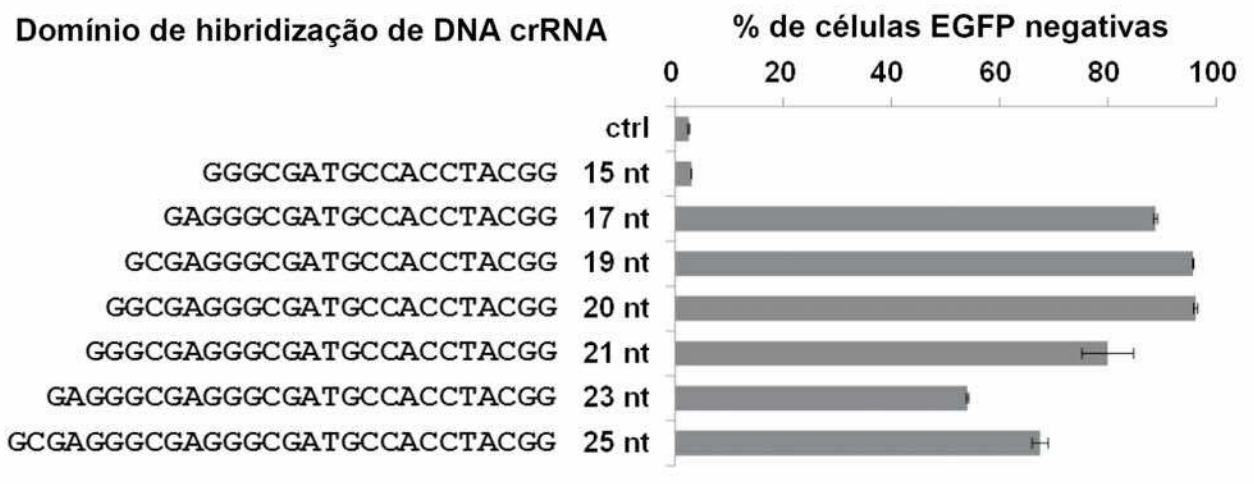
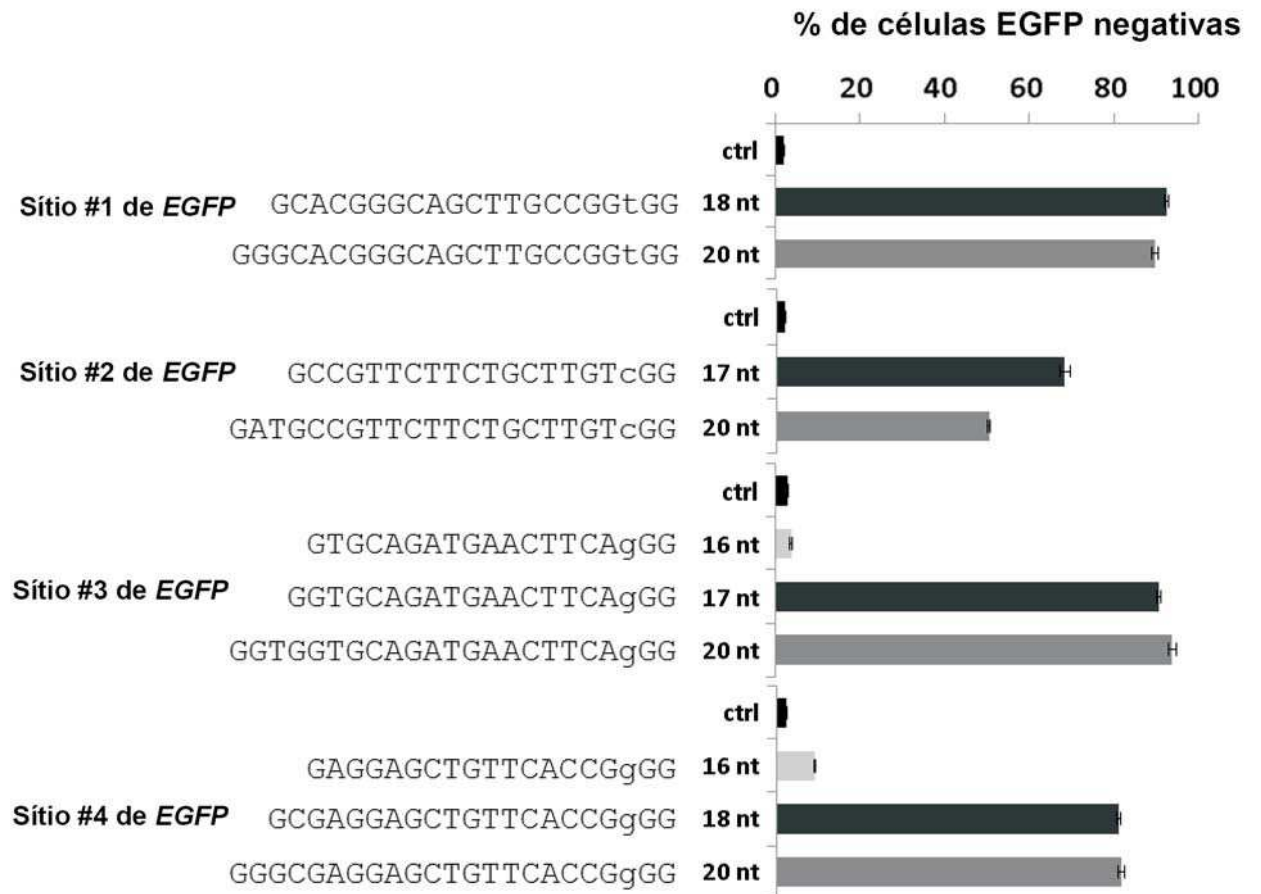


FIG. 2H

**FIG. 3A**

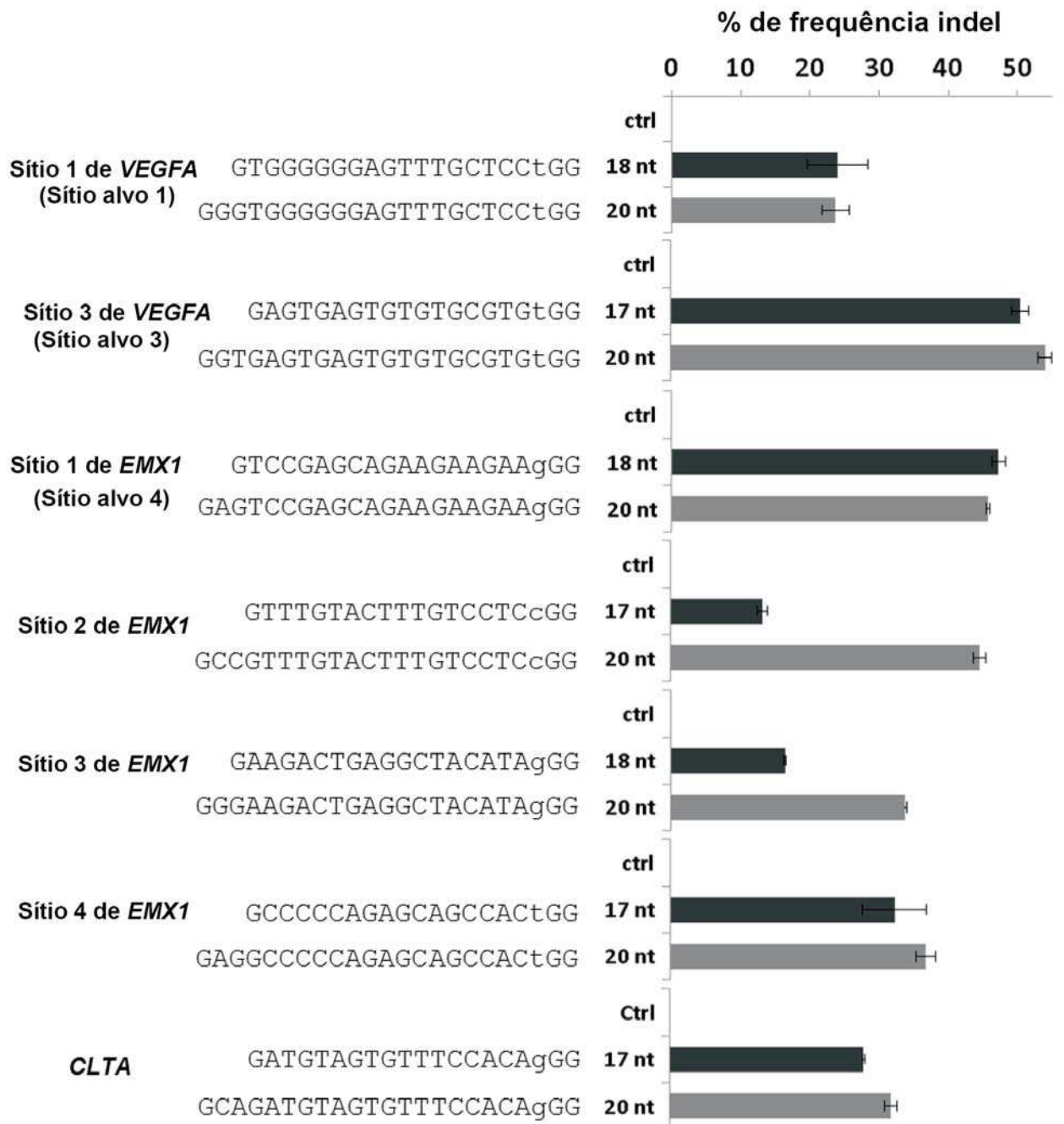


FIG. 3B

gRNA truncado de *EMX1*

GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAgGGCTCCCATCACATCAACCGGTGG	tipo selvagem	x24
GAAGCTGGAGGAGGA----->	Δ365	
<-----TCAACCGGTGG	Δ181	
GAAGCTGGAGGAGGAAGG----->	Δ138	
<----->	Δ126	
<-----GGGCTCCCATCACATCAACCGGTGG	Δ114	
GAAGCTGGAGGAGGAAGGGCCTGA----->	Δ101	
GAAGCTGGAGGA-----GG	Δ53	
GAAGCTGGAGGAGGAAGGG-----CCCATCACATCAACCGGTGG	Δ28	
GAAGCTGGAGGAGGAAGGGC-----TCGCACACATCAACCGGTGG	Δ27	
GAAGCTGGAGGAGGAAGGGC-----CTTCCATCACATCAACCGGTGG	Δ25	
GAAGCTGGAGGAGGAAGGGCCTGAG-----TCCCATCACATCAACCGGTGG	Δ21	x2
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAG-----TCCCATCACATCAACCGGTGG	Δ15	
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAG-----TCCCATCACATCAACCGGTGG	Δ9	
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGCGTTTGTAG-----CCATCACATCAACCGGTGG	Δ8	
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGA-----GCTCCCATCACATCAACCGGTGG	Δ6	
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAA-----CTCCCATCACATCAACCGGTGG	Δ6	
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGC-----AGAAGAAGGGCTCCCATCACATCAACCGGTGG	Δ3	x3
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGA--AAGGGCTCCCATCACATCAACCGGTGG	Δ2	
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAACAGAAGGGCTCCCATCACATCAACCGGT	+2	

gRNA de comprimento total de *EMX1*

GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAgGGCTCCCATCACATCAACCGGTGG	tipo selvagem	x35
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAG----->	Δ202	
<----->	Δ115	
GAA----->	Δ94	
<----->	Δ78	
GAAGCTGGAGG----->	Δ72	
GAAGCTGGA-----GG	Δ56	
GAAGCTGGAGGAGGAAGGGCCTGA-----GTGG	Δ39	
GAAGCTGGAGGAG-----GAAGGGCTCCCATCACATCAACCGGTGG	Δ26	x2
GAAGCTGGAGGAGGAAGGGCCTGAGT-----CCATCACATCAACCGGTGG	Δ22	
GAAGCTGGAGGAGGAAGGGCCTGAG-----TCCCATCACATCAACCGGTGG	Δ21	x3
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAG-----CATCACATCAACCGGTGG	Δ18	
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGA-----GCTCCCATCACATCAACCGGTGG	Δ14	
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGC-----AGAAGGGCTCCCATCACATCAACCGGTGG	Δ6	x3
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGC---AGAAGAAGGGCTCCCATCACATCAACCGGTGG	Δ3	x3
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGA--AAGAAGGGCTCCCATCACATCAACCGGTGG	Δ2	x2
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAACAGAAGGGCTCCCATCACATCAACCGGT	+2	

FIG. 3C

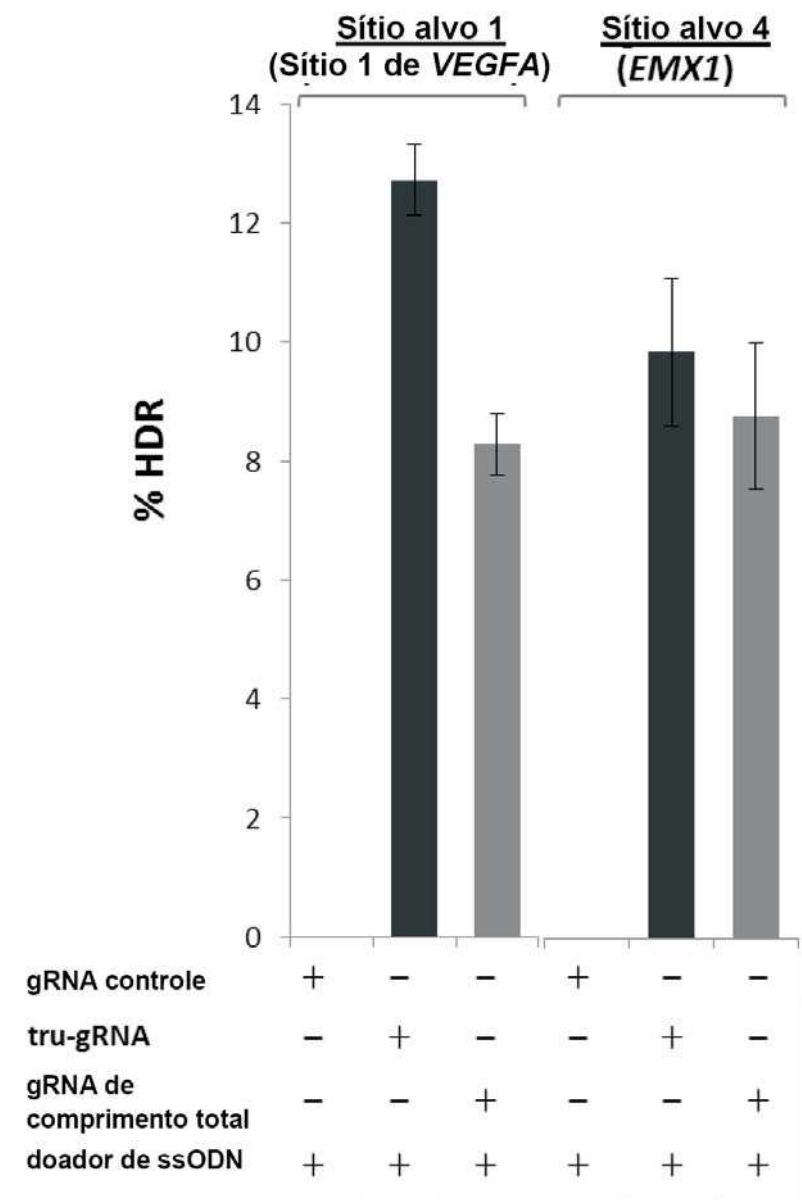
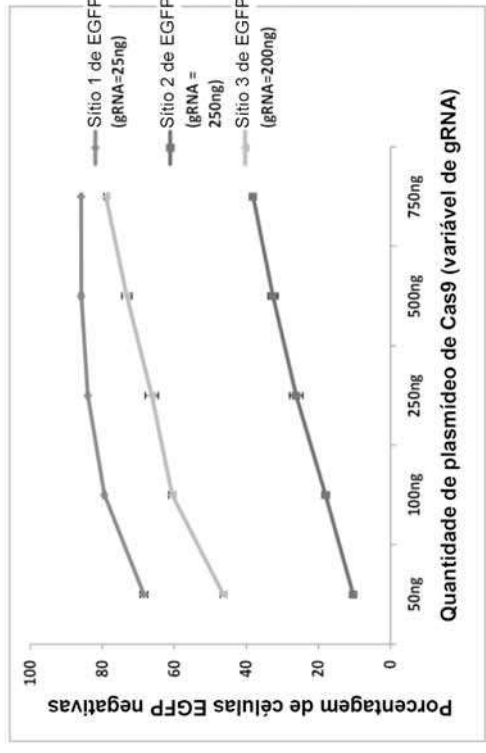
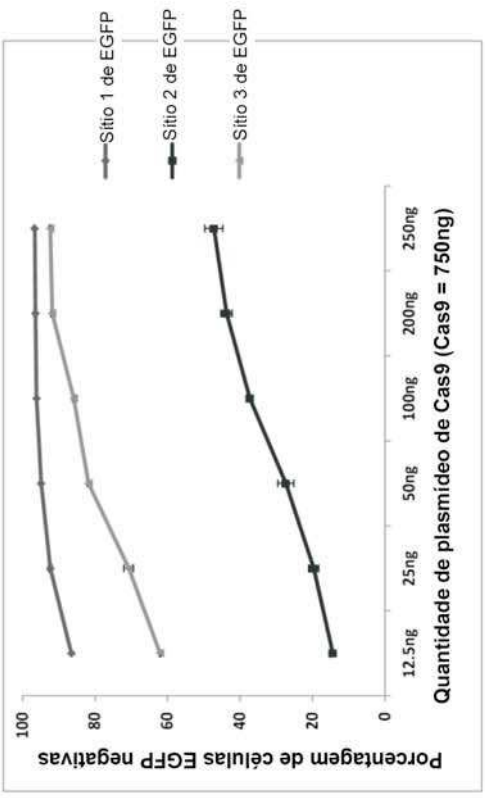


FIG. 3D



12/28

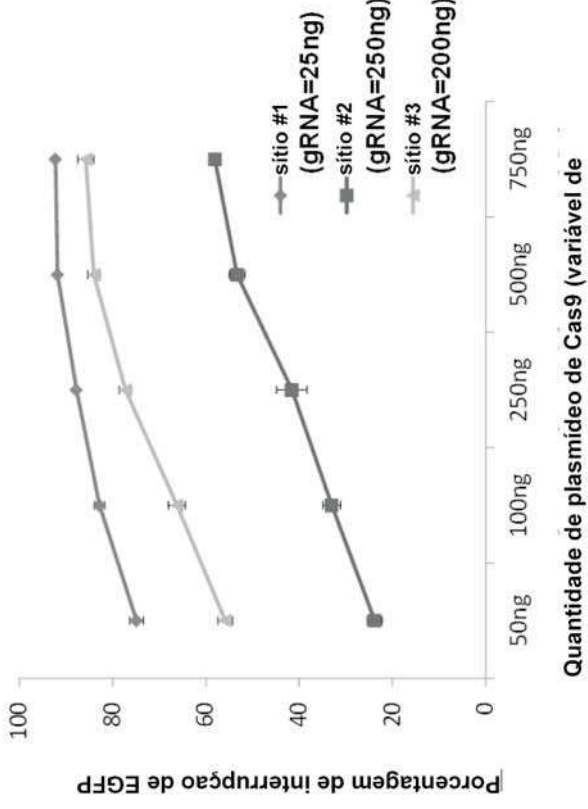
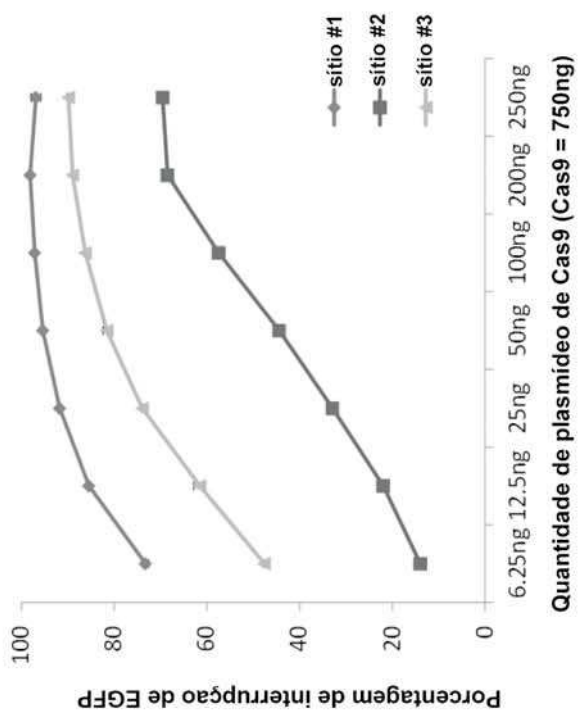


Figura 3E

Figura 3F

Sítio 4 de VEGFA

AGGTCAGAAATAGGGGGTCCAGGAGCAAACTCCCCCCCACCCCCTTCCAAAGCCCATTCCTCTTTAGCCAGAGCCG⁵GGTGTGCAGACGGCAGTCA

TCCAGTCTTTATCCCCCAGGtCCTCGTTTGAAGGGGGTGGGGAAAGGTTTCGGGTAAAGGAGAAATCGGTCTCGGCCCCACACGTCTGCCGTCAGT

Sítio 1 de VEGFA

FIG. 4A

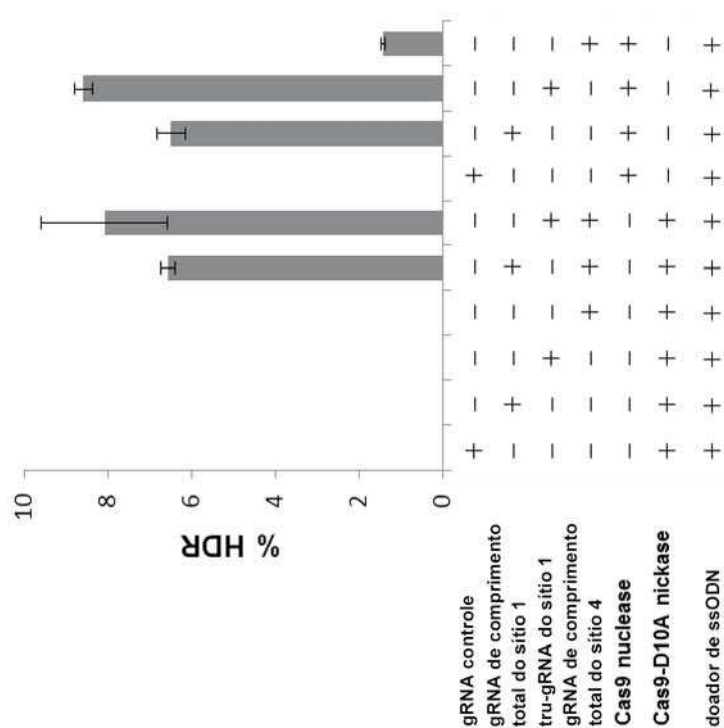
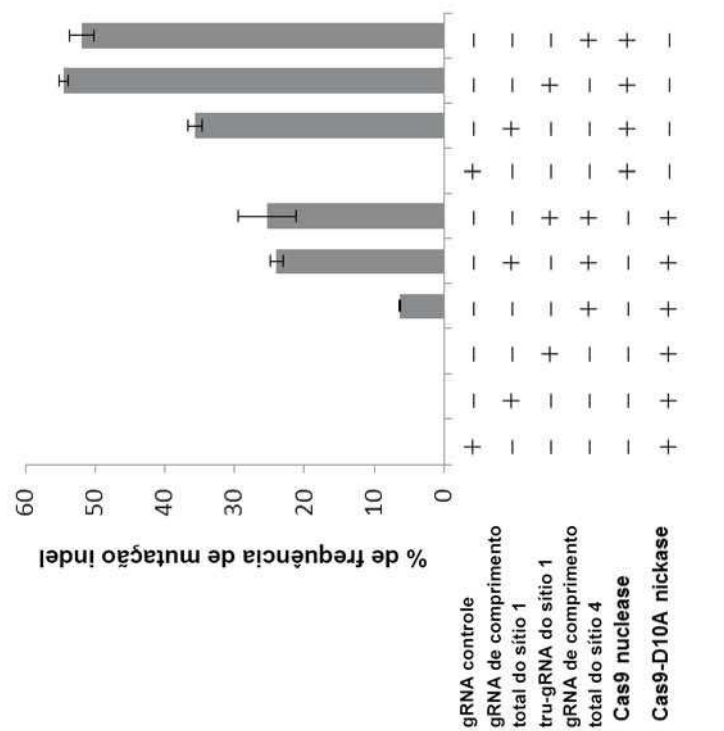


FIG. 4B

FIG. 4C

Figura 5A

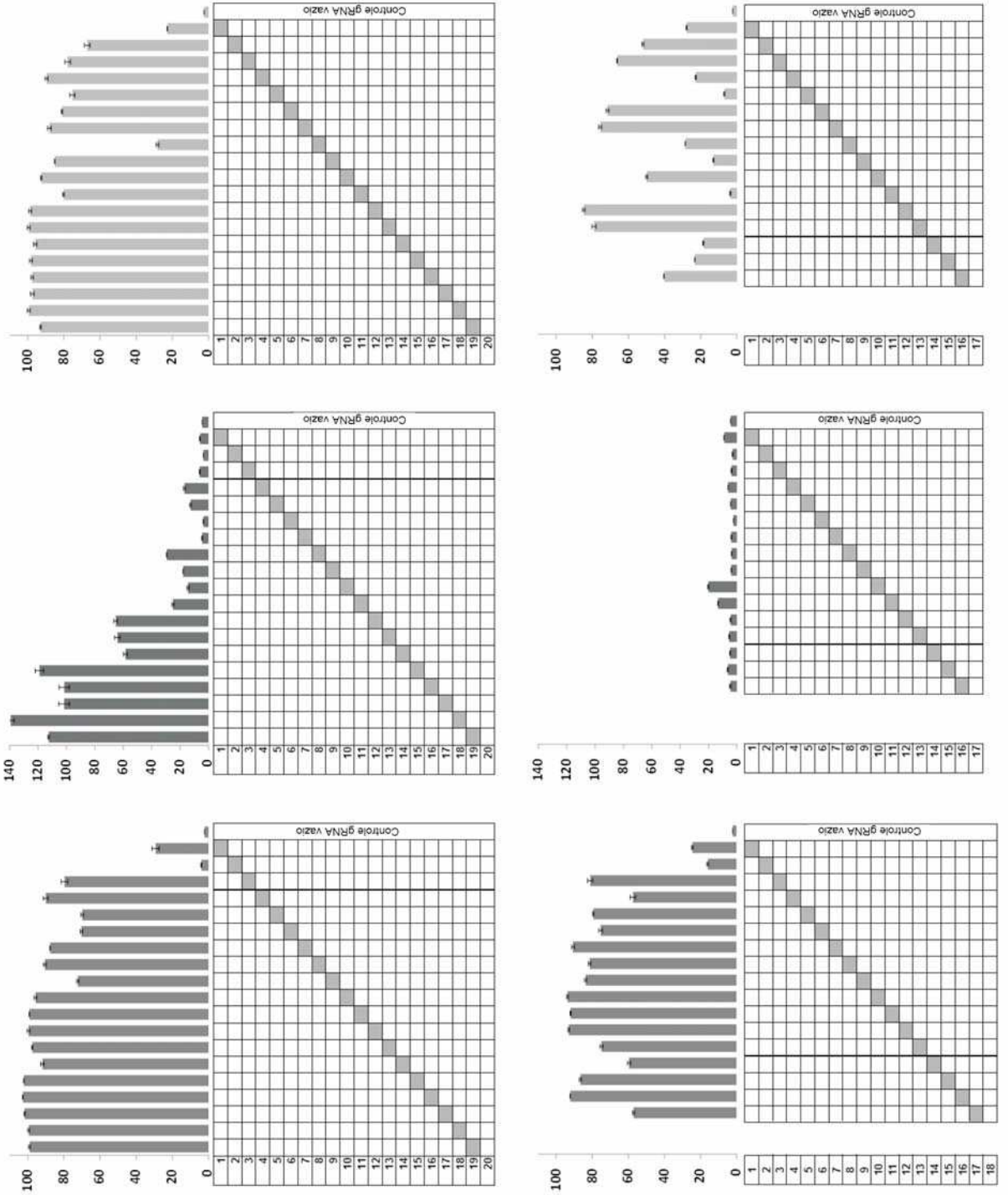
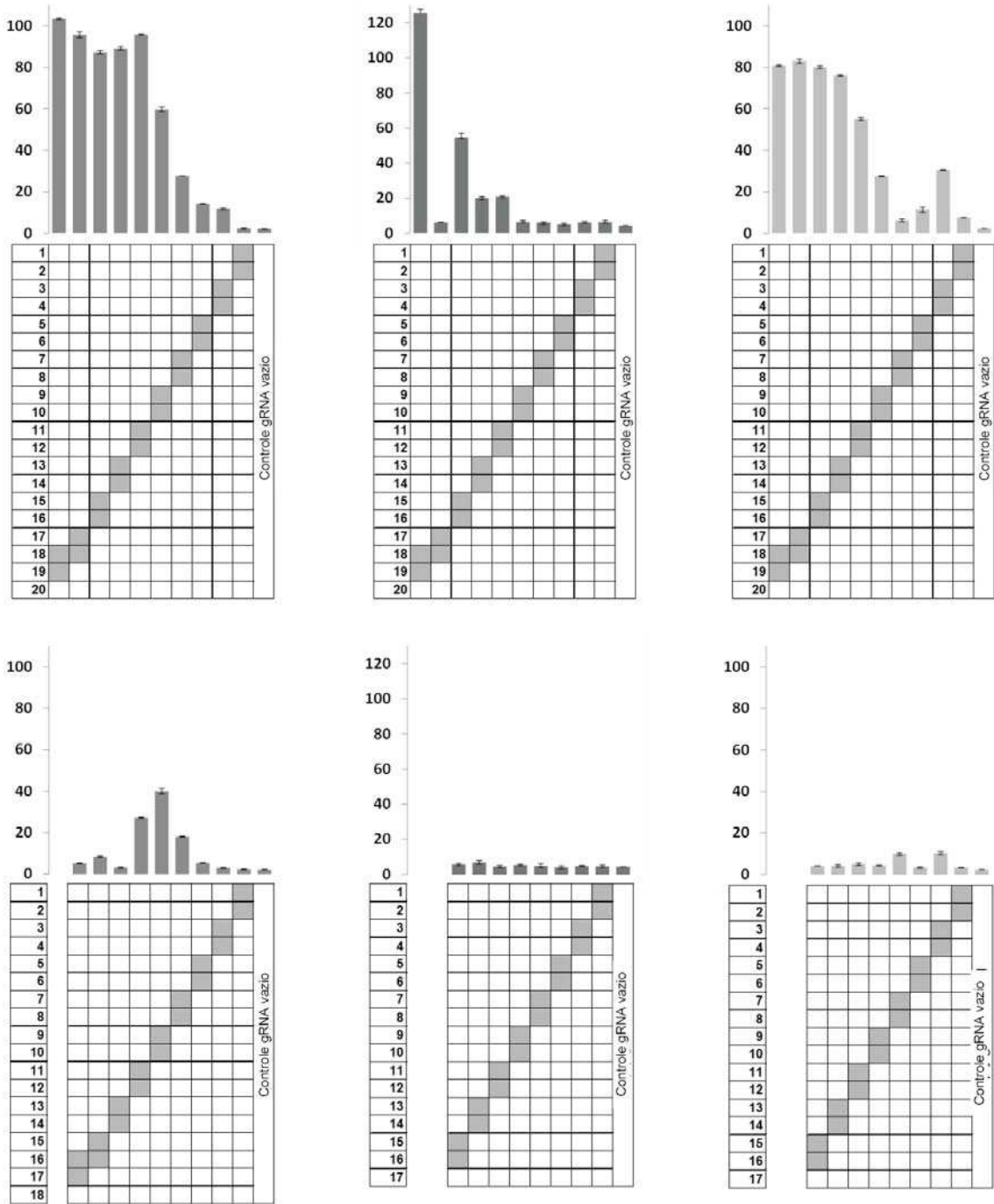


Fig. 5B



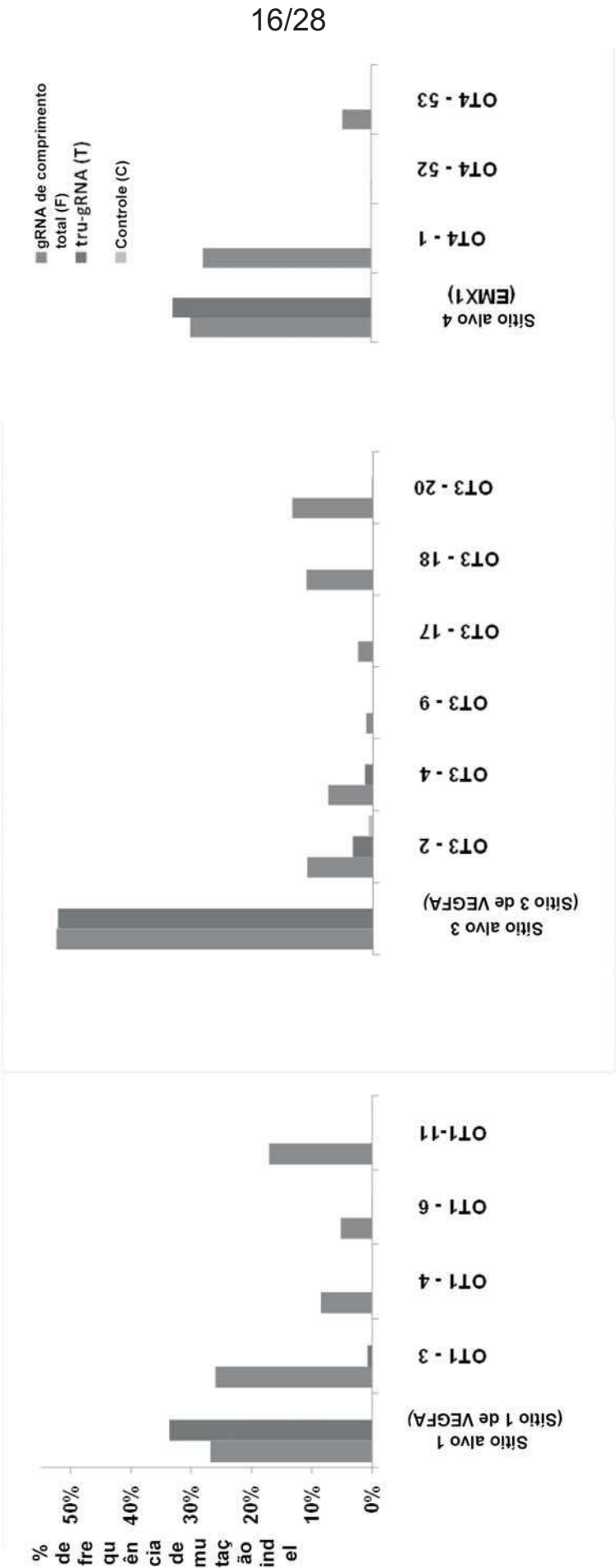


Fig. 6A

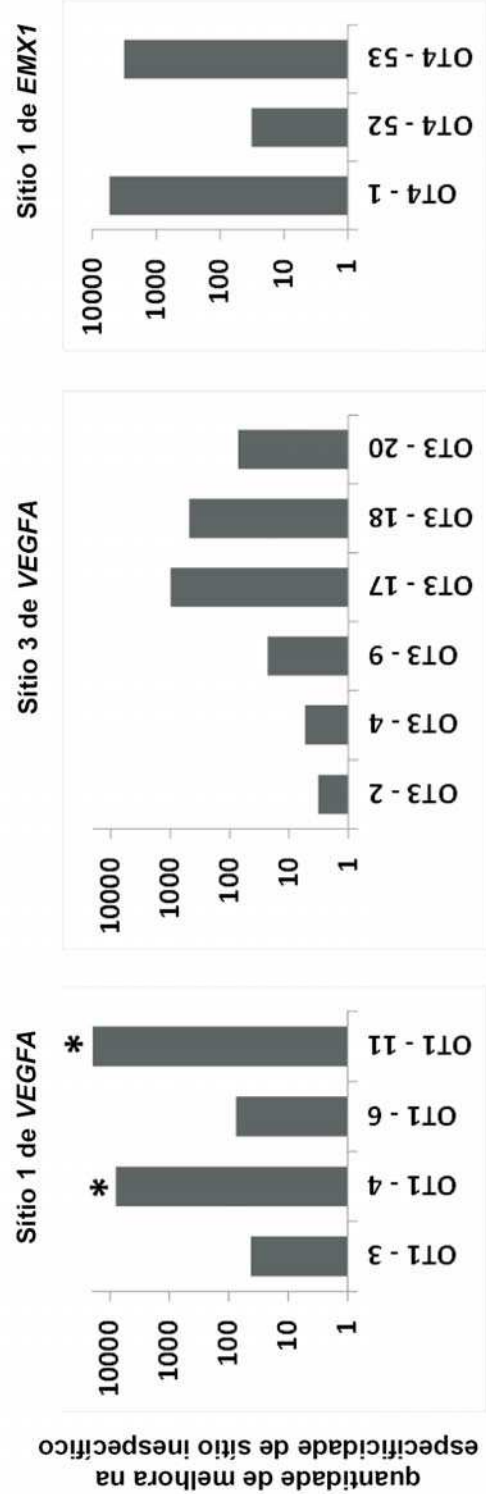


FIG. 6B

Sítio 1 de VEGFA (no alvo):
 Sítio inespecífico:
 Alvo de gRNA de comprimento total:
 Alvo de tru-gRNA:

5' -GGGTGGGGGGAGTTTGCTCCTGG-3'
 5' -TAGTGGAGGGAGCTTGCTCCTGG-3'
 XX X X
 X X

	Alvo	Frequência de mutação indel (%) ± s.e.m.
Full-length gRNA	TAGTGGAGGGAGCTTGCTCCTGG	3.88 ± 0.20
tru-gRNA	GTGGAGGGAGCTTGCTCCTGG	6.88 ± 0.19

FIG. 6C

Figura 7A

gRNA de comprimento total do Sítio1 de VEGFA

TAGCTGTTTGGGAGGTCAGAAATAGGGGGTCCaGGAGCAAAC TCCCCCACCCTTTCCAAAGCCC Tipo
 Selvagem x87

TAGCTGTT-----> Δ205
 <-----CCTTTCCAAAGCCC Δ71
 <-----CCAAAGCCC Δ67
 <-----CCCCACCCCTTTCCAAAGCCC Δ55
 TAGCTG-----TTTCCAAAGCCC Δ49
 TAG-----CCCTTTCCAAAGCCC Δ49
 TAGCTGTTTGGGAGGTCAGA----- Δ47
 TAGCTGTT-----CCCCTTTCCAAAGCCC Δ43
 TAGCTGTTTGGGAGGTCAGAAAT-----AGCCC Δ39
 TAGCTGTTTGGGTGG-----CCTTTCCAAAGCCC Δ38
 TAGCTGTTTGGGAGGTCAGAAATAGG-----CAAAGCCC Δ33
 TAGCTGTTTGGGAGGT-----CACCCCTTTCCAAAGCCC Δ32
 TAGCTGTTTCTGA-----CCTCCCACCCCTTTCCNAAGCCC Δ30
 TAGCTGTTTGGGAGGTCAGAAATAGGGGGTCCAG-----ANCCC Δ28
 TAGCTGTTTGGGAGGTCAGAAAT-----ACCCCTTTCCAAAGCCC Δ26
 TAGCTGTTTGGGAGGTCAGAAATAGGGGGT-----CCCTTTCCAAAGCCC Δ22
 TAGCTGTTTGGGAGGTCAG-----AAACTCCCCCACCCTTTCCAAAGCCC Δ19 x2
 TAGCTGTTTGGGAGGTCAGAAATAGGGGGT-----CCACCCCTTTCCAAAGCCC Δ17
 TAGCTGTTTGGGAGGTCAGAAATAGGGGGT-----CCCACCCCTTTCCAAAGCCC Δ16
 TAGCTGTTTGGGAGGTCAGAAATAGGG-----TCCCCCNCCCCCTTTCCNAANCCC Δ15
 TAGCTGTTTGGGAGGTCAGA-----AGCAAAC TCCCCCACCCTTTCCAAAGCCC Δ15
 TAGCTGTTTGGGAGGTCAGAAAT-----AAACTCCCCCACCCTTTCCAAAGCCC Δ15
 NAGCTGNNTGGGAGGNCNNA-----NNGCAAAC TCCCCCACCCTTTCCAAANCCC Δ14
 TAGCTGTTTGGGAGGTCAGAAATAGGGGGT-----CCCCACCCCTTTCCAAAGCCC Δ14
 TAGCTGTTTGGGAGGTCAGAAATAGGGGGTCCA-----CCCACCCCTTTCCAAAGCCC Δ13
 TAGCTGTTTGGGAGGTCAGAAAT-----AGCAAAC TCCCCCACCCTTTCCAAAGCCC Δ12
 TAGCTGTTTGGGAGGTCAGAAATAGGGGGTCCAGGA-----CCCCACCCCTTTCCAAAGCCC Δ9
 TAGCTGTTTGGGAGGTCAGAAATAGGGGGTCC-----AACTCCCCCACCCTTTCCAAAGCCC Δ7
 TAGCTGTTTGGGAGGTCAGAAATAGGGGGTCC-----CACCCCTTTCCACCCCTTTCCAAAGCCC Δ6
 TAGCTGTTTGGGAGGTCAGAAATAGGGGGTCCAGG-----ACTCCCCCACCCTTTCCAAAGCCC Δ5 x2
 TAGCTGTTTGGGAGGTCAGAAATAGGGGGTCCAGG-----AACTCCCCCACCCTTTCCAAAGCCC Δ4 X2
 TAGCTGTTTGGGAGGTCAGAAATAGGGGGTCC---AGCAAAC TCCCCCACCCTTTCCAAAGCCC Δ3
 TAGCTGTTTGGGAGGTCAGAAATAGGGGGTCCAG--GCAAAC TCCCCCACCCTTTCCAAAGCCC Δ2 X3

TAGCTGTTTGGGAGGTCAGAAATAGGGGGTCCAGGAGTGCTAGATCTTCATCTAAAGACATTTCTGA +191
 GAAAAATGTATCTGTTTTCTTTCAGAAAGAAATTTACACTTAATAGATATTATGGTAAC TAAAGTAAG
 GCAGATAATTTTGGCCATCAGCTTATATTGTGGGATAATCTCTTTTTTGCTGACCTTGAAAAGNTGTG
 GCATATTACAACAAGTAGGAAAATTGCAAAC TCCCCCACCCTTTCCAAAGCCC

Figura 7B

gRNA truncado do SítioI de *VEGFA*

TGGGAGGTCAGAAATAGGGGGTCCaGGAGCAAACCTCCCCCACCCTTTCCAAAGCCCATTCCCTC Tipo
Selvagem x85

TGGGAGGTCAGAAATAGGGGGTCCAGGA-----> Δ144
 <-----CACCTCTTTCCAAAGCCCATTCCCTC Δ112
 <-----CAAACCTCCCCCACCCTTTCCAAAGCCCATTCCCTC Δ111
 <-----AAGCCCATTCCCTC Δ80
 <-----TCCCTC Δ80
 <-----TTCCAAAGCCCATTCCCTC Δ61
 TGGGAGGTC-----> Δ61
 <-----CCCACCCCTTTCCAAAGCCCATTCCCTC Δ56
 <-----CCCACCCCTTTCCAAAGCCCATTCCCTC Δ44
 TGGG-----CCCCTTTCCAAAGCCCATTCCCTC Δ39
 TGGGAG-----CCCCCTTTCCAAAGCCCATTCCCTC Δ36
 TGGGAGGTCAGAAA-----CACCCCTTTCCAAAGCCCATTCCCTC Δ26
 TGGGAGGTCAGAAATAGGGGGT-----CCCCTTTCCAAAGCCCATTCCCTC Δ20
 TGGGAGGTCAGAA-----AAACTCCCCCACCCTTTCCNAAGCCCATTCCCTC Δ17
 TGGGAGGTCAGAAATAGGGGGT-----CCACCCCTTTCCAAAGCCCATTCCCTC Δ17
 TGGGAGGTCAGAAATAGGGGGT-----CCCACCCCTTTCCAAAGCCCATTCCCTC Δ16
 TGGGAGGTCAGAAATAGG-----TCCCCCACCCTTTCCAAAGCCCATTCCCTC Δ16
 TGNAGGTCAGAA-----AGCAAACCTCCCCCACCCTTTCCAAAGCCCATTCCCTC Δ15
 TGGGAGGTCAGAAATAGGGGGT-----CCCCACCCCTTTCCAAAGCCCATTCCCTC Δ15
 TGGGAGGTCAGAAATAGGGGG-----TCCCCCACCCTTTCCAAAGCCCATTCCCTC Δ13 x2
 TGGGAGGTCAGAAATAGGGGGTCCAGGA-----CACCCCTTTCCAAAGCCCATTCCCTC Δ12
 TGGGAGGTCAGAAATAG-----GCAAACCTCCCCCACCCTTTCCAAAGCCCATTCCCTC Δ11 x2
 TGGGAGGTCAGAAATAGGGGGTCCAGGA-----CCCCACCCCTTTCCAAAGCCCATTCCCTC Δ9
 TGGGAGGTCAGAAATAGGGGGTCCAG-----NCCCCACCCCTTTCCAAAGCCCATTCCCTC Δ8
 TGGGAGGTCAGAAATAGGGGGTCCAG-----CTGGCCCCACCCTTTCCAAAGCCCATTCCCTC Δ7
 TGGGAGGTCAGAAATAGGGGGTCC-----CAAACCTCCCCCACCCTTTCCAAAGCCCATTCCCTC Δ6 x5
 TGGGAGGTCANAAATAGGGGGTCCAG-----ACTCCCCCACCCTTTCCAAAGCCCATTCCCTC Δ6
 TGGGAGGTCAGAAATAGGGGGTCCAG-----ACTCCCCCACCCTTTCCAAAGCCCATTCCCTC Δ5
 TGGGAGGTCAGAAATAGGGGGTCC---AGCAAACCTCCCCCACCCTTTNNNNNNCCCATTTNNNTC Δ3
 TGNAGGTCAGAAATAGGGGGTCCAGGA---AAACTCCCCCACCCTTTCCAAAGCCCATTCCCTC Δ2

TGGGAGGTCa-----CCCCCTTTCCAAAGCCCACCCTTTCCAAAGCC Δ9 (+14
 Δ23)
 TGGGAGGTCAGAAATAGGGGGTCCAGGAAGCAAACCTCCCCCACCCTTTCCAAAGCCCATTCCC +2

Figura 7C

gRNA de comprimento total do Sítio3 de *VEGFA*

GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGTGTGCGTGtGGGGTTGAGGGTGTGGAGCGGGGA	Tipo
Selvagem x35	
GAGGACGTGTGTGTCTGTGTG----->	Δ117
GAGGACGTGTGTGTTGG----->	Δ84
GAGGACGTGTGTGTCTGTGTG----->	Δ75
GAGGACGTGTGTGTCTGTGTG----->	Δ49
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGTGTG----->	Δ43
GAGGACGTGTGTGTCTGTGTGGGTGAGTG----->	Δ40
GAGGACGTGTGTGTCTGTGTGAGT-----GGGA	Δ39
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGTNNG----->	Δ37
GAGGACGTGTGTGTCTGTGTGGGTGAGT-----GAGNNGNGN	Δ30 x2
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAG-----TGGGGCGGGGA	Δ25
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTG-----TGGAGCGGGGA	Δ23
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGT-----GTTGGAGCGGGGA	Δ22
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGT-----GTTGGAGCGGGGA	Δ20
GAGGACGTGTGTGTCTGTGTGGGTGAG-----TGAGGGTGTGGAGCGGGGA	Δ20 x2
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGT-----GGGCGTTGGAGCGGGGA	Δ18
GAGGACGTGTGTGTCTGTGTGGGTGA-----NNGTGGGGTTGAGGGTGTGGAGCGGGGA	Δ12
GAGGACGTGTGTGTCTGTGTGGGTGAGTGA-----GTGTGGGGTTGAGGGCGTTGGAGCGGGGA	Δ8 x3
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGTGT-----GGGTTGAGGGCGTTGGAGCGGGGA	Δ7
GAGGACGTGTGTGTCTGTGTGG-TGAGTGAGTGTGT-----GGGGTTGAGGGTGTGGAGCGGGGA	Δ6
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGT-----GNGTGGGGTTGAGGGTGTGGAGCGGGGA	Δ6 x5
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGT----GTGTGGGGTTGAGGGTGTGGAGCGGGGA	Δ4
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGTGTGC---TGGGGTTGAGGGTGTGGAGCGGGGA	Δ3
GAGGANGNGTGTGTCTGTGTGGGTGAGTGAGTGTGTGTGGGTGAGTGAGTGTGTGTCTGTGGGGTTG	+20
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGTGTGCGTCGTGTGGGGTTGAGGGTGTGGAGCGG	+3
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGTGTGCAAGTGTGGGGTTGAGGGTGTGGAGCGG	+3
GAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGTGTGCGTGTGCGGGTTGAGGGTGTGGAGCGGG	+2

Figura 7D

gRNA truncado do Sítio3 de *VEGFA*

```

TGAGGACGTGTGTCTGTGTGGGTGAGTGAAGTGTGTGCGGTtGGGGTTGAGGGTGTGGAGCGGGG Tipo
Selvagem x47
GGGAGGTCAGAAATAGGGGGT-----> Δ324
<-----> Δ201
TGAGGACGTGTGTCTGTGTGGGTGAGTGAAGTGTGTG-----> Δ157
TGAGGACGTGTGTGT-----> Δ153
TGAGGACGTGTGTCTGTGTGGGAGANNGANNGNNG-----> Δ88
TGAGGACGTGTGTG-----> Δ87
TGAGGACGTGTGTGTCTGTGTGG-----> Δ83
TGAGGACGTGTGTGT-----GG Δ50
TGAGGACGTGTGTGTCTGTGTGGGTGAGTG-----> Δ40
TGAGGACGTGTGTGTCTGTGTGGGTGA-----GG Δ38
TGAGGACGTGTGTGTCTGTGTGGGTGAGT-----G Δ37
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGA-----GGAGCGGGG Δ27
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGA-----GTTGGAGCGGGG Δ24
TGAGGACGTGTGTGTCTGTGTGGGTGAGT-----GTGTTGGAGCGGGG Δ24
TGAGGACGTGTGTGTCT-----GTGTGGGGTTGAGGGCGTTGGAGCGGGG Δ22
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGAGT-----GTTGGAGCGGGG Δ22
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGT-----GTTGGAGCGGGG Δ20 X2
TGAGGACGTGTGTGTCTGTGTGGGTGAG-----TGAGGGCGTTGGAGCGGGG Δ20 X2
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGA-----GTTGAGGGTGTGGAGCGGGG Δ15
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGAGT-----GTTGAGGGCGTTGGAGCGGGG Δ13
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGT-----GTTGAGGGCGTTGGAGCGGGG Δ11
TGAGGACGTGTGTGTCTGTGTGGGTGAGT-----NNNGGGGGGNNGANGNGTGNNNNNNGGG Δ9
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGA-----NNNGGGGNTGAGGGTGTGGAGCGGGG Δ9
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGA-----GTGTGGGGTTGAGGGGNTGNNNNNNNGGG Δ8 X3
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGTGT-----GGGTTGAGGGCGTTGGAGCGGGG Δ7
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGAGT-----GTGTGGGGTTGAGGGTGTGGAGCGGGG Δ6 X5
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGTGTGNNGNTGNGGTTGA-----NGGAGNNGGA Δ6
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGT----GTGTGGGGTTGAGGGTGTGGAGCGGGG Δ4
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTG---CGTGTGGGGTTGAGGGTGTGGAGCGGGG Δ4
TGAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGAGT--GTGTGGGGTTGAGGGCGTTGGAGCGGGG Δ2

TGAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGAGTGAGT GNGTGGGGTTGAGGGCGTTGGAGCGG +2

TGAGGACGTGTGTGTCTGTGTGGGTGAGTGAGTGAGTGAGT C TCGTGTGGGGTTGAGGGCGTTGGAGCGG +2

GGGTGAATGGAGCGAGCAGCGTCTTCGNGNGNGAGGACGTGNNGTCTGNGTGNGTNNGNGAGTGTG +234 x2
TGC TAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGG
GTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTANANN
TTAGGGTTNNNNNNNAGGNNNANNGTTANGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGT
TAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGT

```

FIG. 7E

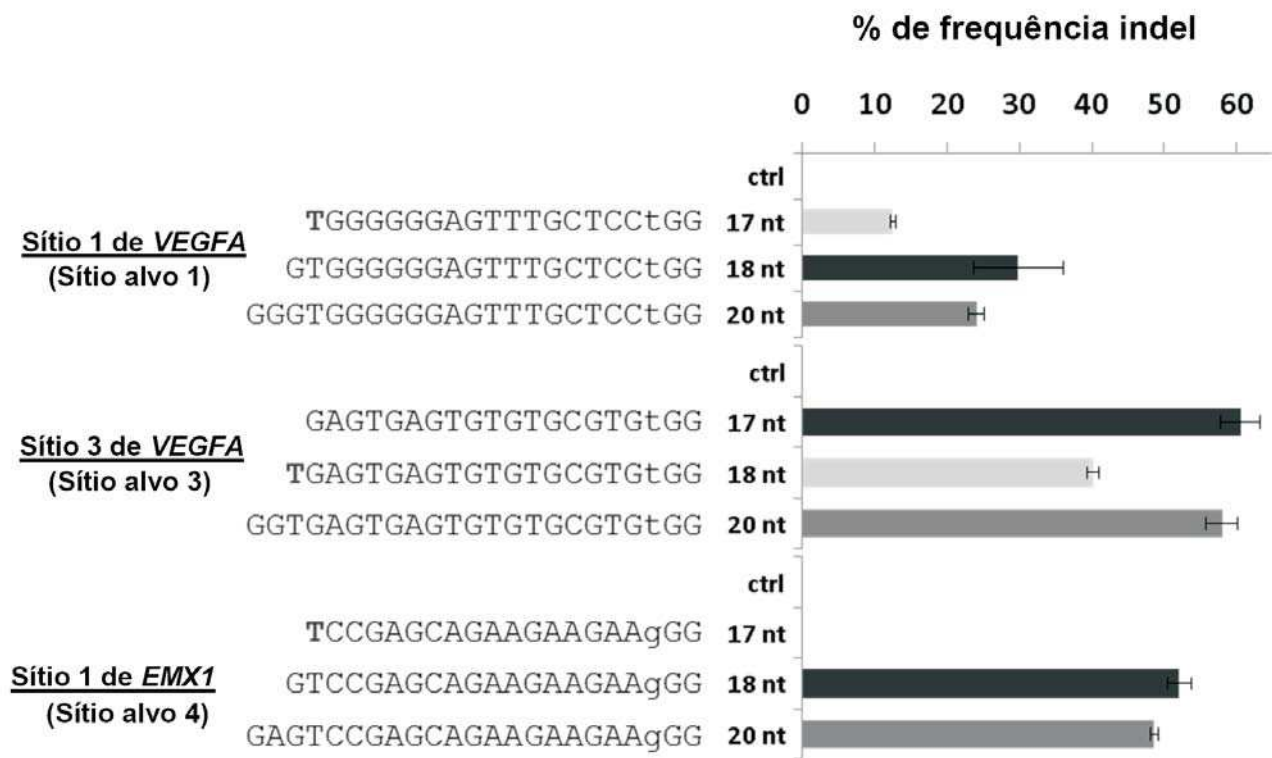


Fig. 8A**Alvo 1 (Sítio 1 de VEGFA):****OT1-3**

AGACAGGACATTCTGACACCCCAGGAGCAAACCTCCCTCCATCCCACAAATCCGTCCTTAGATGTGCA Tipo
Selvagem x18

<-----A AACTCCCTCCATCCCACAAATCCGTCCTTAGATGTGCA Δ53
AGACAGGACATTCTGACACC-----CCATCCCACAAATCCGTCCTTAGATGTGCA Δ17
AGACAGGACATTCTGACACCCCAGGA-----CCCACAAATCCGTCCTTAGATGTGCA Δ15
AGACAGGACATTCTGACACCCCAG--GCAAACCTCCCTCCATCCCACAAATCCGTCCTTAGATGTGCA Δ2

AGACAGGACATTCTGACACCCCAGGAGGGCAAACCTCCCTCCATCCCACAAATCCGTCCTTAGATGTGCA +2

OT1-6

GAGAGAGGCTCCCATCACGGGGGAGGGAGTTTGCTCCTGGGGAACTGTGATCCCCACAGGGAACA Tipo
Selvagem x87

GAGAGAGGCTCCCATCACGGGGG-----AGGGAACA Δ35 x3
GAGAGAGGCTCCCATCACGGGGGA-----GGGGAACCTGTGATCCCCACAGGGAACA Δ14 x1

OT1-11

TGGACTCTACCCACTGAATGCCAGGAGCAAACCTTCCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Tipo
Selvagem x27

TGGACTCTACCCACTG-----AATGTCTC Δ43
TGGACTCTA-----CCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ25
TGGACTCTAC-----CCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ24
TGGACTCTACCCACTGAATG-----CCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ15
TGGACTCTACCCACTGAATG-----CCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ15
TGGACTCTACCCACTGAAT-----GCAAACCTTCCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ7
TGGACTCTACCCACTGAATGCCAGGA-----TTCCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ6
TGGACTCTACCCACTGAATGCCTGG-----CATCCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ6
TGGACTCTACCCACTGAATGCCAGG---AAACTTCCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ3

TGGACTCTACCCACTGAATGCCAGGATGGAAGATAATTTTTTCCATAGACCAGGGTGGGGGAATGGTTTCGGGGAT
GATTCAAGCACATCACATTTATTGTGCACTTTATTTCTATTACTATTATATTGTAATGTATACTAAAAATAATTA
TACAACCTACCATAATGTAGAACCAGTGGGAGCCGCAAACCTTCCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC
A +158

TGGACTCTACCCACTGAATGCCAGGCAAACCTTCCCCTCCCCGAGTTGTGACAGCAAAAATGTCTGGCCTAATGTC
TGGCAAAAATGTCTCAAGACATTGCCAAATGTCCCCT +23 (Δ2 +25)

Fig. 8B**Alvo 2 (Sítio 2 de VEGFA):****OT2-2**

ACCCACCTCCCTATCCTCAAAACTTGGCCAGAGGCGGGGTGGAGGGGGCCCTAGGAGCGCCTTGGTG Tipo
Selvagem x30

ACCCAC-----CGCCTTGGTG Δ51
ACCCACCTCC-----TATCCTAGGAGCGCCTTGGTG Δ36
ACCCACCTCCCTATCCTCAAAACTT-----CCCTTGGTG Δ33
ACCCACCTCCCTATCCTCAAAACTTGGCCAGAG-----TAGGAGCGCCTTGGTG Δ18
ACCCACCTCCCTATCCTCAAAACTT-----GGGGTGGAGGGGGCCCTAGGAGCGCCTTGGTG Δ10
ACCCACCTCCCTATCCTCAAAACTT-----GGCGGGTGGAGGGGGCCCTAGGAGCGCCTTGGTG Δ7
ACCCACCTCCCTATCCTCAAAACTTGGCCAGA-----GTGGAGGGGGCCCTAGGAGCGCCTTGGTG Δ6
ACCCACCTCCCTATCCTCAAAACT-----
CTGGCAGTCTGTGAGTGCCTTATCTTGTACACTTCTACAAGGGGGCTCTCCCTGCATTCTGA +21 (Δ40,
+61)

ACCCACCTCCCTATCCTCAAAACTTGGCCAGAGGGCGCCTCCCCAGGAAGTGCTCCGGCCAGCCCAGGGTAAACA
CGCTAGCCCTGCCCCCTCTGGGACCATAGCCCGGGGACCCAGACTCTTGGCCACGCTCATTCCCACCGCGGGGTGG
AGGGGCCCCCTAGGAGCGCCTTGGTG +108

OT2-15

TGACTGTGCGGTGCCCCACATGTGGCAGATGCCCAGAGGCGGGGTGTGGGGGGTACTTTGTGGGCGTT Tipo
Selvagem X71

TGACTGTC-----GGTGTGGGGGGTACTTTGTGGGCGTT Δ33
TGACTGTGCGGTGCCCCACATGTGGCAGATGCCCAGAG-----GGGCGTT Δ23
TGACTGTGCGGTGCCCCACATGTGGCAGA-----TGTGGGGGGTACTTTGTGGGCGTT Δ15
TGACTGTGCGGTGCCCCACATGT-----GGCGGGGTGTGGGGGGTACTTTGTGGGCGTT Δ14
TGACTGTGCGGTGCCCCACATGTGGCAGATG---GAGGCGGGGTGTGGGGGGTACTTTGTGGGCGTT Δ4
TGACTGTGCGGTGCCCCACATGTGGCAGATGCCCAGA---GGGTGTGGGGGGTACTTTGTGGGCGTT Δ4
TGACTGTGCGGTGCCCCACATGTGGCAGATGCCCAGAG--GGGTGTGGGGGGTACTTTGTGGGCGTT Δ2
TGACTGTGCGGTGCCCCACATGTGGCAGATGCCCAGAGTTGCGGGGTGTGGGGGGTACTTTGTGGGCGTT +2
TGACTGTGCGGTGCCCCACATGTGGCAGATGCCCAGAGCTGCGCGGGGTGTGGGGGGTACTTTGTGGGCGTT +2

OT2-24

ACAAGATGACTATGTCCCTCTGGGCCCCATCCTCCCCCTCCCCACCCACCCCGCCTCAGGCTTGAAGA Tipo
Selvagem x8

ACAAGA-----> Δ121
<-----CTCAGGCTTGAAGA Δ82
ACAAGATG-----> Δ80
<-----> Δ79
ACAAGATGACTATGTCCCTCTGGGCCCCATCCTCC-----> Δ42
ACAAGATGACTATGTCCCTCTGGGC-----CTTGAAGA Δ34
ACAAGATGACTATGTCCCTCTGGGCCCCATCCTCCCCCTCC-----CAGGCTTGAAGA Δ14
ACAAGATGACTATGTCCCTCTGGGCCCCATCCTCCCCCTCC-----CCTCAGGCTTGAAGA Δ12
ACAAGATGACTATGTCCCTCTGGGCCCCATCCTCCCCCTCCCCA-----GCAGGCTTGAAGA Δ11
ACAAGATGACTATGTCCCTCTGGGCCCCATCCTCCCCCTCCCCACCCAC-----CTTGAAGA Δ11
ACAAGATGACTATGTCCCTCTGGGCCCCATCCTCCCCCTCC-----GCTCAGGCTTGAAGA Δ11

ACAAGATGACTATGTCCCTCTGGGCCCCGCTCAAGTGATCCAGCTGCCTTGGCCTCCAAAAGTGCTAGCAGTACA
GATGTGAGCCTCCATGCCTGGCCTATTGCAACATCCCATCTCTGTGAAGCAGGGTTTTCTGCAGTGACAGCAAGA
AGAGCACAGGGCCAAAAAACTTTGTCCTTTAGAAAGGATCTACCTTTTAGGCTGAGAATGGCA +76 (Δ43
+119)

ACCTGTACATCTGCACAAGATTGCCTTTACTCCATGCCTTTCTTCTTCTGCTCTAACTCTGACAATC Tipo
Selvagem x20
-----ATC Δ64
ACCTGTACATCTGCACAAGATTGCCTTTACTCC-----ACAATC Δ28
ACCTGTACATCTGCACAAGATTGCCTTTACTCCAT-----ACTCTGACAATC Δ20
ACCTGTACATCTGCACAAGATTGC-----CTTCTGCTCTAACTCTGACAATC Δ20
ACCTGTACATCTGCACAAGATTGCCTTTACTCCATGCCTTTCT-----CAATC Δ19
ACCTGTACATCTGCACAAGATTGCCTTTACTCCA-----TGCTCTAACTCTGACAATC Δ14
TCCTGTACATCTGCACAAGATTGCCTTTACTCC-----CTTCTTCTGCTCTAACTCTGACAATC Δ8

Fig. 9A

Alvo 1 (Sítio 1 de VEGFA) :**OT1-3**

TCAGACAGGACATTCTGACACCCAGGAGCAAACCTCCCTCCATCCACAAATCCGTCCTTAGATGTG Tipo Selvagem
x41

TCAGACAGGACATT-----CAAACCTCCCTCCATCCACAAATCCGTCCTTAGATGTG Δ15
TCAGACAGGACATTCT-----GAGCAAACCTCCCTCCATCCACAAATCCGTCCTTAGATGTG Δ10
TCAGACAGGACATTCTGAC-----GCAAACCTCCCTCCATCCACAAATCCGTCCTTAGATGTG Δ9
TCAGACAGGACATCCTGACAC-----GCAAACCTCCCTCCATCCACAAATCCGTCCTTAGATGTG Δ7
TCAGACAGGACATTCTGACACCCAG--GCAAACCTCCCTCCATCCACAAATCCGTCCTTAGATGTG Δ2 x6

TCAGACAGGACATTCTGACACCCAGGATGTCCTCCTCCCTCCATCCACAAATCCGTCCTTAGATGTG +2

TCAGNACAGGACATTNNGNCACCCAGGAAACNNGAGTTTTCGNTNCNNGANNGTCAGACCCAGNAGCAAACCTCCCTCCATC
CCACAAATCCGTCCTTAGATGTG +38

TCAGACAGGACATTNTGACACCCAGGAGTNTGCACNTCAGTTTTCTTTANTATGTNGNNNNGGGGCANGNACAAANNTTNN
GCAAACCTCCCTCCATCCACAAATCCGTCCTTAGATGTG +54

TCAGACAGGACATTCTGACACCCAGGATGTTTGTGTTGAGTCAGAGTCTCTCTTTTGTACCCAGGCTGGAGTGCACTGGAA
ACCTGTGCCTTTTGTATATCCTCTTTGAAGGTTAAAGAGTCATCATGGATCANCNNCATAAAGCAAACCTCCNTCCATCC
+116

TCAGACAGGACATTCTGACACCCAGGATAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACG
CCTGGTATCTTTATAGTCCTGTGCGGGTTTCGCCACCTCTAGGGGAGTTCTGAGCTCTGACGCCGGGCGTGTAGGAGATAGC
AGGCCGTTAATGACCATCCAGCCGAATTCCTCACTGTGCAGATGAGGAAGTGAGCTCAGGGAGGCTGAGTGTCCTCAGGCCT
GTTGCCAGATGAGGCCACGCTGAGACTGTGCAAACCTCCCTCCATCCACAAATCCGTCCTTAGATGTG +247

OT1-6

TGGAGAGAGGCTCCCATCACGGGGGAGGGAGTTTGCTCCTGGGAACCTGTGATCCCCACAGGGAAC Tipo Selvagem
x88

TGGAGAGAGGCTCCCATCACGGGGGAGGGAGTTTG-----CCTGTGATCCCCACAGGGAAC Δ11
TGGAGAGAGGCTCCCATCACGGGGGAGGGAGTTT-----GGGAACCTGTGATCCCCACAGGGAAC Δ7

OT1-11

AGCATCGCTGGACTCTACCCACTGAATGCCAGGAGCAAACCTCCCTCCCGAGTTGTGACAGCAAA Tipo Selvagem
x84

AGCATCGCTGGACTCTACCCACTGAATGCCAGGA-----CCCGAGTTGTGACAGCAAA Δ14
AGCATCGCTGGACTCTACCCACTGA-----GCAAACCTCCCTCCCGAGTTGTGACAGCAAA Δ9
AGCATCGCTGGACTCTACCCACTGAATGCCAG--GCAAACCTCCCTCCCGAGTTGTGACAGCAAA Δ2

AGCATCGCTGGACTCTACCCACTGAATGCCAGGAGTTTCAGACGATTGAATGTATCAACTTGGCACATTGCCTATCAACTGGT
GAGTGCTCAAAAATATCCATTGCTGTGATCAGTAATGCCACAGGGTGACCATTTAAGGACAGAGTCCATGTTTTATCCATCC
TTAGCAAACCTCCCTCCCGAGTTGTGACAGCAAA +133

AGCATCGCTGGACTCTACCCACTGAATGCCAGAGCCCTTCCTTCTCCCTCTCTCTCCAGAGGTCCTGCCGAGATCAGGTT
GGAGGTCCTCTTTGTTCTTATGCCCATTCTCCCGAGGCACTTGGAGGAGGGCACTGTTTTTGTGAGTGTGCAAGTCTTTCTC
TGTTACTGTTGGGCAAACCTCCCTCCCGAGTTGTGACAGCAAA +142

AGCATCGCTGGACTCTACCCACTGAATGCCAGGATGTTTTGTTTGGCAGCGANTCTCACTCTGTGCGNCCGGGCTGGAGTGCA
NNGGCACANTTCTCANCTGACTGCNATGTCCGCCTCCCGATTCAAGTGATTCTCCTGCCCGAGCCTCCCGAGTAGCTGGG
ATTATAGGTGCCTGCCACCATGCCTGGCTAATTTTTTTTTTTTTTAAATGGAGTCTCACTCTGTTGCCCCGAGTTGTGA
CAGCAAA +186

Fig. 9B**Alvo 2 (Sítio 2 de VEGFA):****OT2-2**

CACCTCCCTATCCTCAAAACTTGGCCAGAGGCGGGGTGGAGGGGCCCTAGGAGCGCCTTGGTGGGA
 CACCTCCCTATCCTCAAAACTTGGCCAGAGGCGGGGT-----GGAGCGCCTTGGTGGGA Δ13
 CACCTCCCTATCCTCAAAACTT-----GGCGGGGTGGAGGGGCCCTAGGAGCGCCTTGGTGGGA Δ7
 CACCTCCCTATCCTCAAAACTTG-----GACGGGGTGGAGGGGCCCTAGGAGCGCCTTGGTGGGA Δ6
 CACCTCCCTATCCTCAAAACTTGGCC--AGGCGGGGTGGAGGGGCCCTAGGAGCGCCTTGGTGGGA Δ2

CACCTCCCTATCCTCAAAACTTGGCCAGAGACTTANACCTAANACCTCAAACTATGAGACTGCTACNAGAGAACA
 TCANAAAACTTTCCAGGACATTCTTCTGGNNGGGTGGAGGGGCCCTANGAGCGCCTTGNNGGGA +74

CACCTCCCTATCCTCAAAACTTGGCCAGAGTCTACAGATTTATAAAATATTACCAGTTAATCATGACACATATTG
 TTTATTTTCAAATATTTTCTAGTTAAACCCACCATTTATATAACCAATTATATTTGATATTATTTAAAAATTT
 TGTATTAACACCCACCAAATCATTTTACAGCGGGGTGGAGGGGCCCTAGGAGCGCCTTGGTGGGA +150

OT2-15

GTCGGTGCCCCACATGTGGCAGATGCCAGAGGCGGGGTGTGGGGGTACTTTGTGGGCGTTTTGGG Tipo
 Selvagem x79
 GTCGGTGCCCCAC-----GCGGGGTGTGGGGGTACTTTGTGGGCGTTTTGGG Δ19
 GTCGGTGCCCCACATGTGGC-----GCGGGGTGTGGGGGTACTTTGTGGGCGTTTTGGG Δ12
 GTCGGTGCCCCACATGTGGCAGATGCCAGA-----GGTGTGGGGGTACTTTGTGGGCGTTTTGGG Δ5
 GTCGGTGCCCCACATGTGGCAGATGCCA-----GGGTGTGGGGGTACTTTGTGGGCGTTTTGGG Δ5
 GTCGGTGCCCCACATGTGGCAGATGCCAGA-----GGGTGTGGGGGTACTTTGTGGGCGTTTTGGG Δ4
 GTCGGTGCCCCACATGTGGCAGATGCC-----GCGGGGTGTGGGGGTACTTTGTGGGCGTTTTGGG Δ4
 GTCGGTGCCCCACATGTGGCAGATGCC--AGGCGGGGTGTGGGGGTACTTTGTGGGCGTTTTGGG Δ2 x2

OT2-24

GTCCCTCTGGGCCCCATCCTCCCCTCCCCACCCACCCCGCCTCAGGCTTGAAGAGGAAAGAAGAGCA
 GTCCCTCTGGGCCCC-----ANAGNANNNANNANN Δ36
 GTCCCTCTGGGCCCCATCCTCCCCTCCC-----TGAAGAGGAAAGAAGAGCG Δ19
 GTCCCTCTNNNCNNCNT-----CCNCNCCTCAGGCTTGAAGAGGAAAGAAGAGCG Δ17
 GTCCCTCTGGGCCCCATCCTCCCCTC-----CCTCAGGCTTGAAGAGGAAAGAAGAGCG Δ13
 GTCCCTCTGGGCCCCATCCTCCCCTCCCCA-----CTCAGGCTTGAAGAGGAAAGAAGAGCG Δ10
 GTCCCTCNNNNNNNCCTCCTCCC-----NNCCNCNCTCAGGCTTGAAGAGGAAAGAAGAGCG Δ8
 GTCCCTCTGGGCCCCATCCTCCCCTCCCCACCCA-----CTCAGGCTTGAAGAGGAAAGAAGAGCG Δ6
 GTCCCTCTGGGCCCCATCCTCCCCTCCCCACCCA-----CCTCAGGCTTGAAGAGGAAAGAAGAGCG Δ5

GTCCCTCTGGGCCCCATCCTCCCCTCCCCACCCACCCCGCATCGTACGTGTCCTTACTAAGCTGCAATTTGGCAT
 CTCAGCTAAGTCGAAGTTCGACCTCAGGCTTGAAGAGGAAAGAAGAGCG +58

Fig. 9C

Alvo 3 (Sítio 3 de VEGFA):

OT3-2*

[illegible]

OT3-9

TGGAGGTGTTGGGATGCGGGAGTGGGTGAGTGAGTGCGTGCGGGTGGCGATGCAAGCGTGTGCGAAT Tipo
Selvagem x101
TGGAGGTGTTGGGATGCGGGAGTGG-----GTGCGTGCGGGTGGCGATGCAAGCGTGTGCGAAT Δ8
TGGAGGTGTTGGGATGCGGGAGTGGGTGA-----GTGCGGGTGGCGATGCAAGCGTGTGCGAAT Δ8

OT3-18

CAAAGACAGTAGATCTTAAATGTCCTCACGCACACACTCACCCACACATAAAAGGTGGTAACTGTGT Tipo
 Selvagem x64
 CAAAGACAGTAGATCTTAA-----GCACACACTCACCCACACATAAAAGGTGGTAACTGTGT Δ10
 CAAAGACAGTAGATCTTAAATGTCCTCAG-----TCACCCACACATAAAAGGTGGTAACTGTGT Δ7
 CAAAGACAGTAGATCTTAAATGTC-----GCACACACTCACCCACACATAAAAGGTGGTAACTGTGT Δ5
 CAAAGACAGTAGATCTTAAATGTCCTCACGCCGACNATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGC
 TGCTTCGCGATGTACGGGCCAGATAGCACACACTCACCCACACATAAAAGGTGGTAACTGTGT +71

CAAAGACAGTAGATCTTAAATGTCCTCACGCAAATTTTATTTTGGTTCATGATATGGCTTGGCGTGTATGCTTTT
CATTGTGAAAATTGCTGTTCTTTTGACAATTTAAGTGACTGTTTCATTGACTACAAGTTTGAAAATAAAAATTAA
TTAAGAAAAAAATTCCAATGACTGTGCTGTGGTTGGGCACACACTCACCCACACATAAAAGGTGGTAACTGTGT
+157

CAAAGACAGTAGATCTTAAATGTCCTCACGTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAAC
GATTTATCACCTCANAAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTAACTACTTGGCTCTTGC
CCATATGATAAAGTTCCGTGGGCACCTTCTCATTGAGGGTGATCTAAATCCGGACAACCTCGGATGTCGAGCACAC
ACTCACCCACACATAAAAGGTGGTAACTGTGT +190
CAAAGACAGTAGATCTTAAATGTCCTCACG-----
TTATTTAGAGACAGAGTCTCACTCTGTTGCCAGGCTGGGGTGAGTGGTACGAACTCGGCTCACTGCAACCTCC
GTCTCCTGGGCTCAAGTGATTATCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGTGCCACCACCACACCC
GGCTAATTTTTGTATTTTCAGTAGAGCTGGGGTTTCACCATGTTGGCCAGCCTGTTCTCGGCACACACTCACCCA
CACATAAAAGGTGGTAACTGTGT +211 ($\Delta 16$ +227)

Alvo 4 (*EMX1*):

OT4-1

GATTGCCTTTACTCCATGCCTTTCTTCTTGCTCTAACTCTGACAATCTGTCTTGCCATGCCATAA Tipo
Selvagem x74

GATTGCCTTTACTC-----CTTCTTCTTGCTCTAACTCTGACAATCTGTCTTGCCATGCCATAA Δ9
GATTGCCTTTACTCCATGCCT-----TTCTTGCTCTAACTCTGACAATCTGTCTTGCCATGCCATAA Δ6 x2
GATTGCCTTTACTCCATGC-----TCTTCTTGCTCTAACTCTGACAATCTGTCTTGCCATGCCATAA Δ6
GATTGCCTTTACTCCATGCCT---TTCTTCTTGCTCTAACTCTGACAATCTGTCTTGCCATGCCATAA Δ3 x3