



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}s 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° 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)GUUUUAGAGCUAUGC (SEQ ID NO: 2408);

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

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUAAAAUAAG GCUAGUCCGUUAUC(X_N) (SEQ ID NO: 2);

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCUGUUUUGGAAACAAACAGCAUAGC AAGUUAAAAUAAGGCUAGUCCGUUAUC(X_N) (SEQ ID NO: 3);
 (X₁₇₋₁₈)GUUUUAGAGCUAGAAAAGCAAGUUAAAAUAAG GCUAGUCCGUUAUC(X_N) (SEQ ID NO: 4),

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAGAAAAGCAAGUUAAAAUAAGGCUAGUC GUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID NO: 5);

(X₁₇₋₁₈ ou

$X_{17-19})$ GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAU
AAGGCUAGUCGUUAUCAACUUGAAAAAGUGGCACCGGUCGGU
GC (SEQ ID NO: 6); ou

(X_{17-18} ou

$X_{17-19})$ GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAU
AAGGCUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG
UGC (SEQ ID NO: 7);

em que X_{17-18} ou X_{17-19} 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, 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 U, 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. 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 (**Protospacer Adjacent Motif - 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) 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 X₁₇₋₁₈ ou X₁₇₋₁₉. Em algumas modalidades, alguns ou todos os tracrRNA ou crRNA, por exemplo, dentro ou fora da região de complementaridade ao alvo X₁₇₋₁₈ ou X₁₇₋₁₉, 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 (**Protospacer Adjacent Motif - PAM**), por exemplo, NGG, NAG ou NNGG, por exemplo, um ácido ribonucleico conforme descrito aqui. 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, 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)GUUUUAGAGCUAUGC (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

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

UAGCAAGGUAAAAAUAGGCUAGGUCCGUUAUCAACUUGAAAAAGU
GGCACCGAGUCGGUGC (SEQ ID NO: 2405); ou o cRNA é (X₁₇₋₁₈ ou X₁₇₋₁₉) GUUUUAGAGCUAUGC (SEQ ID NO: 2408) e o tracrRNA é
AGCAUAGCAAGGUAAAAAUAGGCUAGGUCCGUUAUCAACUUGAAA
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, 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 (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 X₁₇₋₁₈ ou X₁₇₋₁₉. Em algumas modalidades, alguns ou todos dos tracrRNA ou crRNA, por exemplo, dentro ou fora da região de complementaridade ao alvo X₁₇₋₁₈ ou X₁₇₋₁₉, 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 transcracional, por exemplo, um domínio de ativação transcracional 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 GATGCCGTTCTCTGCTTGTGG
 (SEQ ID NO: 10)

Sítio 3 de EGFP GGTGGTGCAGATGAACCTTCAGGG
 (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 a 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: Sequencias 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 *BamHI* 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: Comparaçao 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) àquelas 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 (qualsquer) 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 (nímeros 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 (%) ± SEM	12,5 ng de gRNA / 50 ng de Cas9 Frequência média de indel (%) ± SEM
T2 (sobre o alvo)	GACCCCCCTCCACCCCGCCTCCGG	12	50,2 ± 4,9	25,4 ± 4,8
OT2-1	GACCCCC<u>C</u>CCACCCCG<u>CC</u>CGG	13	14,4 ± 3,4	4,2 ± 0,2
OT2-2	G<u>GG</u>CCCCCTCCACCCCGCCTCTGG	14	20,0 ± 6,2	9,8 ± 1,1
OT2-6	CTA CCCCCTCCACCCCGCCTCCGG	15	8,2 ± 1,4	6,0 ± 0,5
OT2-9	GCCCCC<u>A</u>CCCACCCCGCCTCTGG	16	50,7 ± 5,6	16,4 ± 2,1
OT2-15	TACCCCC<u>C</u>ACACCCCGCCTCTGG	17	9,7 ± 4,5	2,1 ± 0,0
OT2-17	AC<u>A</u>CCCC<u>CC</u>ACACCCCGCCTCAGG	18	14,0 ± 2,8	7,1 ± 0,0
OT2-19	AT<u>T</u>CCCC<u>CC</u>ACACCCCGCCTCAGG	19	17,0 ± 3,3	9,2 ± 0,4
OT2-20	CCCC<u>A</u>CC<u>CC</u>ACACCCCGCCTCAGG	20	6,1 ± 1,3	N.D.
OT2-23	CG<u>CC</u>CT<u>CC</u>ACACCCCGCCTCCGG	21	44,4 ± 6,7	35,1 ± 1,8
OT2-24	CT<u>CCCC</u><u>A</u>CCCACCCCGCCTCAGG	22	62,8 ± 5,0	44,1 ± 4,5
OT2-29	T<u>G</u>CCCC<u>T</u>CCCACACCCCGCCTCTGG	23	13,8 ± 5,2	5,0 ± 0,2
OT2-34	AG<u>GG</u>CCCC<u>C</u>ACACCCCGCCTCAGG	24	2,8 ± 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 (%) ± SEM	12,5 ng de gRNA / 50 ng de Cas9 Frequência média de indel (%) ± SEM
T3 (sobre o alvo)	GGTGAGTGAGTGTGCGTGTGG	25	49,4 ± 3,8	33,0 ± 3,7
OT3-1	GGTGAGTGAGTGTG <u>T</u> GAGG	26	7,4 ± 3,4	N.D.
OT3-2	A GTGAGTGAGTGTG <u>T</u> GTGGGG	27	24,3 ± 9,2	9,8 ± 4,2
OT3-4	G C TGAGTGAGTGT <u>A</u> TGCGTGTGG	28	20,9 ± 11,8	4,2 ± 1,2
OT3-9	GGTGAGTGAGTGC <u>G</u> TGCG <u>G</u> TGG	29	3,2 ± 0,3	N.D.
OT3-17	G TGAGTG <u>A</u> TGTGTGCGTGAGG	30	2,9 ± 0,2	N.D.
OT3-18	T GT <u>G</u> GTGAGTG <u>T</u> GCGTGAGG	31	13,4 ± 4,2	4,9 ± 0,0
OT3-20	A GAGTGAGTGTG <u>C</u> ATGAGG	32	16,7 ± 3,5	7,9 ± 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 contraintuitiva 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 comprehende um crRNA fundido a um tracrRNA normalmente transcodificado, 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 (**Protospacer Adjacent Motif - 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)GUUUUAGAGCUAUGC (SEQ ID NO: 2408); (X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAGAAAAGCAAGGUAAAAUAAGGCUAGUC

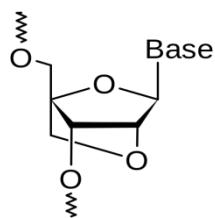
CG(X_N) (SEQ ID NO: 1);
 (X₁₇₋₁₈) ou
 X₁₇₋₁₉)GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUAAAAAUAA
 GGUAGUCCGUUAUC(X_N) (SEQ ID NO: 2);
 (X₁₇₋₁₈) ou
 X₁₇₋₁₉)GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGC
 AAGUUAAAAAUAGGCUAGUCCGUUAUC(X_N) (SEQ ID NO: 3); (X₁₇₋₁₈)
 ou
 X₁₇₋₁₉)GUUUUAGAGCUAGAAAAGCAAGUUAAAAAUAGGCUAGUC
 CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(X_N) (SEQ
 ID NO: 4),
 (X₁₇₋₁₈) ou
 X₁₇₋₁₉)GUUUUAGAGCUAGAAAAGCAAGUUUUAAAAGGCUAGUC
 CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID
 NO: 5);
 (X₁₇₋₁₈) ou
 X₁₇₋₁₉)GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUUAAAU
 AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG
 UGC (SEQ ID NO: 6); ou (X₁₇₋₁₈) ou
 X₁₇₋₁₉)GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUUAAAU
 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)GUUUUAGAGCUAUGCUC(X_N) (SEQ ID NO: 2408);

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAGAAAAGCAAGUUAAAAUAAGGCUAGUC

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

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUAAAAUAAGGCUAGUC

GGCUAGUCCGUUAUC(X_N) (SEQ ID NO: 2);

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCUGUUUUGGAAACAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUC

AAGUUAAAAUAAGGCUAGUCCGUUAUC(X_N) (SEQ ID NO: 3);

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAGAAAAGCAAGUUAAAAUAAGGCUAGUC

GCUAGUCCGUUAUCACUUGAAAAAGUGGCACCGAGUCGGUGC(X_N) (SEQ ID NO: 4),

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAGAAAAGCAAGUUAAAAUAAGGCUAGUC

CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID NO:5);

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 6); ou

(X₁₇₋₁₈ ou X₁₇₋₁₉)GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (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 (**Protospacer Adjacent Motif - PAM**), por exemplo, NGG, NAG ou NNNG, 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)GUUUUAGAGCUAUGC (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: UAGCAAGUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 2405) ou uma porção ativa da mesma; AGCAUAGCAAGUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGC (SEQ ID NO: 2407) ou uma porção

ativa da mesma; CAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 2409) ou uma porção ativa da mesma; UAGCAAGUUAAAAUAAGGCUAGUCGUUAUCAACUUGAAAAAGU G (SEQ ID NO: 2410) ou uma porção ativa da mesma; UAGCAAGUUAAAAUAAGGCUAGUCGUUAUCA (SEQ ID NO: 2411) ou uma porção ativa da mesma; ou UAGCAAGUUAAAAUAAGGCUAGUCCG (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: GGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCC 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: UAGCAAGUUAAAAUAAGGCUAGUCGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC (SEQ ID NO: 2405) ou uma porção ativa da mesma. Em algumas modalidades em que (X_{17-18} ou X_{17-19}) GUUUUAGAGCUAUGC (SEQ ID NO: 2408) é usada como um crRNA, o seguinte tracrRNA é usado: AGCAUAGCAAGUUAAAAUAAGGCUAGUCGUUAUCAACUUGAAA 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)GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUC

CG(X_N) (SEQ ID NO:1);
 (X₁₇₋₁₈ ou
 X₁₇₋₁₉)GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUAAAAAUAA
 GGUAGUCGUUAUC(X_N) (SEQ ID NO:2);
 (X₁₇₋₁₈ ou
 X₁₇₋₁₉)GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGC
 AAGUUAAAAUAAGGCUAGGUCCGUUAUC(X_N) (SEQ ID NO:3);
 (X₁₇₋₁₈ ou
 X₁₇₋₁₉)GUUUUAGAGCUAGAAAAGCAAGUUAAAAUAAGGCUAGUC
 CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(X_N) (SEQ
 ID NO:4),
 (X₁₇₋₁₈ ou
 X₁₇₋₁₉)GUUUUAGAGCUAGAAAAGCAAGUUUUAAAAGGUAGUC
 CGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID
 NO:5);
 (X₁₇₋₁₈ ou
 X₁₇₋₁₉)GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUUAAAU
 AAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG
 UGC (SEQ ID NO:6); ou
 (X₁₇₋₁₈ ou
 X₁₇₋₁₉)GUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUUAAAU
 AAGGUAGUCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG
 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**rotospacer **A**djacent **M**otif - PAM), por exemplo, NGG, NAG ou NNNG, e ainda em que um ou mais dos nucleotídeos são desoxirribonucleotídeos, 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 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, 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</i> subsp. <i>vincentii</i>
374307738	<i>Filifactor alocis</i> ATCC 35896

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

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. Buddy
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 DSM 1740</i>
218563121	<i>Campylobacter jejuni subsp. jejuni</i>
291276265	<i>Helicobacter mustelae 12198</i>
229113166	<i>Bacillus cereus Rock1-15</i>
222109285	<i>Acidovorax ebreus TPSY</i>
189485225	<i>Cupim, grupo 1, não cultivado</i>
182624245	<i>Clostridium perfringens D str.</i>
220930482	<i>Clostridium cellulolyticum H10</i>
154250555	<i>Parvibaculum lavamentivorans DS-1</i>
257413184	<i>Roseburia intestinalis L1-82</i>
218767588	<i>Neisseria meningitidis Z2491</i>
15602992	<i>Pasteurella multocida subsp. multocida</i>
319941583	<i>Sutterella wadsworthensis 3 1</i>
254447899	<i>gama proteobactéria HTCC5015</i>
54296138	<i>Legionella pneumophila str. Paris</i>
331001027	<i>Parasutterella exrementihominis YIT 11859</i>
34557932	<i>Wolinella succinogenes DSM 1740</i>
118497352	<i>Francisella novicida U112</i>

[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
MDKKYSIGA IGTNSVGWAV ITDEYKVPSK KFKVLGNTDR HSIKKNLIGA LLFDSGETAE

70 80 90 100 110 120
ATRLKRTARR RYTRRKNRIC YLQEIFSNEM AKVDDSSFFHR LEESFLVEED KKHERHPIFG

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

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

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

310 320 330 340 350 360

```

LLSDILRVNT EITKAPLSAS MIKRYDEHHQ DLTLKALVR QQLPEKYKEI FFDQSKNGYA

370 380 390 400 410 420

GYIDGGASQE EFYKFIKPIL EKMDGTEELL VKLNREDLLR KQRTFDNGSI PHQIHLGELH

430 440 450 460 470 480

AILRRQEDFY PFLKDNREKI EKILTFRIPY YVGPLARGNS RFAWMTRKSE ETITPWNFEE

490 500 510 520 530 540

VVDKGASAQS FIERMTNFDK NLPNEKVLPK HSLLYEYFTV YNELTKVKYV TEGMRKPAFL

550 560 570 580 590 600

SGEQKKAIVD LLFKTNRKVT VKQLKEDYFK KIECFDSVEI SGVEDRFNAS LGTYHDLLKI

610 620 630 640 650 660

IKDKDFLDNE ENEDILEDIV LTTLTFEDRE MIEERLKTYA HLFDDKVMKQ LKRRRTGWG

670 680 690 700 710 720

RLSRKLINGI RDQSGKTIL DFLKSDGFAN RNFMQLIHDD SLTFKEDIQK AQVSGQGDSDL

730 740 750 760 770 780

HEHIANLAGS PAIKKGILQT VKVVDELVKV MGRHKPENIV IEMARENQTT QKGQKNSRER

790 800 810 820 830 840

MKRIEEGIKE LGSQILKEHP VENTQLQNEK LYLYYLQNGR DMYVDQELDI NRLSDYDVDA

850 860 870 880 890 900

IVPQSFLKDD SIDNKVLTRS DKNRGKSDNV PSEEVVKKMK NYWRQLLNAK LITQRKFDSL

910 920 930 940 950 960

TKAERGGLSE LDKAGFIKRQ LVETRQITKH VAQILDSRMN TKYDENDKLI REVKVITLKS

970 980 990 1000 1010 1020

KLVSDFRKDF QFYKVREINN YHHAHDAYLN AVVGTALIKK YPKLESEFVY GDYKVYDVRK

1030 1040 1050 1060 1070 1080

MIAKSEQEIG KATAKYFFYS NIMNFFKTEI TLANGEIRKR PLIETNGETG EIWWDKGRDF

1090 1100 1110 1120 1130 1140

ATVRKVLSMP QVNIVKKTEV QTGGFSKESI LPKRNSDKLI ARKKDWDPKK YGGFDSPTVA

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

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

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

1330 1340 1350 1360
 PAAFKYFDTT IDRKRYTSTK 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 transcrecional, por

exemplo, a partir de VP64 ou NF-κB 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 transpcionais (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:

Nos. Acesso GenBank		
Gene	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.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, 3^a 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 transfetadas, tais como timidina quinase, higromicina B fosfotransferase e di-hidrofolato

reduzir. 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).

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	1	1	1	1	1	1	1	1	1	oligonucleotídeo 1	#	oligonucleotídeo 2	#
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	(5' a 3')		(5' a 3')	
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGGGCACGG	36	AAAACCCGCAAG	23
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	c	GCAGCTTCCCGG		CTGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	c	ACACCGGGCACGG	37	AAAACCCGCAAG	23
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	c	G	GCAGCTTCCCGG		CTGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	c	G	ACACCGGGCACGG	38	AAAACCCGCAAG	23
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	c	G	GCAGCTTCCCGG		CTGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	c	g	G	ACACCGGGCACGG	39	AAAACCCGCAAG	23
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	g	G	G	GCAGCTTCCCGG		CTGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	g	G	G	G	ACACCGGGCACGG	40	AAAACCCGCAAG	23
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	c	C	C	G	G	GCAGCTTCCCGG		CTGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	a	G	C	C	G	ACACCGGGCACGG	41	AAAACCCGGAAG	23
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	c	C	C	G	G	GCAGCTTCCCGG		CTGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	c	C	C	G	G	ACACCGGGCACGG	42	AAAACCCGCTAG	23
G	G	G	C	A	C	G	G	G	C	A	G	C	T	a	G	C	C	G	G	GCAGCTAGCCGG		CTGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	C	A	G	C	a	T	G	C	C	G	G	ACACCGGGCACGG	43	AAAACCCGGCATG	23
G	G	G	C	A	C	G	G	G	C	A	G	g	T	T	G	C	C	G	G	ACACGGGACCGG	44	AAAACCCGCAAC	23
G	G	G	C	A	C	G	G	G	C	A	G	g	T	T	G	C	C	G	G	GCAGGTTCCCGG		CTGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	C	A	c	C	T	T	G	C	C	G	G	ACACGGGACCGG	45	AAAACCCGCAAG	23
G	G	G	C	A	C	G	G	G	C	A	g	T	T	T	G	C	C	G	G	GCACCTTCCCGG		GTGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	C	t	G	C	T	T	G	C	C	G	G	ACACGGGACCGG	46	AAAACCCGGAAG	24
G	G	G	C	A	C	G	G	G	C	t	G	C	T	T	G	C	C	G	G	GCTGCTGCCCG		CAGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	g	A	G	C	T	T	G	C	C	G	G	ACACGGGACCGG	47	AAAACCCGCAAG	24
G	G	G	C	A	C	G	G	G	c	A	G	C	T	T	G	C	C	G	G	CCAGCTTCCCGG		CTCCCCGTGCCCG	
G	G	G	C	A	C	G	G	G	c	C	A	G	C	T	T	G	C	C	G	ACACGGGACCGG	48	AAAACCCGCAAG	24
G	G	G	C	A	C	G	G	G	c	C	A	G	C	T	T	G	C	C	G	CCAGCTTCCCGG		CTGGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	c	G	A	G	C	T	T	G	C	C	G	ACACGGGACCGG	49	AAAACCCGCAAG	24
G	G	G	C	A	C	G	G	G	c	G	A	G	C	T	T	G	C	C	G	GCAGCTTCCCGG		CTGCGCTGCCCG	
G	G	G	C	A	C	c	G	G	G	C	A	G	C	T	T	G	C	C	G	ACACGGGACCGG	50	AAAACCCGCAAG	24
G	G	G	C	A	C	c	G	G	G	C	A	G	C	T	T	G	C	C	G	GCAGCTTCCCGG		CTGCCGTGCCCG	
G	G	G	C	A	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACGGGACGGG	51	AAAACCCGCAAG	24
G	G	G	C	t	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACGGGCTCGG	52	AAAACCCGCAAG	24
G	G	G	G	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	GCAGCTTCCCGG		CTGCCGAGCCCG	
G	G	G	g	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACGGGACCGG	53	AAAACCCGCAAG	24
G	G	c	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACGGGACCGG	54	AAAACCCGCAAG	24
G	c	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGCAGACGG	55	AAAACCCGCAAG	24
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	c	c	G	ACACGGGACCGG	56	AAAACGGGACAG	25
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	c	G	GCAGCTTCCCGG		CTGCCGTGCCCG	
G	G	G	C	t	g	G	G	G	C	A	G	C	T	T	G	g	g	G	G	ACACGGGACCGG	57	AAAACCCCAAGC	25
G	G	G	C	A	C	G	G	G	C	A	G	C	T	T	G	g	g	G	G	GCAGCTGGGGGG		TGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	C	A	G	C	T	a	c	C	C	G	G	ACACGGGACCGG	58	AAAACCCGGTAG	25
G	G	G	C	A	C	G	G	G	C	A	G	C	T	a	c	C	C	G	G	GCAGCTACCGGG		CTGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	C	A	G	g	a	T	G	C	C	G	G	ACACGGGACCGG	59	AAAACCCGCACTCC	25
G	G	G	C	A	C	G	G	G	C	t	c	C	T	T	G	C	C	G	G	GCAGGATCCGGG		TGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	c	g	A	G	C	T	T	G	C	C	G	ACACGGGACCGG	60	AAAACCCGCAAG	25
G	G	G	C	A	C	G	G	G	c	g	A	G	C	T	T	G	C	C	G	GCTCTTCCCGG		GAGCCGTGCCCG	
G	G	G	C	A	C	G	G	G	c	g	A	G	C	T	T	G	C	C	G	ACACGGGACCGG	61	AAAACCCGCAAG	25
G	G	G	C	A	C	G	G	G	c	g	A	G	C	T	T	G	C	C	G	CGAGCTTCCCGG		CTCGCGTGCCCG	
G	G	G	C	A	C	c	c	G	C	A	G	C	T	T	G	C	C	G	G	ACACGGGACCCC	62	AAAACCCGCAAG	25
G	G	G	C	t	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACGGGCTGGG	63	AAAACCCGCAAG	25
G	G	c	g	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACGGGACCGG	64	AAAACCCGCAAG	25
G	c	c	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	GCAGCTTCCCGG		CTGCCGTGCCCG	
G	c	c	g	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACGGGACCGG	65	AAAACCCGCAAG	25
G	c	c	g	t	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	GCAGCTTCCCGG		CTGCCGTGCCCG	
G	c	c	g	t	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACGGGACCGG	66	AAAACCCGCAAG	25
G	c	c	g	t	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACGGGACCGG	67	AAAACCCGCAAG	26
G	c	c	g	t	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACGGGACCGG	68	AAAACCCGCAAG	26
G	c	c	g	t	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACGGGACCGG	69	AAAACCCGCAAG	26
G	c	c	g	t	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACGGGACCGG	70	AAAACCCGCAAG	26

Tabela A

Tabela A																			Oligos para Geração de plasmídeo de expressão de gRNA			
Posição de Sequência Alvo 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	299
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	109 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	109 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	ACACCGATGCCGTT 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	G	C	G	T	T	C	T	T	C	T	G	C	T	T	G	ACACCGATGCCGTT 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	ACACCGATGCCGTT 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	ACACCGATGCCGTT 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 CTTCTGCAAGTG	118 AAAACACAAGCAGA AGAACGGCATCG	312
G	A	T	G	C	C	G	T	T	C	T	T	C	T	G	C	a	a	G	T	ACACCGATGCCGTT CTTCTGCTTGTG	119 AAAACACAAGCAGA AGAACGGCATCG	313
G	A	T	G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGATGCCGTT CTTCTGCTTGTG	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	ACACCGATGCCGAT 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 CTCTCTGCTTGTG	124 AAAACACAAGCAGA AGAACGGCATCG	318
G	A	a	c	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGAACCCTGTT 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	ACACCGTAGCCGTT CTTCTGCTTGTG	127 AAAACACAAGCAGA AGAACGGGTACG	321
G	t	a	c	g	G	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGTAGCCGTT 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	ACACCGTAGCCGTT CTTCTGCTTGTG	129 AAAACACAAGCAGA AGAACCCGTACG	323
G	t	a	c	g	g	c	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGTAGCCGTT CTTCTGCTTGTG	130 AAAACACAAGCAGA AGAACCCGTACG	324
G	t	a	c	g	g	c	a	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGTAGCCGAT 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	ACACCGTAGCCGCA 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	ACACCGTAGCCGCA AGTTCTGCTTGTG	133 AAAACACAAGCAGA ACTTGGCGTACG	327
G	A	T	G	C	C	G	T	T	C	T	T	C	T	G	C	T	a	G	a	ACACCGATGCCGTT CTTCTGCTAGAG	134 AAAACTCTAGCAGA AGAACGGCATCG	328
G	A	T	G	C	C	G	T	T	C	T	T	C	T	G	g	T	a	G	T	ACACCGATGCCGTT CTTCTGCTAGTG	135 AAAACACTACCAGA AGAACGGCATCG	329
G	A	T	G	C	C	G	T	T	C	T	T	C	a	G	C	T	a	G	T	ACACCGATGCCGTT CTTCTGCTAGTG	136 AAAACACTAGCTGA AGAACGGCATCG	330
G	A	T	G	C	C	G	T	T	C	T	T	C	a	G	C	T	a	G	T	ACACCGATGCCGTT CTTCTGCTAGTG	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 AGATGGCATCG	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 AGAACCGCATCG	334
G	A	T	c	C	C	G	T	T	C	T	T	C	T	G	C	T	a	G	T	ACACCGATGCCGTT CTTCTGCTAGTG	141 AAAACACTAGCAGA AGAACCGCATCG	335

Tabela A

Tabela A																			Oligos para Geração de plasmídeo de expressão de gRNA			
Posição de Sequência Alvo de gRNA																						
G	G	a	G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGAGGTGC	17	AAAAGTGAAGTTCA
																				TCTGCACCTCCG	37	
G	c	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCTGGTGC	17	AAAAGTGAAGTTCA
																				TCTGCACCAAGCG	37	
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	g	t	ACACCGTGGTGC	17	AAAACACAAGTTCA
																				TCTGCACCACCG	37	
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	C	a	a	C	A	ACACCGTGGTGC	17	AAAAGTGAACACAG
																				TCTGCACCACCG	37	
G	G	T	G	G	T	G	C	A	G	A	T	G	A	t	g	T	T	C	A	ACACCGTGGTGC	18	AAAAGTGAACATCA
																				TCTGCACCACCG	37	
G	G	T	G	G	T	G	C	A	G	A	T	c	t	A	C	T	T	C	A	ACACCGTGGTGC	18	AAAAGTGAAGTGA
																				TCTGCACCACCG	37	
G	G	T	G	G	T	G	C	A	G	t	a	G	A	A	C	T	T	C	A	ACACCGTGGTGC	18	AAAAGTGAAGTTCT
																				ACTGCACCACCG	37	
G	G	T	G	G	T	G	C	t	c	A	T	G	A	A	C	T	T	C	A	ACACCGTGGTGC	18	AAAAGTGAAGTTCA
																				TGAGCACCACCG	37	
G	G	T	G	G	T	c	g	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGTGGTGC	18	AAAAGTGAAGTTCA
																				TCTCGACCACCG	37	
G	G	T	G	c	a	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGTGGTGC	18	AAAAGTGAAGTTCA
																				TCTGTCACCACCG	37	
G	G	a	c	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGAGCTGC	18	AAAAGTGAAGTTCA
																				TCTGCACTTCGG	38	
G	c	a	G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCAGGTGC	18	AAAAGTGAAGTTCA
																				TCTGACCTACGG	38	
G	c	a	c	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCACCTGC	18	AAAAGTGAAGTTCA
																				TCTGCAAGTGCG	38	
G	c	a	c	c	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TCTGCTGGTGC	38	
G	c	a	c	c	a	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TCTGCTGGTGC	38	
G	c	a	c	c	a	c	g	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TCTGCTGGTGC	38	
G	c	a	c	c	a	c	g	t	G	A	T	G	A	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	38	
G	c	a	c	c	a	c	g	t	c	A	T	G	A	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	38	
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	a	C	t	ACACCGGTGGTGC	19	AAAACAGTAGTTCA
																				TCTGCACCACCG	38	
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGTGGTGC	19	AAAACGTACTTCA
																				TCTGCACCACCG	39	
G	G	T	G	G	T	G	C	A	G	A	T	G	t	A	C	T	a	C	A	ACACCGGTGGTGC	19	AAAACGTAGTACA
																				TCTGACCTACAG	39	
G	G	T	G	G	T	G	C	A	G	A	T	G	t	A	C	T	a	C	A	ACACCGGTGGTGC	19	AAAACGTAGTACA
																				TCTGACCTACAG	39	
G	G	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGTGGTGC	19	AAAACGTAGTACA
																				TCTGACCTACAG	39	
G	G	T	c	G	T	G	C	A	G	A	T	G	A	A	C	T	a	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TCTGCACTTCAG	39	
G	c	T	G	G	T	G	C	A	G	A	T	G	A	A	C	T	a	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGACGAACTACAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	A	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGAAGTTCA
																				TGTGACCTACAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A	C	T	T	C	A	ACACCGCACCCAG	19	AAAAGTGAAGTTCA
																				TGATGAACTTCAG	39	
G	G	T	G	G	T	G	C	A	c	A	T	G	t	A								

Tabela A																										
Posição de Sequência Alvo de gRNA																								Oligos para Geração de plasmídeo de expressão de gRNA		
G	c	T	G	G	T	G	C	A	G	A	T	G	A	A	g	T	T	C	A	AGATGAACTTCG		TCTGCACCAGCG				
G	c	T	G	G	T	G	C	A	G	A	T	G	t	A	C	T	T	C	A	ACACCGCTGGTGC	213	AAAACTGAACTTC	40			
G	c	T	G	G	T	G	C	A	G	A	a	G	A	A	C	T	T	C	A	AGATGAACTTCAG		TCTGCACCAGCG				
G	c	T	G	G	T	G	g	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCTGGTGC	214	AAAACTGAAGTACA	40			
G	c	T	G	G	T	G	g	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCTGGTGG	215	AAAACTGAAGTTCT	40			
G	c	T	G	G	T	G	a	G	C	A	G	A	T	G	A	A	C	T	T	AGATGAACTTCAG		TCTCCACCAGCG				
G	c	T	c	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCTGTGC	216	AAAACTGAAGTTCA	41			
																				AGATGAACTTCAG		TCTGCACGAGCG				
Alvo 1 endógeno (Sítio 1 de VEGFA)																										
2	1	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')			
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	ACACCGGGTGGGG	217	AAAACGGAGCAAAC	41			
G	G	G	T	G	G	G	G	G	G	A	G	T	T	T	G	C	T	C	C	GGAGTTGCTCG		TCCCCCACCAG				
																				220			414			
Alvo 2 endógeno (Sítio 2 de VEGFA):																										
2	1	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')			
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	ACACCGACCCCT	221	AAAACGAGGCCGG	415			
G	A	C	C	C	C	C	C	T	C	C	A	C	C	C	C	G	C	C	T	C	CCACCCCGCCTCG		GTGGAGGGGGTCG			
																				222			416			
Alvo 3 endógeno (Sítio 3 de VEGFA):																										
2	1	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')			
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	ACACCGGTGAGTG	223	AAAACACGCACAC	417			
G	G	T	G	A	G	T	G	A	G	T	G	T	G	T	G	C	G	T	G	AGTGTGTGCGTGG		ACTCACTCACCG				
																				224			418			
Alvo 4 endógeno (EMX1):																										
2	1	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')			
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	ACACCGAGTCCGA	225	AAAACCTTCTCTCT	419			
G	A	G	T	C	C	G	A	G	C	A	G	A	A	G	A	A	G	A	A	GCAGAGAAGAA		GCTCGGACTCG				
																				226			420			
Alvo 5 endógeno (RNF2):																										
2	1	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')			
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	ACACCGTGTCTTA	227	AAAACCAGGTAATG	421			
G	T	C	A	T	C	T	T	A	G	T	C	A	T	T	A	C	C	T	G	GTCATTACCTGG		ACTAAGATGACG				
																				228			422			
Alvo 6 endógeno (FANCF):																										
2	1	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')			
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	ACACCGGAATCCCT	229	AAAACGGTGCCTGA	423			
G	G	A	A	T	C	C	C	T	T	C	T	G	C	A	G	C	A	C	C	TCTGCAGCACCG		GAAGGGATTCCG				

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:.

Ensaios de Atividade de EGFP

[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 transfeções, 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 LSRII. As transfeções para otimização da concentração de plasmídeo de gRNA/Cas9 foram realizadas em triplicata e todas as outras as transfeções foram realizadas em duplícata.

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
	TCCAGATGGCACATTGTC AG	436.	AGGGAGCAGGAAAGTGAG GT	748.	DMSO			
	GGGGCCCCACTCTTCTTC CAT	437.	ACCCAGACTCCTGGTGTG GC	749.	No DMSO	0	0	1
	GCTAACAGAGATGCCT ATGCC	438.	ACCACCCCTTCCCCAGAA AA	750.	DMSO	2	0	0
	ACCCCCACAGCCAGGTTT CA	439.	GAATCACTGCACCTGGCC ATC	751.	DMSO	0	0	2
	TGCGGCAACTCAGACAA CC	440.	TAAAGGGCGTGCTGGGAG AG	752.	DMSO	1	1	0
	GCATGTCAGGATCTGAC CCC	441.	TGCAGGGCATCTTGTGT GT	753.	DMSO	0	2	0
	CCACCCACATGTTCTGGGT GC	442.	CTGGGTCTGTTCCCTGTG GG	754.	DMSO	1	1	1
	GGCTCTCCCTGCCCTAG TTT	443.	GCAGGTCAAGTTGAAACC CG	755.	DMSO	0	2	1
	GGGGCTGAGAACACATG AGATGCA	444.	AGATTGTCAGTGCCTG CCT	756.	DMSO	1	0	2

TABELA B								
Alvo real em células U2OS. EGFP	Iniciador de PCR direto	SEQ ID N O	Iniciador de PCR reverso	SEQ ID N O:	Condições de PCR	Transformações de Watson-Crick	transversões que não Watson-Crick	Transformações
	CCCGACCTCCGCTCAA AGC	445.	GGACCTCTGCACACCTG GC	757.	DMSO	2	1	0
	TGCAAGGTGCGATAGTC CCA	446.	CAGGAGGGGAAGTGTGT CC	758.	DMSO	1	1	1
	GCCCATTCTTTGCACT GGA	447.	GAGAGCAAGTTGTTCCC CAGG	759.	DMSO	0	1	2
	GCCCCCAGCCCCCTGT TTC	448.	GCTGCTGGTAGGGAGCT GG	760.	DMSO	1	2	0
	CGGCTGCCCTCCCTGAG TCC	449.	GGGTGACGCTTGCCATGA GC	761.	72C recoze r, DMSO a 3%	1	2	0
	TGACCCCTGGAGTACAAA TGTTCCA	450.	GCTGAGACAACCAGCCCA GCT	762.	72C recoze r, DMSO a 3%	2	1	0
	TGCCTCCACCCTTAGCCC CT	451.	GCAGCCGATCCACACTGG GG	763.	DMSO	1	0	2
	AACTCAGGACAACACTGC CTGT	452.	CCCAAGGAGCAGGGTACAA TGC	764.	DMSO	0	1	2
	TCCTCCTGGAGAGGGG CCC	453.	CCTTGGAAAGGGCCTTGG TGG	765.	DMSO	0	3	0
	CCGAGGGCATGGCAAT CCT	454.	GGCTGCTGGAGTTGCCA AC	766.	DMSO	0	1	3
	TGCTTTCGATGGGTCTC AGACA	455.	GGGTTGCTTGCCTCTGT GT	767.	DMSO	0	2	2
	AGCTCCTCTCATTTCTC TTCTGCTG	456.	CACAGAAGGATGTGTGCA GGTT	768.	DMSO	0	2	2
	AGCAGACACAGGTGAAT GCTGCT	457.	GGTCAGGTGTGCTGCTAG GCA	769.	DMSO	1	1	2
	CCTGTGGGCTCTCAGG TGC	458.	ACTGCCTGCCAAAGTGGG TGT	770.	Sem DMSO TD	1	1	2
	AGCTGCACTGGGAATG AGT	459.	TGCCGGGTAATAGCTGGC TT	771.	DMSO	0	1	3
	CCAGCCTGGCACAAA GCG	460.	GGGGGCTTCCAGGTACA GG	772.	72C recoze r, DMSO a 3%, 6% DMSO	0	3	1
	TACCCCCACTGCCATT GC	461.	ACAGGTCCATGCTTAGCA GAGGG	773.	DMSO	0	1	3
GGGTGATTGAAGTTGCTCCAGG GGGTGATTGAAGTTGCTGCAGG (SEQ ID NO:424)	ACGGATTCAAGCACGGAG GTGC	462.	CCGAGTCCGTGGCAGAGA GC	774.	DMSO	0 / 1	2	2
	TGTGGTTGAAGTAGGGG ACAGGT	463.	TGGCCCAATGGAAAGTGA TTCTGT	775.	DMSO	3	1	0
	TGGGATGGCAGAGTCAT CAACGT	464.	GGCCCAATCGGTAGAGGA TGCA	776.	DMSO	0	3	1
	ATGGGGCGCTCCAGTCT GTG	465.	TGCACCCACACAGCCAGC AA	777.	DMSO	0	3	1
	GGGGAGGGAGGACAG GGAA	466.	AATTAGCTGGCGCGGTG GT	778.	72C recoze r, DMSO a 3%	0	1	3
	ATCCCGTGCAGGAAGTC GCC	467.	CAGGCAGGCCCTTGAGGA AT	779.	DMSO	3	1	0
	CCCCAACCTTGCTCAG CG	468.	TGAGGAGAACACACAGG CAGA	780.	DMSO	1	2	1
	ATCGACGAGGAGGGGC CTT	469.	CCCCTCACTCAAGCAGGC CC	781.	DMSO	0	3	1
	TGCTCAAGGGCCTGTT CCA	470.	CAGGGCAGTGGCAGGA GTC	782.	No DMSO	1	3	0
	TGCCTGGCACCGAGTAG GTG	471.	GGGAAGGGGAAACAGGT GCA	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
ACCTGGGCTTGCCACTAGGG	472.	GCTGCTCGCAGTTAACCA	784.	DMSO	1	3		1
GTGGCCGGGCTACTGCTACC	473.	GGTTCACAAAGCTGGGGGCA	785.	DMSO	3	2	0	
Não otimizado						1	3	1
GCAAGAGGCCGGAGGAGACCC	474.	AGAGTCATCCATTCTGGGGC	786.	DMSO	2	3	0	
GGGGTCAGTGGTGTATATCCCCCT	475.	AGGGAATCCTTTCCATTGCTTGT	787.	Betain a a 1M, TD	1	4	0	
AGAGAGGCCACGTGGAGGGT	476.	GCCTCCCCCTCCCTCCCTCC	788.	DMSO	1	3	1	
GACAGTGCCTTGCATGCAC	477.	TCTGACCGGTATGCCTGACG	789.	DMSO	3	2	0	
TGTGTGAACGCAGCTGGCT	478.	TGGTCTAGTACTCCCTCCA	790.	DMSO	3	1	1	
GGTTCTCCCTGGCTCTGTGA	479.	CCCACTGCTCTAGCCCTGC	791.	DMSO	1	3	1	
TGAAGTCAACAATCTAACCTTCCACCT	480.	AGCTTTGGTAGTTGGAGCTTTGAAGG	792.	DMSO	3	1	2	
TGATTGGGCTGCAGTTCA	481.	GCACAGCCTGCCCTTGGAG	793.	DMSO	2	1	3	
TCCATGGGCCCTCTGA	482.	AGCGGCTCTGCTTCTGCAGA	794.	DMSO	1	0	5	
GCGGTTGGTGGGTTGATGC	483.	GAGTTCCCTCCCGCCA	795.	DMSO	2	0	4	
AGGCAAGATTTCCAGTGTGCAAGA	484.	GCTTTGCCTGGACTCCAGC	796.	DMSO	2	0	4	
GCTGCTGGCTGGGCTCTGCT	485.	GCTCTGCCCCACTTCCCCCTGG	797.	Sem DMSO TD	3	1	2	
GCTGCGAGGCTCCCGTGA	486.	CGCCCCCTAGAGCTAAGGGGGT	798.	DMSO	3	2	1	
CCAGGAGCCTGAGAGCTGCC	487.	AGGGCTAGGACTGCAGTGAGC	799.	DMSO	1	3	2	
CTGTGCTCAGCCTGGGT	488.	GCCTGGGCTGTGAGTAGTT	800.	DMSO	2	3	1	
AGCTCGGCCAGATCTGTGG	489.	ACTTGGCAGGCTGAGGCA	801.	72C recone r, DMSO a 3%	4	2	0	
AGAGAAAGTCGAGGAAGAGAG	490.	CAGCAGAAAGTTCATGGTTTCG	802.	DMSO				
TGGACAGCTGCAGTACTCCCTG	491.	ACTGATCGATGATGGCCTATGGT	803.	DMSO	0	0	2	
CAAGATGTCACTGGGCTA	492.	GCAGCTATGTCTCTGT	804.	DMSO	1	0	1	
GTCCAGTGCCTGACCCCTGGC	493.	AGCATCATGCCCTCAGCTCA	805.	DMSO	1	1	1	
GCTCCCGATCCCTCTGCCACC	494.	GCAGCTCCCACCAACCCCTCAG	806.	DMSO	1	2	0	
GGGGACAGGCAGGCAAGAGAG	495.	GTGCGTGTCCCGTTACCCCT	807.	DMSO	1	1	1	
AAGGGGCTGCTGGGTAGGAC	496.	CGTGTGAGTTCCCTGGCA	808.	DMSO	2	1	0	
GACCTCTAGGAAGCTGGAGAG	497.	CTGCGAGATGCCCAAATCG	809.	Betain a a 1M, TD	1	0	2	
CCCGGGCGCTCTGCTAGA	498.	TGCTGGGATTACAGGCGCGA	810.	DMSO	1	1	1	
CCAGGTGGTGTAGCGGAGG	499.	TGCCTGGCCCTCTGTGAGTCT	811.	DMSO	0	2	1	
CGACTCCACGGCGTCTAGG	500.	CAGCGCAGTCCAGCCCGATG	812.	Betain a a 1M, TD	2	1	0	
CTTCCCTCCCCCACACAC	501.	GCTACAGGTTGCACAGTGAGAGGT	813.	DMSO	1	1	1	

TABELA B

Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SEQ ID N O	Iniciador de PCR reverso	SEQ ID N O:	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	CCCCGGGGAGTCTGTCC TGA	502.	CCCAGCCGTTCCAGGTCT TCC	814.	72C Recozer, DMSO a 3%	1	0	2
	GAAGGCGCAAACCGG CTC	503.	TCCAGGGCTCTCGGC CC	815.	DMSO	1	0	2
	AGGGTGGTCAGGGAGGC CTT	504.	CATGGGGCTCGAACCTCG TC	816.	DMSO	2	0	1
	GGGAAGAGGCAGGGCTG TCG	505.	TGCCAGGAAGGAAGCTGG CC	817.	72C Recozer, DMSO a 3%	0	2	1
	GAGTGACGATGAGCCCC GGG	506.	CCCTTAGCTGCAGTCGCC CC	818.	68C Recozer, DMSO a 3%	0	1	3
	CCCATGAGGGGTTTGG AGTGC	507.	TGAAGATGGCAGTTGG GG	819.	DMSO	0	2	2
	CACCTGGGCATCTGGG TGG	508.	ACTGGGGTTGGGGAGGG GAT	820.	DMSO	2	0	2
	TCATGATCCCCAAAGGG CT	509.	CCATTGCTGATCTGTG GGT	821.	DMSO	1	0	3
	TGGTGCCCAGAATAGTG GCCA	510.	AGGAAATGTTGTGCCA GGGC	822.	DMSO	1	2	1
	GCCTCAGACAACCCCTGC CCC	511.	GCCAAGTGTACTCATCAA GAAAGTGG	823.	Sem DMSO TD	2	1	1
	GCCGGGACAAGACTGAG TTGGG	512.	TCCCGAACATCCCGAAAA CG	824.	DMSO	1	2	1
	TGCTGCAGGTGGTCCG GAG	513.	CTGGAACCGCATCCTCCG CA	825.	Sem DMSO TD	1	0	3
	ACACTGGTCAGGTCCC GTCT	514.	GGCTGTGCCCTCCGATGG AA	826.	DMSO	2	1	1
CTCTCCCCCACCCCCCTCTGG (SEQ ID NO:425)	ATCGCGCCAAAGCACA GGT	515.	AGGCTCTGGAAAAGTCC TCAATGCA	827.	DMSO	3	0	2
Não otimizado						1	1	2
	CCCTCATGGGGTCTTAC GGCA	516.	AGCCACACATCTTCTGGT AGGG	828.	DMSO	1	1	2
	TGGTCGCTCATGCTGG GAG	517.	AGGGTGGGTGTACTGGC TCA	829.	DMSO	0	3	1
	GAGCTGAGACGGCACCA CTG	518.	TGGCTTGAACCTTGGG CT	830.	betaín 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
	GGGAGAACCTGTCCAG CCT	520.	AAGCCGAAAAGCTGGCA AA	833.	DMSO	0	2	3
	CTTCCCAGTGGCCCG TCC	521.	ACACAGTCAGAGCTCCCG CG	834.	DMSO	1	1	3
Não otimizado						1	0	4
	CTGAGAGGGGAGGGG GAGG	522.	TCGACTGGTCTTGTCCCT CCA	835.	68C Recozer, DMSO a 3%	3	0	2
	CAGCCTGCTGCATCGGA AAA	523.	TGCAGCCAAGAGAAAAAG CCT	836.	betaín a a 1M, TD	1	0	4
	TCCCTCTGACCCGGAAC CCA	524.	ACCCGACTTCCCTCCCAATT 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 GCTCC	527.	ACCTCTCCAAGTCTCAGTA ACTCTCT	840.	DMSO	1	3	1
	GGTCCCTCTGTGCAGTG GAA	528.	CTTTGGTGGACCTGCACA GC	841.	DMSO	2	2	2
	GCGAGGCTGCTGACTTC CCT	529.	GCTGGGACTACAGACATG TGCCA	842.	DMSO	2	2	2
ATTTCTCCCCCCC-CCTCAGG (SEQ ID NO:426)	ATTGCAGGCGTGTCCAG GCA	530.	AAATCCTGCATGGTGATG GGAGT	843.	DMSO	1	1	5
	TGCTCTGCCATTATGTC CTATGAAC	531.	ACAGCCTCTTCATGAC TGAGC	844.	DMSO	1	3	2
	TCCGCCAACAGGAGG CAG	532.	GCGGTGGGAAGGCCATTG AG	845.	DMSO	2	3	1
	GGGGGTCTGGCTCACCT GGA	533.	CCTGTCGGAGAGTCCT GC	846.	DMSO	3	1	2
	TCCTGTTCATTTGCTAG AACTCTGGA	534.	ACTCCAGATGCAACCAGG GCT	847.	DMSO	3	2	1
	CGTGTGGTAGGCTGAG TCT	535.	GCTTCACCGTAGAGGCTG CT	848.	DMSO	3	0	3
	AGGCCCTGATAATTATG CTACCAA	536.	TCAGTGACAACTTTGTG TTCGGCA	849.	DMSO	0	2	4
	Não otimizado					2	2	2
		537.						
	TCCAGATGGCACATTGTC AG	538.	AGGGAGCAGGAAAGTGAG GT	850.	DMSO			
	GCAGGCAAGCTGTCAG GGT	539.	CACCGACACACCACTCA CC	851.	DMSO	0	0	1
	GAGGGGAAGTCACCGA CAA	540.	TACCCGGCCGCTGTGA GA	852.	DMSO	0	0	2
	GACACCCACACACTCTC ATGC	541.	TGAATCCCTCACCCCCAA G	853.	DMSO	1	0	1
	TCCTTGAGGTTCATCCC CC	542.	CCAATCCAGGATGATTCC GC	854.	DMSO	1	0	1
	CAGGGCCAGAACACAG GAA	543.	GGGAGGTATGCGGGGA GTG	855.	DMSO	1	1	0
	TGCAGCCTGAGTGAGCA AGTGT	544.	GCCCAGGTGCTAACCCCC TC	856.	DMSO	1	0	1
	TACAGCCTGGGTATGG AGC	545.	TGTGTATGGACTTTCCA TTGT	857.	betaín a a 1M, TD	1	1	0
	GGCAGGCATTAACACTCAT CAGGTC	546.	TCTCCCCAAGGTATCAG AGAGCT	858.	DMSO	1	1	0
	GGGCCTCCCTGCTGGTT CTC	547.	GCTGCCCTCGAACCCAA GA	859.	DMSO	0	1	1
	ACAAACGCAGGGGACATTG GAA	548.	ACTCCGAAATGCCCGC AGT	860.	DMSO	1	1	0
	AGGGGAGGGGACATTG CTC	549.	TTGAGAGGTTCACTGGT TGC	861.	DMSO	1	0	1
	CTAATGCTTAAGGCTGGC GG	550.	AGCCAACGGCAGATGCAA AT	862.	DMSO	1	0	1
	GAGCGAAGTTAACCCAC CGC	551.	CACACATGCACATGCCCG TG	863.	68C, DMSO a 3%	2	0	0
	GCATGTGCTAACTGGAG ACAATAGCA	552.	TCCCCCATACACACACA CA	864.	DMSO	2	0	0
	GCCCCCTCCGCCTTTGT GT	553.	TGGGCAAAGGACATGAAA CAGACA	865.	DMSO	2	0	0
	GCCTCAGCTCTGCTCTTA AGCC	554.	ACGAAACAGATCATTTCGA TGGCTCC	866.	DMSO	2	0	0
	CTCCAGAGCCTGGCCTA CCA	555.	CCCTCTCCGGAAGTGCCT TG	867.	DMSO	0	1	1
	TCTGTCACACACAGTTA CCACC	556.	GTTGCCTGGGATGGGT AT	868.	DMSO	0	1	1
	GGGGACCCCTCAAGAGGC 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	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
	AGCTGAGGCAGAGTCCC CGA	560.	GGGAGACAGAGCAGCGC CTC	872.	DMSO	1	1	1
	ACCACCAGACCCCACCT CCA	561.	AGGACGACTTGTGCCCA 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.	GCCCCCTCACCATGAT GCT	875.	DMSO	2	0	1
	GGGGGACATTCCCTCTGCT GCA	564.	TGGGGATCCTGCTCATG GC	876.	DMSO	3	0	0
	ACACACTGGCTCGCATTC ACCA	565.	CCTGCACGAGGCCAGGTG TT	877.	DMSO	2	1	0
	TGGGCACGTAGTAAACT GCACCA	566.	CTCGCCGCCGTGACTGTA GG	878.	DMSO	0	3	1
	TCAGCTGGCTCTGGCT TGG	567.	AGAGCACTGGGTAGCACT CAGT	879.	DMSO	2	1	0
	AGACACAGCCAGGGCCT CAG	568.	GGTGGGCGTGTGTGTGA CC	880.	68C, DMSO a 3%	1	1	1
	ACACTCTCACACACGCAC CAA	569.	GAGAAGTCAGGGCTGGCG GG	881.	72C Recoz er, DMSO a 3%	1	2	0
	ACTGCCTGCATTCCCCG GT	570.	TGGTGAGGGCTTCAGGGA GC	882.	DMSO	1	1	1
	GCCAGGTTCATGACTGC CC	571.	TCCTTCTACACATCGGCG GC	883.	DMSO	2	1	0
	CGAGGGAGCCGAGTTCG TAA	572.	CTGACCTGGGCTCTGGT AC	884.	DMSO	1	2	0
	TCCCTGGGAAGTCATGG CTTCA	573.	GCAC TGAGCAACCAGGAG CAC	885.	DMSO	2	1	0
Não otimizado						1	0	3
	TAACCGTGTCCCCCGC CTC	574.	GCTCCCTGCCAGGTGAA CC	886.	DMSO	2	1	1
	CCTGCTGAGACTCCAGG TCC	575.	CTGCGGAGTGGCTGGCTA TA	887.	DMSO	2	0	2
	CTCGGGGACTGACAAGC CGG	576.	GGAGCAGCTTCCAGGG CC	888.	DMSO	3	0	1
	CCCGGACCAAAGCAGGA GCA	577.	CTGGCAGCCTCTGGATGG GG	889.	DMSO	1	2	1
Não otimizado						0	3	1
	ATTCAGAGCCCCGGGG AAA	578.	AGGCCGCAGGTGTTATGGT TA	890.	DMSO	1	2	1
	GCCAGTGGCTTAGTGTCT TTGTGT	579.	TGACATATTTCTGGGCC ATGGGT	891.	DMSO	2	1	1
	TGCCAGAAGAACATGGG CCAGA	580.	CCATGCTGACATCATATA C TGGGAAGC	892.	DMSO	3	1	0
	GCGTGTCTGTGTGCG TGC	581.	CCAGGCTGGCACACAGG TT	893.	DMSO	3	1	0
Não otimizado						2	2	0
	TGCCCAGTCCAATATTC AGCAGCT	582.	AGGATGAGTTCATGTCCTT TGTTGGG	894.	DMSO	2	2	0
	GGGTGAAAATTTGGTACT GTTAGCTGT	583.	AATGACTCATTCCCTGGGT ATCTCCA	895.	DMSO	2	2	0
	TGCCCCATCAATCACCTC GGC	584.	CAAGGTGGCAGGGCACT GA	896.	DMSO	1	2	2
	GCCTCCTCTGCCGCTGG TAA	585.	TGAGAGTCTGTGTGCTC CACACT	897.	DMSO	1	2	2
Não otimizado						2	2	1
	GCCACCAAAATAGCCAG CGT	586.	ACATGCATCTGTGTGTC GT	898.	DMSO	3	0	2
	ACAGACTGACCCCTGAAA AATACCACT	587.	TGTATTTCTTGCCAATG GTTTCCC	899.	DMSO	2	1	2

TABELA B

Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SEQ ID N O	Iniciador de PCR reverso	SEQ ID N O:	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	AGCCAAATTCCTAACAG CAGCACT	588.	TCCTGGAGAGCAGGCATT TTTGT	900.	DMSO	3	1	1
	ACCTCTTGTGCTGCCCTG GC	589.	GGCGGGAAGGTAACCCCTG GG	901.	DMSO	2	1	2
	CACAAAGCTCACCTTC CAGTAGTGT	590.	TGATCCGATGGTTGTTCAC AGCT	902.	DMSO	3	1	1
	TGTGGGGATTACCTGCCT GGC	591.	ACGCACAAAATGCCCTT GTCA	903.	DMSO	2	2	1
	TGAGGCAGACCAGTCAT CCAGC	592.	GCCCGAGCACAGTGTAGG GC	904.	DMSO	2	3	0
	ATTAGCTGGCGTGGCC GAG	593.	ACTGCATCTCATCTCAGG CAGCT	905.	DMSO	2	1	3
	TGAAGCAGAAGGAGTGG AGAAGGA	594.	TCAGCTTCACATCTGTTTC AGTTCACT	906.	DMSO	4	0	2
	TGGTGGAGTGTGTGTGT GGT	595.	AGAGCAGAAAGAGAGTGCC AAG	907.	DMSO	1	3	2
	GCCCCCTGTACGCTCTGA CAGC	596.	TGCACAAGCCACTTAGCC TCTCT	908.	DMSO	3	1	2
	AGCGCAGGTAACAGGC CCA	597.	TCTCTGCCCGTTCCTT GT	909.	DMSO	3	1	2
	ATGGGTGCCAGGTACCA CGC	598.	ACAGCAGGAAGGAGGCCGC AG	910.	DMSO	2	3	1
	CGGGCGGGTGGACAGAT GAG	599.	AGGAGGTCTCGAGGCCAGG GG	911.	DMSO	2	3	1
	TCAACCTAGTGAACACAG ACCACTGA	600.	GTCTATATACAGCCCACAA CCTCATGT	912.	DMSO	1	2	3
	GCCAGGGCCAGTGGATT GCT	601.	TGTCACTTCTTAGTATGTC AGCCGGA	913.	DMSO	2	4	0
	GAGCCCCACCGGTTCA TAG	602.	GCCAGAGCTACCCACTCG CC	914.	DMSO	1	3	2
		603.						
	GGAGCAGCTGGTCAGAG GGG	604.	GGGAAGGGGACACTGG GGA	915.	DMSO			
	TCTCTCCTCAACTCATG ACCAGCT	605.	ATCTGCACATGTATGTACA GGAGTCAT	916.	DMSO	0	1	1
AAGACAGAGGAGAAGAAGGG (SEQ ID NO:427)	TGGGGATCTCAAAGAA CCC	606.	AGGGTGTACTGTGGGAAC TTGCA	917.	DMSO	2	1	1
	GATGGCCCCACTGAGCA CGT	607.	ACTTCGTAGAGCCTAAC ATGTGGC	918.	DMSO	1	0	2
	AGGATTAATGTTAAAGT CACTGGTGG	608.	TCAAAACAGGTGCAGATA CAGCA	919.	betaín a a 1M, TD	1	0	2
	TCCAAGCCACTGGTTCT CAGTCA	609.	TGCTCTGGGATCATATTT TGGGGA	920.	DMSO	0	1	2
	ACTTTAGAGCTGGGG CAGGT	610.	CCACGCTGAAGTGCAAT GGC	921.	DMSO	1	1	1
	CAAAGCATGCCCTTCAGC CG	611.	GGCTCTCGATTTGGCAC CT	922.	betaín 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.	CCGACCTTCATCCCTCTT GG	924.	DMSO	0	1	2
	TGATTCTGCCCTAGAGTC CCAGGT	614.	TGGGCTCTGTGTCCCTAC CCA	925.	DMSO	0	3	0
	Não otimizado					2	1	0
	AGGCAGGAGAGCAAGCA GGT	615.	ACCCGTGACTACTGACTGA CCGCT	926.	DMSO	0	1	2
	CTCCCCATTGCGACCCG AGG	616.	AGAGGCATTGACTTGGAG CACCT	927.	DMSO	1	2	0
	CTGGAGCCCCAGCAGGAA GGC	617.	CCTCAGGGAGGGGGCCT GAT	928.	DMSO	1	2	0
	ACTGTGGCGCTTGTCCC CAC	618.	AGGTGGTGCAGGGTTTA AGGA	929.	DMSO	1	0	3
	GGCGCTCCCTTTCCCT TTGT	619.	CGTCACCCATCGTCTCGT GGA	930.	DMSO	2	0	2

TABELA B

Alvo real em células U2OS. EGFP	Iniciador de PCR direto	SEQ ID N O	Iniciador de PCR reverso	SEQ ID N O:	Condições de PCR	Transformações de Watson-Crick	transformações que não Watson-Crick	Transformações
	TGCCATCTATAGCAGCCC CCT	620.	GCATCTTGCTAACCGTACT TCTTCTGA	931.	DMSO	1	0	3
	GTGGAGACGCTAAACCT GTGAGGT	621.	GCTCCTGGCTTCCTCA CAGC	932.	DMSO	1	2	1
	CCGAACCTCTGCTGAGCT TGATGC	622.	CCAAGTCAATGGGCAACA AGGGA	933.	DMSO	0	2	2
	Não otimizado					1	1	2
	TGCCCCCAAGACCTTCT CC	623.	ATGGCAGGCAGAGGAGGA AG	934.	DMSO	2	0	2
	GGGTGGGCCATTGTGG GTT	624.	CTGGGGCCAGGGTTCTG CC	935.	DMSO	3	0	1
	TGGAGAACATGAGAGGC TTGCAA	625.	TCCTTCTGAGGCAATGG GAACAA	936.	DMSO	3	0	1
	GCCACATGGTAGAACGT GGC	626.	GGCAGATTCCCCATGC TG	937.	betaín a a 1M, TD	1	2	1
	TGTACACCCCAAGTCCTC CC	627.	AAGGGGAGTGTGCAAGCC TC	938.	DMSO	3	1	0
	AGGTCTGGCTAGAGATG CAGCA	628.	AGTCCAACACTCAGGTGA GACCT	939.	DMSO	3	1	0
	CCAAGAGGACCCAGCTG TTGGA	629.	GGGTATGAAATTCTGGAT TAGCAGAC	940.	DMSO	0	2	2
	ACCATCTCTCATGATG AGTCCCAA	630.	ACACTGTGAGTATGCTTG GCGT	941.	DMSO	2	2	0
	GGCTGGGGAGATGAG CTC	631.	TCGGATGCTTTCCACAG GGCT	942.	DMSO	2	2	1
	TCTTCCAGGAGGGCAGC TCC	632.	CCAATCCTGAGCTCCTAC AAGCT	943.	DMSO	1	0	4
	GAGCTGCACTGGATGGC ACT	633.	TGCTGGTTAAGGGGTGT TTGGA	944.	DMSO	1	1	3
	TCTGGGAAGGTGAGGAG GCCA	634.	TGGGGGACAATGGAAAAG CAATGA	945.	DMSO	0	2	3
	CTTGCTCCCAGCCTGAC CCC	635.	AGCCCTTGCCATGCAGGA CC	946.	DMSO	3	1	1
	GGGATTTTATCTGTTGG GTGCGAA	636.	AACACACAGATGTACCCCT AACGCT	947.	DMSO	2	2	1
	ACCCATCAGGACCGCAG CAC	637.	TCTGGAACCTGGGAGGG GA	948.	72C Recoz er, DMSO a 3%	3	1	1
	CGTCCCTCACAGCCAGC CTC	638.	CCTCCTGGCCTGGGT TC	949.	DMSO	1	3	1
	CCCTCTGCAAGGTGGAG TCTCC	639.	AGATGTTCTGCCCCAGG CCT	950.	DMSO	1	3	1
	GGCTTCCACTGCTGAAG GCCT	640.	TGCCGCTCACATACCCCT CC	951.	DMSO	2	1	2
	AGCATTGCGCTGCGGT GATGT	641.	AGCACCTATTGGACACTG GTTCTCT	952.	DMSO	1	3	1
	TCTAGAGCAGGGCACA ATGC	642.	TGGAGATGGAGCCTGGTG GGA	953.	DMSO	2	2	1
	GGTCTCAGAAAATGGAG AGAAAGCAGC	643.	CCCACAGAACCTGGGCC CT	954.	DMSO	1	2	3
	GGTTCTGATACCAAAAC GTTGCCT	644.	TGGGTCCCTCCACCTCT GCA	955.	DMSO	0	3	3
	ACTCTCTTAAGTACTGA TATGGCTGT	645.	CAGAACATCTGCTCTGTTGC CCA	956.	DMSO	0	4	2
	Não otimizado					2	2	2
	Não otimizado					2	2	2
	CAATGCCCTGCAGTCCTCA GGA	646.	TCCCAAGAGAAAACTCTGT CCTGACA	957.	DMSO	4	1	1
	GCATTGGCTGCCAGGG AAA	647.	TGGCTGTGCTGGCTGTG TT	958.	DMSO	2	2	2
	CCACAAGCCTCAGCCTA CCCG	648.	ACAGGTGCCAAACACTG CCT	959.	DMSO	2	1	3
TCATTGCAGCAGAAGAAGAAGG TCATTGTAGCAGAAGAAGAAGG (SEQ ID NO:428)	GCCTCTTGCAGGACAC TCCCTTT	649.	CGATCAGTCCCCTGGCGT CC	960.	DMSO	2 / 1	2 / 3	2
	TCCCAGAACATGCTCCG CA	650.	AGGGGTTCCAGGCACAT GGG	961.	DMSO	0	4	2

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	Transformações de Watson-Crick	transversões que não Watson-Crick	Transformações
		651.		962.				
	TCCTAAAAATCAGTTTG AGATTTACTCC	652.	AAAGTGTAGCCAACATAC AGAACGTAGGA	963.	DMSO			
GGTATCTAAGTCATTACCTGTGG GGTATCTAAGTCATAACCTGTGG (SEQ ID NO:429)	ACATCTGGGAAAGCAA AAGTCAACA	653.	TGTCTGAGTATCTAGGCTA AAAGTGGT	964.	DMSO	1 / 2	1	1
	ACGATCTTGCTTCATITC CCTGTACA	654.	AGTGCCTTGTGAACTGAAA AGCAAACA	965.	DMSO	0	3	0
	GCACCTTGGTGTCTAA ATGCC	655.	GGGCAACTGAACAGGCAT GAATGG	966.	DMSO	1	2	0
	AACTCTCTGCATCCCCG CC	656.	GGTCACCTGGATCCACC CA	967.	DMSO	1	1	1
Não otimizado						1	1	1
	CATCACCCCTCACCAAGG CCC	657.	ACCACTGCTGCAGGCTCC AG	968.	72C Recozer, DMSO a 3%	0	3	0
Não otimizado						2	0	2
	CCTGACCCGTGGTCCC GAC	658.	TGGTGCCTGGTGTGTGT GT	969.	72C Recozer, DMSO a 3%	1	2	1
	TGGGAACATTGGAGAAG TTTCTGA	659.	CCATGTGACTACTGGCT GCCC	970.	DMSO	1	1	2
	AGCCTTGGCAAGCAACT CCTT	660.	GGTTCTCTCTCAGAAAA GAAAGAGG	971.	DMSO	1	0	3
	GGCAGCGGACTTCAGAG CCA	661.	GCCAGAGGCTCTCAGCAG TGC	972.	DMSO	1	0	3
	CCAGCCTGGTCAATATG GCA	662.	ACTGTGCCAGCCCCATA TT	973.	DMSO	2	1	1
	ATGCCAACACTCGAGGG GCC	663.	CGGGTTGTGGCACCGGGT TA	974.	DMSO	2	1	1
	TTGCTCTAGGGGAGG GGG	664.	AGAGTTCAAGCATGAAAA GAAGCAACA	975.	DMSO	3	0	1
	AGCTGAAGATAGCAGTGT TTAAGCCT	665.	TGCAATTGAGGGCTCT CTTCA	976.	DMSO	1	1	2
	AGTCACTGGAGTAAGCCT GCCT	666.	TGCCAGCCAAAAGTTTA GTGTGT	977.	DMSO	2	0	2
	GGGTCTCCCTCAGTGCC CTG	667.	TGTGTGTAGGGAGCAA ACGACA	978.	DMSO	2	0	2
	TGGGGGCTTTAAGAGG CACA	668.	TGACCACACACACCCCCA CG	979.	DMSO	1	2	1
	TCAAAACAGATTGACCAA GGCAAAAT	669.	TGTGTTTTAAGCTGCACC CCAGG	980.	DMSO	1	0	3
	TCTGGCACCAGGACTGA TTGTACA	670.	GCACGCAGCTGACTCCCC GA	981.	DMSO	1	2	1
Não otimizado						1	0	3
	AGCATCTGTGATACCTA CCTGTCT	671.	ACCAGGGCTGCCACAGAG TC	982.	DMSO	1	0	3
	TAGCTTGTGCCCCAGGC TG	672.	CTCGGCCCCCTGAGAGTT AT	983.	DMSO	1	2	1
TCCATCTCACTCATTACCTGAGGTCCAT CTCACTCATTACCTGTATG (SEQ ID NO:430)	CTGCAACCAGGGCCCTT ACC	673.	GAGCAGCAGCAAAGCCAC CG	984.	DMSO	1	1	2
	GCCTGGAGAGCAAGCCT GGG	674.	AGCCGAGACAATCTGCC CG	985.	DMSO	1	1	2
TTTATATTAGTGTATTACCTGCGG (SEQ ID NO:431)	AGTGAACAAACAAGCAG CAGTCTGA	675.	GGCAGGTCTGACCAGTGG GG	986.	Sem DMSO TD	1	2	1
	AGGCTCAGAGAGGTAAAG CAATGGA	676.	TGAGTAGACAGAAATGTTA CCGGTGTT	987.	DMSO	3	0	2
	TCAGAGATGTTAAAGCCT TGGTGGG	677.	AGTGAACCAAGGAAATGG GGGG	988.	DMSO	3	0	2
	TGTGCTTCTGGGGTAGT GGCA	678.	CACCTCAGCCCTGTAGTC CTGG	989.	DMSO	0	4	1
	CCATTGGGTGACTGAATG CACA	679.	GCCACTGTCCCCAGCCTA TT	990.	betaein a a1M, TD	1	3	1

TABELA B

Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SEQ ID N O	Iniciador de PCR reverso	SEQ ID N O:	Condições de PCR	Transversões de Watson-Crick	transversões que não Watson-Crick	Transições
	ACCAAGAAAGTGAAG GAAACCC	680.	TGAGATGGCATACGATTTC CCCA	991.	DMSO	1	2	2
	AGGGTGGGACTGAAAG GAGCT	681.	TGGCATCACTCAGAGATT GGAACACA	992.	DMSO	3	1	1
	ACCAGTGCTGTGACCT TGGA	682.	TCCTATGGGAGGGAGGC TTCT	993.	DMSO	3	1	1
	CCAGGTGTTGTTCA TGAC	683.	GCATACGGCAGTAGAATG AGCC	994.	68C, DMSO a 3%	4	0	1
	CAGGCCTGGTTCTTA GCCT	684.	CCTTCCTGGCCCCATGG TG	995.	DMSO	2	3	0
	TGGGGTCCAAGATGTCC CCT	685.	TGAAACTGTTGATGAGG TGTGGA	996.	DMSO	1	2	2
	GCTGGCTTGGTGTAT ATGC	686.	ACTTGCAAAGCTGATAACT GACTGA	997.	DMSO	5	0	1
	AGTTGGTCACTGACAA TGGGA	687.	CGCAGCCACGAGTTCAT CA	998.	DMSO	3	0	3
	AGAGGAGGCACAATTCA ACCCCT	688.	GGCTGGGAGGCCCTACA AT	999.	DMSO	1	1	4
	GGGAAAGTTGGGAAAG TCAGCA	689.	AGGACAAGCTACCCACA CC	1000	DMSO	1	3	2
	TGGTGCATCAAAGGGTT GCTTCT	690.	TCATTCCACGACGCCGG AG	1001	DMSO	0	3	3
	CCCAGGCTGCCATCAC ACT	691.	TGGAGTAAGTACCTTG GGGACCT	1002	DMSO	1	3	2
	TCAGTGCCTGGTCC TCA	692.	TGTGCAAATACCTAGCAC GGTGC	1003	DMSO	4	2	0
	AGCACTCCCTTGAATT TTGGTGC	693.	ACTGAAGTCCAGCCTCTT CCATTTC	1004	DMSO	2	1	3
	GAACCCGGCTCTGGT CCA	694.	GGGGAGTAGAGGCTAGTG TTGCC	1005	DMSO	2	0	4
	TTGGGGTCCCTGTGGA GTC	695.	AGGTGCCGTGTTGCC AA	1006	DMSO	1	2	3
		696.		1007				
	GCCCTACATCTGCTCTCC CTCCA	697.	GGGCCGGAAAGAGTTGC TG	1008	DMSO			
	TTGGAGTGTGCCGG TTG	698.	ACCTCTCTCTCTGCC ACTGT	1009	DMSO	0	1	1
	CACACCATGCTGATCCAG GC	699.	GCAGTACGGAGCACGAA GC	1010	DMSO	1	1	1
	CTCCAGGGCTCGCTGTC CAC	700.	CTGGGCTCTGCTGGTCC CC	1011	DMSO	0	2	1
	CTGTGGTAGCCGTGGCC AGG	701.	CCCCATACACCTCTCCG GGA	1012	DMSO	0	2	1
	GGTGGGGACTGAAT GAG	702.	CCAGCGTTCCAAGGG AT	1013	betaín a a 1M, TD	0	1	2
GGAATCCCTCTCCAGCCCCCTGG GGAATCCCTCTCCAGCCTCTGG (SEQ ID NO:432)	CCAGAGTGGGGCCCTG TGA	703.	TTTCCACACTCAGTTCTGC AGGA	1014	DMSO	1	1	1 / 2
GGAATCTCTCTGGCATCTGG (SEQ ID NO:433)	TGTGACTGGTTGCTGC TTCT	704.	GCAGTGTGTTGTTGAT GGGCA	1015	betaín a a 1M, TD	0	1	5
	CTGGCCAAGGGTGAGT GGG	705.	TGGGACCCAGCAGCCAA TG	1016	DMSO	1	0	2
	ACGGTGTGCTGGCTGCT CTT	706.	ACAGTGTGACCGTGCTG GG	1017	DMSO	1	1	1
	TGGTTGGCCTCAGGG ATGG	707.	TGCCCTCCACAAAAATGTC TACCT	1018	DMSO	0	0	3
	TGGTTGGCCTCAGGG ATGG	708.	ACCCCTTACCCAGAAC CATGA	1019	DMSO	0	0	3
	TCCAAGTCAGCGATGAG GGCT	709.	TGGGAGCTGTTCTTTG GCCA	1020	DMSO	0	3	0
	CACCCCTCTCAGCTCCC AA	710.	GCTAGAGGGTCTGCTGCC TT	1021	DMSO	1	2	0
	AGACCCCTTGCCAAGC ACA	711.	CTTGCTCTCACCCGCCT CC	1022	DMSO	2	1	0
	ACATGTGGAGGGGAC AGA	712.	TCTCACTTGTGTTACCG ATGTCG	1023	DMSO	0	1	3

TABELA B

Alvo real em células U2OS.EGFP	Iniciador de PCR direto	SEQ ID N O	Iniciador de PCR reverso	SEQ ID N O:	Condições de PCR	Transformações de Watson-Crick	transformações que não Watson-Crick	Transformações
	GGACGACTGTGCCTGGACA	713.	AGTGCCAGAGTGTGTA ACTGCT	1024	72C Recozer, DMSO a 3%	0	1	3
	GGAGAGCTCAGGCCAGGTC	714.	CAGCGTGGCCGTGGGATA	1025	DMSO	1	1	2
	GCTGAAGTGCTCTGGGTTGCT	715.	ACCCCAGTGATGAATTGGTAC	1026	DMSO	1	1	2
	TGGGGTGCACATGGCCATC	716.	TTGCCTCGCAGGGAAAGCAG	1027	DMSO	0	1	3
	CTCGGGGAGGCCAACACCT	717.	AGCCACCAACACATACCA GGCT	1028	DMSO	2	0	2
	GCATGCCCTTAATCCGGCT	718.	AGGATTTAGAGTGATGG GGCT	1029	DMSO	2	1	1
	CGCCAGGCCAACAGTCAT	719.	GCAAATTCTGCACCTACT CTAGGCCT	1030	DMSO	1	1	2
	AGCTCACAAAGAATTGGAG GTAACAGT	720.	GCAGTCACCCCTCACTGC CTGT	1031	DMSO	1	1	2
	AAACTGGGCTGGGCTTC CGG	721.	GGGGCTAAGGCATTGTCA GACCC	1032	DMSO	2	0	2
	GCAGGTAGGCAGTCTGG GGC	722.	TCTCCTGCCTCAGCCTCC CA	1033	betaín a a 1M, TD	1	2	1
	GCAGGTAGGCAGTCTGG GGC	723.	TCTCCTGCCTCAGCCTCC CA	1034	betaín a a 1M, TD	1	2	1
	GCAGGTAGGCAGTCTGG GGC	724.	TCTCCTGCCTCAGCCTCC CA	1035	betaín a a 1M, TD	1	2	1
	GCTCTGGGTAGAAGGAGC	725.	GGCCTGTCAACCAACCAA CC	1036	DMSO	2	2	0
	TGACATGTGTGCTGG GC	726.	AAATCCTGCAGCCTCCCC TT	1037	DMSO	0	2	2
	T CCTGGTAGATCGTCCA CAGGA	727.	TCCTCCCCACTCAGCCTC CC	1038	DMSO	0	3	1
	T CCTTAATCCAAGTCC GTTAGACA	728.	AGGGACCAGCCACTACCC TTCA	1039	DMSO	2	2	0
	GGGACACCAGTCCTTC CAT	729.	GGGGGAGATTGGAGTTCC CC	1040	DMSO	1	0	4
	ACACCACTATAAGGCAG AGTAGGT	730.	TCTGCCTGGGTGCTTC CC	1041	DMSO	1	1	3
	CTGGGAGCGGAGGGAA TGC	731.	GCCCCGACAGATGAGGCC TC	1042	DMSO	1	2	2
CAGATTACTGCTGCAGCACCGGG (SEQ ID NO:434)	CGGGTCTGGAATGCCT CCA	732.	ACCCAGGAATTGCCACCC CC	1043	DMSO	1	2	3
	TTGCTGTGGTCCCGGT GTG	733.	GCAGACACTAGAGCCGC CC	1044	DMSO	3	2	0
	GGTGTGGTAGGGT GGT	734.	ACCTGCCTCTGTGCTG CA	1045	DMSO	2	3	0
	CTCCCAGGACAGTGCTC GGC	735.	CCTGGCCCCATGCTGCCT G	1046	DMSO	2	2	1
	TGGTAGGTGGCTCT GTGA	736.	AGGAATGATGTTCCAC CCCCT	1047	DMSO	2	3	0
	CTCCGCAGGCCACCGTG GTA	737.	TGCATTGACGTACGATGG CTCA	1048	DMSO	1	3	1
	ACCTTGAGCATGA CGCA	738.	ACCTTGAGCAACATGACTC ACCTGG	1049	DMSO	2	1	2
ACACAAACTCTGCAGCACCTGG ACACAAACTCTGCAGCACGTGG (SEQ ID NO:435)	TCTCCAGTTCTGCTCT CATGG	739.	ACCATTGGTAACCTGGAC CA	1050	betaín a a 1M, TD	3 / 2	3	1
	TGGGGTGGTGGCTTG ATCCA	740.	TCAGCTATAACCTGGAC TTGTGCT	1051	DMSO	2	1	3
	AGCAGCCAGTCAGTGT CCTG	741.	CCCTTTCATCGAGAACCC CAGGG	1052	DMSO	3	1	2
	TGGACGCTGCTGGGAGG AGA	742.	GAGGTCTCGGCTGCTCG TG	1053	DMSO	0	3	3
	AGGTTTGACTCTGTTGC CTGG	743.	TGGGGTGATTGGTTGCCA GGT	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
	TCTTCCTTGCCAGGCAG CACA	744.	TGCAGGAATAGCAGGTAT GAGGGT	1055	DMSO	4	0	2
	GGACGCCTACTGCCCTGG ACC	745.	GCCCTGGCAGCCCAGGT AC	1056	DMSO	3	0	3
	AGGCAGTCATGCCCTG CTA	746.	GGTCCCACCTCCCCCTAC AA	1057	DMSO	2	3	1
	Não otimizado					3	1	2
	CCCCAGCCCCCACCGT 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. Ensaios 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克lonados 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 (**Enhanced Green Fluorescent Protein - 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 nuclease 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 (**N**on-**H**omologous **E**nd-**J**oining - NHEJ) de rupturas fita dupla induzidas por nuclease (**D**ouble-**S**tranded **B**reak - 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 GATGCCGTTCTCTGCTTGTGCGG
 (SEQ ID NO:10)

Sítio 3 de EGFP GGTGGTGCAGATGAACCTTCAGGG
 (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 exibia 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 evidenciar de mutações indel mediadas por NHEJ induzidas por nuclease 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	GGGTGGGGGGAGTTGCTCCCTGG	1059.	26,0 ± 2,9	10,5 ± 0,07	3,33 ± 0,42	VEGFA
	OT1-3	GGATGGAGGGAGTTGCTCCCTGG	1060.	25,7 ± 9,1	18,9 ± 0,77	2,93 ± 0,04	IGDCC3
	OT1-4	GGGAGGC<u>I</u>GGAGTTGCTCCCTGG	1061.	9,2 ± 0,8	8,32 ± 0,51	N.D.	LOC116437
	OT1-6	CGGGGGAGGGAGTTGCTCCCTGG	1062.	5,3 ± 0,2	3,67 ± 0,09	N.D.	CACNA2D
	OT1-11	GGGGAGGGAGTTGCTCCCTGG	1063.	17,1 ± 4,7	8,54 ± 0,16	N.D.	
Alvo 2 (Sítio 2 de VEGFA)	T2	GACCCCTCCACCCGCCCTCCGG	1064.	50,2 ± 4,9	38,6 ± 1,92	15,0 ± 0,25	VEGFA
	OT2-1	GACCCCC<u>C</u>CCACCCGCCCTCCGG	1065.	14,4 ± 3,4	33,6 ± 1,17	4,10 ± 0,05	FMN1
	OT2-2	GGGCCCC<u>T</u>CCACCCGCCCTCGG	1066.	20,0 ± 6,2	15,6 ± 0,30	3,00 ± 0,06	PAX6
	OT2-6	CTACCCCTCCACCCGCCCTCGG	1067.	8,2 ± 1,4	15,0 ± 0,64	5,24 ± 0,22	PAPD7
	OT2-9	GCCCC<u>C</u>ACCCACCCGCCCTCGG	1068.	50,7 ± 5,6	30,7 ± 1,44	7,05 ± 0,48	LAMA3
	OT2-15	TACCCCC<u>A</u>CACCCGCCCTCGG	1069.	9,7 ± 4,5	6,97 ± 0,10	1,34 ± 0,15	SPNS3
	OT2-17	ACACCCCC<u>C</u>ACCCGCCCTCAGG	1070.	14,0 ± 2,8	12,3 ± 0,45	1,80 ± 0,03	
	OT2-19	AT<u>I</u>CCCCCCCACCCGCCCTCAGG	1071.	17,0 ± 3,3	19,4 ± 1,35	N.D.	HDLBP
	OT2-20	CCCC<u>A</u>CCCCACCCGCCCTCAGG	1072.	6,1 ± 1,3	N.D.	N.D.	ABLIM1
	OT2-23	CGCCCC<u>T</u>CCACCCGCCCTCGG	1073.	44,4 ± 6,7	28,7 ± 1,15	4,18 ± 0,37	CALY
	OT2-24	CTCCCC<u>A</u>CCCCGCCCTCAGG	1074.	62,8 ± 5,0	29,8 ± 1,08	21,1 ± 1,68	
	OT2-29	T<u>G</u>CCCC<u>T</u>CCACCCGCCCTCGG	1075.	13,8 ± 5,2	N.D.	N.D.	ACLY
	OT2-34	AGGCCCC<u>C</u>ACCCGCCCTCAGG	1076.	2,8 ± 1,5	N.D.	N.D.	
Alvo 3 (Sítio 3 de VEGFA)	T3	GGTAGTGAGTGTTGCGTGTGG	1077.	49,4 ± 3,8	35,7 ± 1,26	27,9 ± 0,52	VEGFA
	OT3-1	GGTAGTGAGTG <u>T</u> GTGAGG	1078.	7,4 ± 3,4	8,97 ± 0,80	N.D.	(abParts)
	OT3-2	AGTAGTGAGTG<u>T</u>GTGGGG	1079.	24,3 ± 9,2	23,9 ± 0,08	8,9 ± 0,16	MAX
	OT3-4	G<u>T</u>GTAGTGAGTG<u>T</u>ATGCGTGTGG	1080.	20,9 ± 11,8	11,2 ± 0,23	N.D.	
	OT3-9	GGTAGTGAGTG <u>C</u> GTGCG <u>G</u> GTGG	1081.	3,2 ± 0,3	2,34 ± 0,21	N.D.	TPCN2
	OT3-17	G<u>T</u>GTAGTG<u>A</u>ATG<u>T</u>GTGCGTGAGG	1082.	2,9 ± 0,2	1,27 ± 0,02	N.D.	SLIT1
	OT3-18	T<u>G</u>TGG<u>T</u>GAGTG<u>T</u>GTGCGTGAGG	1083.	13,4 ± 4,2	12,1 ± 0,24	2,42 ± 0,07	COMDA
	OT3-20	AG<u>A</u>GAGTGAGTG<u>T</u>GTG<u>C</u>ATGAGG	1084.	16,7 ± 3,5	7,64 ± 0,05	1,18 ± 0,01	
Alvo 4 (EMX1)	T4	GAGTCGAGCAGAAGAAGAAGGG	1085.	42,1 ± 0,4	26,0 ± 0,70	10,7 ± 0,50	EMX1
	OT4-1	GAG<u>T</u>AGAGCAGAAGAAGAAGGG	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	GGAA <u>T</u> CCCTCTGCAGCACCAAGG	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 transfetada, 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 transfetadas, enquanto que a atividade de RGN no sítio alvo #2 caiu imediatamente quando a quantidade de plasmídeo expressando gRNA transfetado foi diminuída (**Figura 3E** (superior)).

[0132] A quantidade de plasmídeo que codifica Cas9 (variando de 50 ng a 750 ng) transfetado 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 transfetado 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 transfetado (**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 transfeção. 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, reconhecimento 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 a seguintes iniciadores: iniciador senso D10A de Cas9 5'-tggataaaaagtattctattggtagccatcgactaattccg-3' (SEQ ID NO: 1089); iniciador antissenso D10A de Cas9 5'-cgaaattagtgccatggctaaaccaatagaatacttttatcca-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																									
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')	S E Q ID N O :	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 CAGCTTGGCGGG	1091	AAAACGGGCAA GCTGCCGTGCG	1180.		
		G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	c	ACACCGCACGGG CAGCTTGGCGGG	1092	AAAACGGGCAA GCTGCCGTGCG	1181.		
		G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	c	G	ACACCGCACGGG CAGCTTGGCGGG	1093	AAAACGGGCAA GCTGCCGTGCG	1182.		
		G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	g	G	G	ACACCGCACGGG CAGCTTGGCGGG	1094	AAAACCCGCAA GCTGCCGTGCG	1183.		
		G	C	A	C	G	G	G	C	A	G	C	T	T	G	g	C	G	G	ACACCGCACGGG CAGCTTGGCGGG	1095	AAAACCCGCAA GCTGCCGTGCG	1184.		
		G	C	A	C	G	G	G	C	A	G	C	T	T	c	C	C	G	G	ACACCGCACGGG CAGCTTGGCGGG	1096	AAAACCCGGA GCTGCCGTGCG	1185.		
		G	C	A	C	G	G	G	C	A	G	C	T	a	G	C	C	G	G	ACACCGCACGGG CAGCTTGGCGGG	1097	AAAACCCGCTA GCTGCCGTGCG	1186.		
		G	C	A	C	G	G	G	C	A	G	C	a	T	G	C	C	G	G	ACACCCACGGG CAGCATGCCGGG	1098	AAAACCCGGCAT GCTGCCGTGCG	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

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	ACACCCGACGGG	CACCTTGCCTGGG	1100	AAAACCCGGCAA	GGTGCCCGTGC	1189.
	G	C	A	C	G	G	G	C	t	G	C	T	T	G	C	C	G	G	ACACCCGACGGG	CTGCTTGCCTGGG	1101	AAAACCCGGCAA	GCAGCCCGTGC	1190.
	G	C	A	C	G	G	G	g	A	G	C	T	T	G	C	C	G	G	ACACCCGACGGG	GAGCTTGCCTGGG	1102	AAAACCCGGCAA	GCTCCCCGTGCG	1191.
	G	C	A	C	G	G	g	c	C	A	G	C	T	T	G	C	C	G	ACACCCGACGGC	CAGCTTGCCTGGG	1103	AAAACCCGGCAA	GCTGGCCGTGC	1192.
	G	C	A	C	G	c	G	G	C	A	G	C	T	T	G	C	C	G	ACACCCGACGGC	CAGCTTGCCTGGG	1104	AAAACCCGGCAA	GCTGCGCGTGC	1193.
	G	C	A	C	c	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCCGACCGG	CAGCTTGCCTGGG	1105	AAAACCCGGCAA	GCTGCCGGTGC	1194.
	G	C	A	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGCAGGGG	CAGCTTGCCTGGG	1106	AAAACCCGGCAA	GCTGCCCTGTGCG	1195.
	G	C	t	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGCTCGGG	CAGCTTGCCTGGG	1107	AAAACCCGGCAA	GCTGCCCGAGC	1196.
	G	g	A	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGGACGGG	CAGCTTGCCTGGG	1108	AAAACCCGGCAA	GCTGCCCGTCCG	1197.
	G	C	A	C	G	G	G	C	A	G	C	T	T	G	C	C	c	c	ACACCGCACGGG	CAGCTTGCCTGGG	1109	AAAACGGGGCAA	GCTGCCCGTGC	1198.
	G	C	A	C	G	G	G	C	A	G	C	T	T	G	g	g	G	G	ACACCCGACGGG	CAGCTTGGGGGG	1110	AAAACCCCCCAA	GCTGCCCGTGC	1199.
	G	C	A	C	G	G	G	C	A	G	C	T	a	c	C	C	G	G	ACACCGCACGGG	CAGCTACCCGGG	1111	AAAACCGGGTAA	GCTGCCCGTGC	1200.
	G	C	A	C	G	G	G	C	A	G	g	a	T	G	C	C	G	G	ACACCGCACGGG	CAGGATGCCGGG	1112	AAAACCCGGCAT	CCTGCCCGTGC	1201.
	G	C	A	C	G	G	G	C	t	c	C	T	T	G	C	C	G	G	ACACCGCACGGG	CTCCTTGCCTGGG	1113	AAAACCCGGCAA	GGAGCCCGTGC	1202.
	G	C	A	C	G	G	g	A	G	C	T	T	G	C	C	G	G	G	ACACCGCACGGC	GAGCTTGCCTGGG	1114	AAAACCCGGCAA	GCTCGCGTGC	1203.
	G	C	A	C	c	c	G	C	A	G	C	T	T	G	C	C	G	G	ACACCCGACCCG	CAGCTTGCCTGGG	1115	AAAACCCGGCAA	GCTGGGGTGC	1204.
	G	C	t	g	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGCTGGGG	CAGCTTGCCTGGG	1116	AAAACCCGGCAA	GCTGCCCAAGCG	1205.
	G	g	t	C	G	G	G	C	A	G	C	T	T	G	C	C	G	G	ACACCGGTCTGGG	CAGCTTGCCTGGG	1117	AAAACCCGGCAA	GCTGCCCGACCG	1206.
Sítio alvo 2 de EGFP																								
2	1	1	1	1	1	1	1	1	1	1	1	1	1	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	0	1	1118	AAAACACAAGCA	GAAGAACCGGG
	G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	T	ACACCGCGTTC	TTCTGCTTGTG				1207.
	G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	G	A	A	ACACCGCGTTC	TTCTGCTTGTG	1119	AAAACTCAAGCA	GAAGAACCGGG	1208.
	G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	T	c	T	ACACCGCGTTC	TTCTGCTTGTG	1120	AAAACAGAAGCA	GAAGAACCGGG	1209.
	G	C	C	G	T	T	C	T	T	C	T	G	C	T	a	G	T	T	ACACCGCGTTC	TTCTGCTAGTG	1121	AAAACATAGCA	GAAGAACCGGG	1210.
	G	C	C	G	T	T	C	T	T	C	T	G	C	a	T	G	T	T	ACACCGCGTTC	TTCTGCTATGTG	1122	AAAACACATGCA	GAAGAACCGGG	1211.
	G	C	C	G	T	T	C	T	T	C	T	G	g	T	T	G	T	T	ACACCGCCGTT	TTCTGGTTGTG	1123	AAAACACAACCA	GAAGAACCGGG	1212.
	G	C	C	G	T	T	C	T	T	C	T	c	C	T	T	T	G	T	ACACCGCCGTT	TTCTGCTTGTG	1124	AAAACACAGGA	GAAGAACCGGG	1213.
	G	C	C	G	T	T	C	T	T	C	a	G	C	T	T	T	G	T	ACACCGCGTTC	TTCAGCTTGTG	1125	AAAACACAAGCT	GAAGAACCGGG	1214.
	G	C	C	G	T	T	C	T	T	g	T	G	C	T	T	T	G	T	ACACCGCGTTC	TTGTGCTTGTG	1126	AAAACACAAGCA	CAAGAACCGGG	1215.
	G	C	C	G	T	T	C	T	a	C	T	G	C	T	T	T	G	T	ACACCGCGTTC	TACTGCTTGTG	1127	AAAACACAAGCA	GTAGAACCGGG	1216.
	G	C	C	G	T	T	C	a	T	C	T	G	C	T	T	T	G	T	ACACCGCCGTT	ATCTGCTTGTG	1128	AAAACACAAGCA	GATGAACCGGG	1217.
	G	C	C	G	T	T	g	T	T	C	T	G	C	T	T	T	G	T	ACACCGCGTTC	TTCTGCTTGTG	1129	AAAACACAAGCA	GAACAACCGGG	1218.
	G	C	C	G	T	a	C	T	T	C	T	G	C	T	T	T	G	T	ACACCGCCGTT	TTCTGCTTGTG	1130	AAAACACAAGCA	GAAGTACCGGG	1219.
	G	C	C	G	a	T	C	T	T	C	T	G	C	T	T	T	G	T	ACACCGCCGATC	TTCTGCTTGTG	1131	AAAACACAAGCA	GAAGATCGGG	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	AAAACACAAGCA GAAGAAGGGCG	1221.	
		G	C	g	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGCCGTT TTCTGCTTGTG	1133	AAAACACAAGCA GAAGAACCGCG	1222.	
		G	g	C	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGCCGTT TTCTGCTTGTG	1134	AAAACACAAGCA GAAGAACCGCG	1223.	
		G	C	C	G	T	T	C	T	T	C	T	G	C	T	T	c	a	ACACCGCCGTT TTCTGCTTGTG	1135	AAAACTGAAGCA GAAGAACGGCG	1224.	
		G	C	C	G	T	T	C	T	T	C	T	G	C	a	a	G	T	ACACCGCCGTT TTCTGCTTGTG	1136	AAAACACTTGCA GAAGAACGGCG	1225.	
		G	C	C	G	T	T	C	T	T	C	T	c	g	T	T	G	T	ACACCGCCGTT TTCTGCTTGTG	1137	AAAACACAACGA GAAGAACGGCG	1226.	
		G	C	C	G	T	T	C	T	T	T	g	a	G	C	T	T	G	T	ACACCGCCGTT TTGAGCTTGTG	1138	AAAACACAAGCT CAAGAACGGCG	1227.
		G	C	C	G	T	T	C	a	a	C	T	G	C	T	T	G	T	ACACCGCCGTT AACTGCTTGTG	1139	AAAACACAAGCA GTTAACCGCG	1228.	
		G	C	C	G	T	a	g	T	T	C	T	G	C	T	T	G	T	ACACCGCCGTT TTCTGCTTGTG	1140	AAAACACAAGCA GAACTACGGCG	1229.	
		G	C	C	c	a	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGCCCATC TTCTGCTTGTG	1141	AAAACACAAGCA GAAGATGGCG	1230.	
		G	g	g	G	T	T	C	T	T	C	T	G	C	T	T	G	T	ACACCGGGGTT TTCTGCTTGTG	1142	AAAACACAAGCA GAAGAACCCCG	1231.	
Sítio alvo 3 de EGFP																							
2	1	1	1	1	1	1	1	1	1	1	1	1	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')	
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 2 (5' a 3')			
		G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGTGCAG ATGAACCTTCAG	1143	AAAACTCTAGTT ATCTGCACCG	1232.	
		G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	C	t	ACACCGGTGCAG ATGAACCTTCAG	1144	AAAACTCAAGTT CATCTGCACCG	1233.	
		G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	g	A	ACACCGGTGCAG ATGAACCTTCAG	1145	AAAACTGTAGTT CATCTGCACCG	1234.	
		G	G	T	G	C	A	G	A	T	G	A	A	C	T	a	C	A	ACACCGGTGCAG ATGAACCTACAG	1146	AAAACTGATGTT CATCTGCACCG	1235.	
		G	G	T	G	C	A	G	A	T	G	A	A	C	a	T	C	A	ACACCGGTGCAG ATGAACATCAG	1147	AAAACTGAACTT CATCTGCACCG	1236.	
		G	G	T	G	C	A	G	A	T	G	A	A	g	T	T	C	A	ACACCGGTGCAG ATGAAGTTTCAG	1148	AAAACTGAAGAT CATCTGCACCG	1237.	
		G	G	T	G	C	A	G	A	T	G	A	A	t	C	T	T	C	ACACCGGTGCAG ATGATCTTCAG	1149	AAAACTGAAGTA CATCTGCACCG	1238.	
		G	G	T	G	C	A	G	A	T	G	t	A	C	T	T	C	A	ACACCGGTGCAG ATGTAACCTTCAG	1150	AAAACTGAAGTT GATCTGCACCG	1239.	
		G	G	T	G	C	A	G	A	T	c	A	A	C	T	T	C	A	ACACCGGTGCAG ATCAACTTCAG	1151	AAAACTGAAGTT CTTCTGCACCG	1240.	
		G	G	T	G	C	A	G	A	a	G	A	A	C	T	T	C	A	ACACCGGTGCAG AAGAACATTTCAG	1152	AAAACTGAAGTT CAACTGCACCG	1241.	
		G	G	T	G	C	A	G	t	T	G	A	A	C	T	T	C	A	ACACCGGTGCAG TTGAACCTTCAG	1153	AAAACTGAAGTT CATGTGCACCG	1242.	
		G	G	T	G	C	A	c	A	T	G	A	A	C	T	T	C	A	ACACCGGTGCAGC ATGAACCTTCAG	1154	AAAACTGAAGTT CATCAGCACCG	1243.	
		G	G	T	G	C	t	G	A	T	G	A	A	C	T	T	C	A	ACACCGGTGCTG ATGAACCTTCAG	1155	AAAACTGAAGTT CATCTCCACCG	1244.	
		G	G	T	G	g	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGTGGAG ATGAACCTTCAG	1156	AAAACTGAAGTT CATCTGGACCG	1245.	
		G	G	T	c	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGTCCAG ATGAACCTTCAG	1157	AAAACTGAAGTT CATCTGCTCCG	1246.	
		G	G	a	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGAGCAG ATGAACCTTCAG	1158	AAAACTGAAGTT CATCTGCGACCG	1247.	
		G	c	T	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCTCGAG ATGAACCTTCAG	1159	AAAACTGAAGTT CATCTGCAGCG	1248.	
		G	G	T	G	C	A	G	A	T	G	A	A	C	T	T	g	t	ACACCGGTGCAG ATGAACCTTCAG	1160	AAAACTGAAGTT CATCTGCACCG	1249.	
		G	G	T	G	C	A	G	A	T	G	A	A	C	a	a	C	A	ACACCGGTGCAG ATGAACCTTCAG	1161	AAAACTGTGTTTC ATCTGCACCG	1250.	
		G	G	T	G	C	A	G	t	a	G	A	A	C	T	T	C	A	ACACCGGTGCAG TAGAACAAACAG	1162	AAAACTGAACAT CATCTGCACCG	1251.	
		G	G	T	G	C	A	G	A	T	c	t	A	C	T	T	C	A	ACACCGGTGCAG ATCTACTTCAG	1163	AAAACTGAAGTA GATCTGCACCG	1252.	
		G	G	T	G	C	A	G	t	a	G	A	A	C	T	T	C	A	ACACCGGTGCAG TAGAACACTTCAG	1164	AAAACTGAAGTT CTACTGCACCG	1253.	
		G	G	T	G	C	t	c	A	T	G	A	A	C	T	T	C	A	ACACCGGTGCTC ATGAACCTTCAG	1165	AAAACTGAAGTT CATGAGCACCG	1254.	
		G	G	T	c	g	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGGTGAG ATGAACCTTCAG	1166	AAAACTGAAGTT CATCTCGACCG	1255.	
		G	c	a	G	C	A	G	A	T	G	A	A	C	T	T	C	A	ACACCGCAGCAG ATGAACCTTCAG	1167	AAAACTGAAGTT 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	1	1	1	
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')		oligonucleotídeo 2 (5' a 3')							
		G	T	G	G	G	G	G	A	G	T	T	T	G	C	T	C	C	ACACCGTGGGG GAGTTTGCTCCG	1168	AAAACGGAGCAA ACTCCCCCACG	1257.							
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	1	1	1	
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')		oligonucleotídeo 2 (5' a 3')							
		G	A	G	T	G	A	G	T	G	T	G	T	G	C	G	T	G	ACACCGAGTGAG TGTGTGCGTGG	1169	AAAACCACGCAC ACACTCACTCG	1258.							
Alvo 4 endógeno (tru-gRNA ao sitio 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	1	1	1	
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')		oligonucleotídeo 2 (5' a 3')							
		G	T	C	C	G	A	G	C	A	G	A	A	G	A	A	G	A	ACACCGTCCGAG CAGAAGAAGAAG	1170	AAAACTCTCTT CTGCTCGGACG	1259.							
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	1	1	1	
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')		oligonucleotídeo 2 (5' a 3')							
		G	C	A	G	A	T	G	T	A	G	T	G	T	T	T	C	C	A	C	A	ACACCGCAGATG TAGTGTTCACAG	1171	AAAACTGTGAA ACACTACATCG	1260.				
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	1	1	1	
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')		oligonucleotídeo 2 (5' a 3')							
		G	A	T	G	T	A	G	T	G	T	T	T	T	C	C	A	C	A	ACACCGATGTAG TGTTTCCACAG	1172	AAAACTGTGAA ACACTACATCG	1261.						
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	1	1	1	
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')		oligonucleotídeo 2 (5' a 3')							
		T	C	C	C	T	C	T	T	A	G	C	C	A	G	A	G	C	ACACCTCCCTCT TAGCCAGAGCCG	1173	AAAACCGCTCT GGCTAAAGAGGG	1262.							
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	1	1	1	
0	9	8	7	6	5	4	3	2	1	0	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	G	T	A	C	T	T	T	G	T	C	C	ACACCGCCGTT GTACTTGTCTC	1174	AAAACGAGGACA AAGTACAAACGG	1263.							
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	1	1	1	
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')		oligonucleotídeo 2 (5' a 3')							
		G	T	T	T	G	T	A	C	T	T	T	T	G	T	C	C	T	ACACCGTTTGTA CTTTGCTCG	1175	AAAACGAGGACA AAGTACAAACG	1264.							
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	1	1	1	
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')		oligonucleotídeo 2 (5' a 3')							
		G	G	A	A	G	A	C	T	G	A	G	G	C	T	A	C	A	ACACCGGGAAAGA CTGAGGCTACAT	1176	AAAACATGTAG CCTCAGCTTCC	1265.							
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	1	1	1	
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')		oligonucleotídeo 2 (5' a 3')							
		G	A	A	G	A	C	T	G	A	G	G	C	T	A	C	A	T	ACACCGAAGACT GAGGCTACATAG	1177	AAAACATGTAG CCTCAGCTTCC	1266.							
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	1	1	1	
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')		oligonucleotídeo 2 (5' a 3')							
		G	A	G	C	C	C	C	C	A	G	A	G	C	A	G	C	A	ACACCGAGGCC CCAGAGCAGCCA	1178	AAAACGTGGCTG CTCTGGGGCG	1267.							
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	1	1	1	
0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	oligonucleotídeo 1 (5' a 3')		oligonucleotídeo 2 (5' a 3')							
		G	C	C	C	C	C	C	A	G	A	G	C	A	G	C	C	A	ACACGCCCA GAGCAGCCACG	1179	AAAACGTGGCTG CTCTGGGGCG	1268.							

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 (Reyón

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 transfetadas 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 transfeção), 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 EGFP, respectivamente. Dois dias após transfeção, 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, transfeções 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 transfetados em células HEK293 ou U2OS.EGFP usando as seguintes condições: as células U2OS.EGFP foram transfetadas usando as mesmas condições conforme para o ensaio de ruptura de EGFP descrito acima. As células HEK293 foram transfetadas 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 DNAAdvance (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 transfeções 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 DNAAdvance 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 x 10⁵ 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 *BamHI*). 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).

Ensaios de T7EI para quantificar frequências de mutações indel

[0141] Ensaios 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 num 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 esperadas (Esperadas) - HS GRCh37	Desem- parelha				S				S				Tran- sver		Tran- sver	
		S mentos fora do alvo	E em sítios	I sobre o alvo	Alvo real	em células U2OS.EGFP	iniciador PCR direto	I de	E Iniciador PCR reverso	I de	E Condições de PCR	sôes de Wats	sôes não Wats	nsi on-	Cric Cric	nsi on- s	
Alvo 1	GGGTGGGG GGAGTTG CTCCTGG	126 ^t 0			TCCAGATGGC ACATTGTCAG	127 ^c	AGGGAGCAGG AAAGTGAGGT	127 ^r	DMSO								
OT1 -1	GGGTGGGG GGAGTTG CCCCAGG	127 ^t 1			GGGGCCACT CTTCTTCAT	127 ^c	ACCCAGACTC CTGGTGTGGC	127 ^r	No DMSO	0	0	0	1				
OT1 -2	GCGTGGGG GGTGTTC CTCCCCG	127 ^t 2			GCTAACGAGA GATGCCATG CC	127 ^c	ACCACCCTTTC CCCCAGAAA	127 ^r	DMSO	2	0	0	0				
OT1 -3	GGATGGAG GGAGTTG CTCCTGG	127 ^t 2			ACCCCACAGC CAGGTTTCA	127 ^c	GAATCACTGCA CCTGGCCATC	128 ^t	DMSO	0	0	0	2				
OT1 -4	GGGAGGGT GGAGTTG CTCCTGG	128 ^t 2			TGCGGCAACT TCAGACAACC	128 ^c	TAAAGGGCGT GCTGGGAGAG	128 ^r	DMSO	1	1	1	0				
OT1 -5	GGGTGGGT GGAGTTG CTACTGG	128 ^t 2			GCATGTCAGG ATCTGACCCC	128 ^c	TGCAGGGCCA TCTTGTGTGT	128 ^r	DMSO	0	2	0	0				
OT1 -6	CGGGGGAG GGAGTTG CTCCTGG	128 ^t 3			CCACCATG TTCTGGTGC	128 ^c	CTGGGTCTGTT CCCTGTGGG	128 ^r	DMSO	1	1	1	1				
OT1 -7	GAGTGGGT GGAGTTG CTACAGG	129 ^t 3			GGCTCTCCCT GCCCTAGITT	129 ^c	GCAGGTCAAG TTGGAACCCG	129 ^r	DMSO	0	2	1	1				
OT1 -8	GGGAGGGG AGAGTTGT TCCAGG	129 ^t 3			GGGGCTGAGA ACACATGAGA TGCA	129 ^c	AGAGTTGTGCA CTGCTGCCT	129 ^r	DMSO	1	0	2	1				
OT1 -9	GGGAGGGG GCAGGTG CTCCAGG	129 ^t 3			CCCGACCTCC GCTCCAAAGC	129 ^c	GGACCTCTGC ACACCCCTGGC	129 ^r	DMSO	2	1	0	0				
OT1 -10	GGGAGGGG GGAGTGTG TTCCGGG	129 ^t 3			TGCAAGGTGCG CATAGTCCCA	130 ^c	CAGGAGGGGG AAGTGTGTC	130 ^r	DMSO	1	1	1	1				

TABELA E											
OT1	GGGGAGGG			GAGAGCAAGT							
-11	GAAGTTG	130 ^c 3	CTCCTGG	GCCCCATTCTT	130 ^c	TTGTTCCCGAG	130 ^c	DMSO	0	1	2
-12	GGGGGTGG			TTGCAGTGGAA		G					
-13	GGACTTTG	130 ^c 3	CTCCAGG	GCCCCCAGCC	130 ^c	CCTCTGTTTC	130 ^c	DMSO	1	2	0
-14	GGGTGGGG			GGGCTGCCCT	130 ^c	CCCTGAGTCC	131 ^c	72C Reozer, DMSO a 3%	1	2	0
-15	GGAGTGGG	130 ^c 3	CTCCAGG	TGACCCCTGGA		GCTGAGACAA					
-16	GGGTGGCT	131 ^c 3	CTGCTGG	GTACAAAATG	131 ^c	CCAGCCCAGCT	131 ^c	72C Reozer, DMSO a 3%	2	1	0
-17	GGGTGGGG			TTCCCA							
-18	GGTGAGCTG	131 ^c 3	CTCCAGG	AACTCAGGAC		CCCAGGAGCA					
-19	GGAGCTTG	131 ^c 3	CTCCAGG	AACACTGCCT	131 ^c	GGGTACAATG	131 ^c	DMSO	0	1	2
-20	GTGTGGGT			GT							
-21	GGCGTTTG	132 ^c 3	CTCCAGG	TCCTCCTTGG	132 ^c	AGAGGGGCC	132 ^c	DMSO	0	3	0
-22	AGGTGGTG			CCGAGGGCAT	132 ^c	GGGCTGCTGCG	132 ^c	DMSO	0	1	3
-23	GGAGTTG	132 ^c 4	CTCCAGG	GGGCAATCCT		AGTTGCCAAC					
-24	AGTTGGGG			TGCTTGCAT	132 ^c	GGGGTCTCAG	132 ^c	DMSO	0	2	2
-25	GGAGTTTG	132 ^c 4	CTCCAGG	ACA		CCCTCTGTGT					
-26	ATGTGTGG			AGCTCCTCT		CACAGAAGGA					
-27	GGAATTG	132 ^c 4	CTCCAGG	CATTCTCTTC	133 ^c	TGTGTGCAGG	133 ^c	DMSO	0	2	2
-28	GGAGCTTT	133 ^c 4	CTCCAGG	TGCTGT		TT					
-29	CAGTGGGG			AGCAGACACA		GGTCAGGTGT					
-30	GGAGCTTT	133 ^c 4	CTCCAGG	GGTGAATGCT	133 ^c	GCTGCTAGGC	133 ^c	DMSO	1	1	2
-31	GAGGGGGA			GCT		A					
-32	GCAGTTTG	133 ^c 4	CTCCAGG	CCTGTGGGGC	133 ^c	TCTCAGGTGC	133 ^c	Sem TD	1	1	2
-33	GGAGGAGG			AGCTGCAGT	133 ^c	GGCGGGTAA	134 ^c	DMSO	0	1	3
-34	GGAGTCTG	133 ^c 4	CTCCAGG	GGGAATGAGT		TAGCTGGCTT					
-35	GGAGGGGG			AGCTGCAGT	133 ^c	TGCCGGTAA	134 ^c	DMSO	0	1	3
-36	GGCTTTGC	134 ^c 4	CTCCAGG	GGGAATGAGT		TGCTGGCTT					
-37	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-38	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-39	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-40	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-41	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-42	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-43	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-44	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-45	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-46	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-47	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-48	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-49	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-50	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-51	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-52	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-53	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-54	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-55	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-56	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-57	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-58	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-59	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-60	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-61	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-62	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-63	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-64	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-65	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-66	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-67	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-68	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-69	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-70	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-71	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-72	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-73	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-74	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-75	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-76	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-77	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-78	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-79	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-80	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-81	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-82	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-83	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-84	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-85	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-86	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-87	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-88	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-89	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-90	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-91	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-92	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-93	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-94	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-95	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-96	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-97	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-98	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-99	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-100	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-101	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-102	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-103	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-104	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-105	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-106	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-107	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-108	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-109	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-110	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-111	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-112	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-113	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-114	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-115	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-116	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	0	1	3
-117	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO	1	1	2
-118	GGAGGGGG			AGCTGCAGT	133 ^c	AAAGTGGGTG	133 ^c	DMSO</td			

TABELA E										
OT1	AAGTAAGG GAAGTTG -35 CTCCTGG	137 ^c 5	TGCCTGGCAC GCAGTAGGTG	137 ^c	GGGAAGGGGG AACAGGTGCA	137 ^c	DMSO	0	0	5
OT1	AGAAGAGG GGATTTGC -36 TCCTGG	137 ^c 5	Não otimizado					1	1	3
OT1	ATCTGGGG TGATTTGC -37 TCCTGG	137 ^c 5	ACCTGGGCTT GCCACTAGGG	137 ^c	GCTGCTCGCA GTTAAGCACCA	138 ^c	DMSO	1	3	1
OT1	CTCTGCTG GGAGTTG -38 CTCCTGG	138 ^c 5	GTGGCCGGG CTACTGCTAC C	138 ^c	GGTTCACAAA GCTGGGGCA	138 ^c	DMSO	3	2	0
OT1	CTGGTGGG GGAGCTTG -39 CTCCAGG	138 ^c 5	Não otimizado					1	3	1
OT1	CTTTCGGG GGAGTTG -40 CGCCGGG	138 ^c 5	GCAAGAGGGCG GAGGAGACCC	138 ^c	AGAGTCATCCA TTTCTGGGG GC	138 ^c	DMSO	2	3	0
OT1	CTTTGGGG TTAGTTGC -41 TCCTGG	138 ^c 5	GGGGTCAGTG GTGATATCCC CCT	138 ^c	AGGGAAATCCTT TTCCATTGCT TGT	139 ^c	betaína a 1M, TD	1	4	0
OT1	GCTCTGGG GTAGTTGC -42 TCCAGG	139 ^c 5	AGAGAGGCCA CGTGGAGGGT	139 ^c	GCCTCCCCTC CTCCCTCCCA	139 ^c	DMSO	1	3	1
OT1	GTCTCTCG GGAGTTG -43 CTCCGGG	139 ^c 5	GACAGTGCT TGCATGCAC	139 ^c	TCTGACCGGT TGCCCTGACG	139 ^c	DMSO	3	2	0
OT1	TCCTGAGG GCAGTTG -44 CTCCAGG	139 ^c 5	TGTGTGAACG CAGCCTGGCT	139 ^c	TGGTCTAGTAC TTCCCTCCAGCC TT	139 ^c	DMSO	3	1	1
OT1	TCTTGGGA GAGTTGCT -45 CCAGG	140 ^c 5	GGTTCTCCCT TGGCTCTGT GA	140 ^c	CCCACTGCTC CTAGCCCTGC	140 ^c	DMSO	1	3	1
OT1	ACAACCTG GGAGTTG -46 CTCCTGG	140 ^c 6	TGAAGTCAAC AATCTAAGCTT CCACCT	140 ^c	AGCTTTGGTAG TTGGAGTCCTT GAAGG	140 ^c	DMSO	3	1	2
OT1	ACAAGGTG GAAGTTG -47 CTCCTGG	140 ^c 6	TGATTGGGCT GCAGTTCATG TACA	140 ^c	GCACAGCCTG CCCTTGGAAAG	140 ^c	DMSO	2	1	3
OT1	ACATAGAA GGAGTTG -48 CTCCAGG	140 ^c 6	TCCATGGGCC CCTCTGAAAG A	141 ^c	AGCGCCTCT GCTTCTGCGA	141 ^c	DMSO	1	0	5
OT1	AGACCCAG GGAGTTG -49 CTCCCGG	141 ^c 6	GCGGTTGGTG GGGTTGATGC	141 ^c	GAGTTCCCT CCCGCCAGT	141 ^c	DMSO	2	0	4
OT1	AGACCCAG GGAGTTG -50 CTCCCGG	141 ^c 6	AGGCAAGATT TTCCAGTGTG CAAGA	141 ^c	GCTTTGCCTG GGACTCCCGC	141 ^c	DMSO	2	0	4
OT1	CACGGAGG GGTGTGG -51 CTCCCTGG	141 ^c 6	GCTGCTGGTC GGGCTCTGTG	141 ^c	GCTCTGCTCCA CTTCCCCCTGG	142 ^c	Sem TD	3	1	2
OT1	CAGAGCTT GGAGTTG -52 CTCCAGG	142 ^c 6	GCTGCGAGGC TTCCGTGAGA	142 ^c	CGCCCCCTAGA GCTAAGGGGG T	142 ^c	DMSO	3	2	1
OT1	CTATTGATG GAGTTGCT -53 CCTGG	142 ^c 6	CCAGGAGCCT GAGAGCTGCC	142 ^c	AGGGCTAGGA CTGCAGTGAG C	142 ^c	DMSO	1	3	2
OT1	CTTTCTAGG GAGTTGCT -54 CCTGG	142 ^c 6	CTGTGCTCAG CCTGGGTGCT	142 ^c	GCCTGGGCT GTGAGTAGTTT	142 ^c	DMSO	2	3	1
OT1	GCCATGCT GGAGTTG -55 CTCCAGG	143 ^c 6	AGCTCGGCC AGATCTGTGG	143 ^c	ACTTGGCAGG CTGAGGCAGG	143 ^c	72C Recazer, DMSO a 3%	4	2	0
Alvo 2										
GACCCCGT CCACCCCG -1 CCCCCG										
GACCCCCC CCACCCCG -1 CCCCCG										
GGGCCCCC CCACCCCG -2 CCTCTGG										
AACCCCAT CCACCCGG -3 CCTCAGG										

TABELA E

OT2 -4	CACCCCT CAACACCG CCTCAGG	144 ^c 3	GCTCCCGATC CTCTGCCACC	144 ^c	GCAGCTCCA CCACCCCTCAG	145 ^c	DMSO	1	2	0
OT2 -5	CACCCCT CCCCTCCG CCTCAGG	145 ^c 3	GGGGACAGG CAGGCAAGGA G	145 ^c	GTGCGTGTC GTTCACCCCT	145 ^c	DMSO	1	1	1
OT2 -6	CTACCCCT CCACCCCG CCTCCGG	145 ^c 3	AAGGGCCTGC TGGTAGGAC	145 ^c	CGTGAATTGA GTTCTGGCA	145 ^c	DMSO	2	1	0
OT2 -7	GACCCGCC CCGCCCCG CCTCTGG	145 ^c 3	GACCCCTCAGG AAGCTGGGAG	145 ^c	CTGCGAGATG CCCCAAATCG	145 ^c	betaína a 1M, TD	1	0	2
OT2 -8	GATCGACT CCACCCCG CCTCTGG	146 ^c 3	CCGGGGCCT CTGCTAGA	146 ^c	TGCTGGGATTA CAGGCGCGA	146 ^c	DMSO	1	1	1
OT2 -9	GCCCCCAC CCACCCCG CCTCTGG	146 ^c 3	CCAGGGGTG TCAGCGGAGG	146 ^c	TGCCTGGCCC TCTCTGAGTCT	146 ^c	DMSO	0	2	1
OT2 -10	GCCCCGCT CCTCCCCG CCTCCGG	146 ^c 3	CGACTCCACG GCGTCTCAGG	146 ^c	CAGCGCAGTC CAGCCCGATG	146 ^c	betaína a 1M, TD	2	1	0
OT2 -11	GGCCCCCT CCACCAAGG CCTCAGG	146 ^c 3	CTTCCCTCCC CCAGCACAC	147 ^c	GCTACAGTT GCACAGTGAG AGGT	147 ^c	DMSO	1	1	1
OT2 -12	GGCCCCCT CCTCCTCG CCTCTGG	147 ^c 3	CCCCGGGGA GTCTGTCTG A	147 ^c	CCCAGCGT CCAGGTCTTCC	147 ^c	72C Recoser, DMSO a 3%	1	0	2
OT2 -13	GGCGCCCT CCACCCCTG CCTCGGG	147 ^c 3	GAAGCGCGAA AACCCGGCTC	147 ^c	TCCAGGGTCC TTCTCGGCC	147 ^c	DMSO	1	0	2
OT2 -14	GTCCTCCA CCACCCCG CCTCTGG	147 ^c 3	AGGGTGGTCA GGGAGGCCCT	147 ^c	CATGGGGCTC GGACCTCGTC	148 ^c	DMSO	2	0	1
OT2 -15	TACCCCCC ACACCCCG CCTCTGG	148 ^c 3	GGGAAGAGG CAGGGCTGTC G	148 ^c	TGCCAGGAAG GAAGCTGGCC	148 ^c	72C Recoser, DMSO a 3%	0	2	1
OT2 -16	AACCCATT CACCCCTGC CTCAGG	148 ^c 4	GAGTGACGAT GAGCCCCGG G	148 ^c	CCCTTAGCTGC AGTCGCC	148 ^c	68C Recoser, DMSO a 3%	0	1	3
OT2 -17	ACACCCCC CCACCCCG CCTCAGG	148 ^c 4	CCCATGAGGG GTTTGAGTGC	148 ^c	TGAAGATGGG CAGTTGGGG	148 ^c	DMSO	0	2	2
OT2 -18	AGCCCCCA CCTCCCCG CCTCGGG	149 ^c 4	CACCTGGGGC ATCTGGGTGG	149 ^c	ACTGGGGTT GGGAGGGGAT	149 ^c	DMSO	2	0	2
OT2 -19	ATCCCCCC CCACCCCG CCTCAGG	149 ^c 4	TCATGATCCC CAAAAGGCT	149 ^c	CCATTGGTCT GATCTGTGGG T	149 ^c	DMSO	1	0	3
OT2 -20	CCCCACCC CCACCCCG CCTCAGG	149 ^c 4	TGGTGCCCCAG AATAGTGGCC A	149 ^c	AGGAATGTGT TGTGCCAGGG C	149 ^c	DMSO	1	2	1
OT2 -21	CCCCCCCC CCACCCCG CCCCGGG	149 ^c 4	GCCTCAGACA ACCCCTGCC	150 ^c	GCCAAGTGTAA CTCATCAAGAA AGTGG	150 ^c	Sem TD	2	1	1
OT2 -22	CCCCCCCC CCCCCCCG CCTCAGG	150 ^c 4	GCCGGGACAA GACTGAGTTG GG	150 ^c	TCCCGAACTCC CGCAAACAG	150 ^c	DMSO	1	2	1
OT2 -23	CGCCCTCC CCACCCCG CCTCCGG	150 ^c 4	TGCTGCAGGT GGTCCGGAG	150 ^c	CTGGAAACCGC ATCCTCCGCA	150 ^c	Sem TD	1	0	3
OT2 -24	CTCCCCAC CCACCCCG CCTCAGG	150 ^c 4	ACACTGGTCC AGGTCCCGTC T	150 ^c	GGCTGTGCCT TCCGATGGAA	151 ^c	DMSO	2	1	1
OT2 -25	CTCTCCCC CCACCCCG CCTCTGG	151 ^c 4	CTCTCCCCCACC CCCCCTCTGG (SEQ ID NO:2227)	151 ^c	AGGCTTCTGG AAAAGTCCTCA ATGCA	151 ^c	DMSO	3	0	2
OT2 -26	GCACCCCG CCTCAGG	151 ^c 4	Não otimizado					1	1	2
OT2 -27	GTCACTCC CCACCCCG CCTCTGG	151 ^c 4	CCCTCATGGT GGTCTTACGG CA	151 ^c	AGCCACACATC TTCTGGTAGG G	151 ^c	DMSO	1	1	2
OT2 -28	TGCCCCCT CCCCCCAG CCTCTGG	151 ^c 4	TGCGTCCGTC ATGCTGGAG	151 ^c	AGGGTGGGTT GTACTGGCTCA	152 ^c	DMSO	0	3	1

TABELA E

OT2 -29	TGCCCTC CCACCCCG CCTCTGG	152 ^c 4	GAGCTGAGAC GGCACCACTG	152 ^c	TGGCCTTGAAC TCTTGGGCT	152 ^c	betaína a 1M, TD	0	1	3
OT2 -30	TTCCCCCTC CACCCAGC CTCTGG	152 ^c 4	Não otimizado					1	2	1
OT2 -31	TTCTCCCTC CTCCCCGC CTCGGG	152 ^c 4	AGTGAGAGTG GCACGAACCA	152 ^c	CAGTAGGTGG TCCCTTCCGC	152 ^c	DMSO	2	1	1
OT2 -32	ACCCCTCGC CCACCCCG CCTCAGG	152 ^c 5	Não otimizado					1	1	3
OT2 -33	AGCCAACC CCACCCCG CCTCTGG	152 ^c 5	GGGAGAACCT TGTCCAGCCT	153 ^c	AAGCGGAAAA GCTGGGCAA	153 ^c	DMSO	0	2	3
OT2 -34	AGGCCCCC ACACCCCG CCTCAGG	153 ^c 5	CTTCCCAGTG TGGCCCGTCC	153 ^c	ACACAGTCAGA GCTCCGCG	153 ^c	DMSO	1	1	3
OT2 -35	AGGCCCCC CCGCCCCG CCTCAGG	153 ^c 5	Não otimizado					1	0	4
OT2 -36	ATCTGCCA CCACCCCG CCTCCGG	153 ^c 5	CTGAGAGGGG GAGGGGAG G	153 ^c	TCGACTGGTCT TGTCCCTCCCA	153 ^c	68C Recoser, DMSO a 3%	3	0	2
OT2 -37	CATCTTCCC CACCCCGC CTCTGG	153 ^c 5	CAGCCTGCTG CATGGAAAA	154 ^c	TGCAAGCCAAG AGAAAAAAGCCT	154 ^c	betaína a 1M, TD	1	0	4
OT2 -38	TTTCCCTC CACCCAGC CTCTGG	154 ^c 5	TCCCTCTGAC CCGGAAACCA	154 ^c	ACCCGACTTCC TCCCCATTGC	154 ^c	DMSO	2	1	2
OT2 -39	GTCGAGGT CCACCCCG CCTCAGG	154 ^c 5	TGGGGGTTGC GTGCTTGTCA	154 ^c	GCCAGGAGGA CACCAAGGACC	154 ^c	DMSO	4	1	0
OT2 -40	GTCGAGGT CCACCCCG CCTCAGG	154 ^c 5	ATCAGGTGCC AGGAGGACAC	154 ^c	GGCCTGAGAG TGGAGAGTGG	155 ^c	DMSO	4	1	0
OT2 -41	TCAGACCT CCACCCCG CCTCAGG	155 ^c 5	Não otimizado					1	4	0
OT2 -42	TGCAACCT CCTCCCG CCTCGGG	155 ^c 5	TGAGCCACAT GAATCAAGGC CTCC	155 ^c	ACCTCTCCAAG TCTCAGTAAC CTCT	155 ^c	DMSO	1	3	1
OT2 -43	ACCACTCT GCACCCCG CCTCTGG	155 ^c 6	GGTCCCTCTG TGCAGTGGAA	155 ^c	CTTGGTGGAC TGACACAGC	155 ^c	DMSO	2	2	2
OT2 -44	ACTACCCA CCTCCCG CCTCAGG	155 ^c 6	GCGAGGCTGC TGACTTCCCT	155 ^c	GCTGGGACTA CAGACATGTG CCA	156 ^c	DMSO	2	2	2
OT2 -45	ATTTCCCCC CCCCCGC CTCAGG	156 ^c 6	ATTGCAGGCG TGTCCAGGCA (SEQ ID NO:2228)	156 ^c	AAATCCTGCAT GGTATGGGA GT	156 ^c	DMSO	1	1	5
OT2 -46	CCACCATC CCACCCCG CCTCTGG	156 ^c 6	TGCTCTGCCA TTTATGCTCTA TGAACT	156 ^c	ACAGCCTCTTC TCCATGACTGA GC	156 ^c	DMSO	1	3	2
OT2 -47	CCCAAGCC CCACCCCG CCTCGGG	156 ^c 6	TCCGCCAA CAGGAGGCAG	156 ^c	GGGGTGGGAA AGCCATTGAG	156 ^c	DMSO	2	3	1
OT2 -48	CCGGCGTT CCGCCCCG CCTCTGG	157 ^c 6	GGGGGTCTGG CTCACCTGGA	157 ^c	CCTGTCGGGA GAGTGCCTGC	157 ^c	DMSO	3	1	2
OT2 -49	CCTGCCAT GCACCCCG CCTCAGG	157 ^c 6	TCCTGGTCA TTTGCTAGAA CTCTGGA	157 ^c	ACTCCAGATGC AACCAAGGGCT	157 ^c	DMSO	3	2	1
OT2 -50	CTGCCCTCT CACCCCGC CTCAGG	157 ^c 6	CGTGTGGTGA GCCTGAGTCT	157 ^c	GCTTCACCGTA GAGGCTGCT	157 ^c	DMSO	3	0	3
OT2 -51	TCTCTTTC CACCCCGC CTCAGG	157 ^c 6	AGGCCCTGAT AATTCTAGCTA CCAA	158 ^c	TCAGTGACAAC CTTTGTATTG GGCA	158 ^c	DMSO	0	2	4
OT2 -52	TTGACCCC CCGCCCCG CCTCAGG	158 ^c 6	Não otimizado					2	2	2
Alvo 3	GGTGAGTG AGTGTGTG CGTGTGG	158 ^c 0	TCCAGATGGC ACATTGTCAG	158 ^c	AGGGAGCAGG AAAGTGAGGT	158 ^c	DMSO			

TABELA E											
OT3	GGTGAGTG AGTGTGTG -1 GTGAGG	158 ^c	1	GCAGGCCAAC TGTCAAGGGT	158 ^c	CACCGACACA CCCACCTCAC	158 ^c	DMSO	0	0	1
OT3	GGTGAGTG AGTGTGTG -2 GTGGGG	158 ^c	2	GAGGGGAA GTCACCGACA A	159 ^c	TACCCGGGCC GTCTGTTAGA	159 ^c	DMSO	0	0	2
OT3	AGTGTGTG AGTGTGTG -3 CGTGTGG	159 ^c	2	GACACCCAC ACACTCTCAT GC	159 ^c	TGAATCCCTTC ACCCCCAAG	159 ^c	DMSO	1	0	1
OT3	GCTGAGTG AGTGTATG -4 CGTGTGG	159 ^c	2	TCCTTTGAGG TTCATCCCCC	159 ^c	CCAATCCAGG ATGATCCCGC	159 ^c	DMSO	1	0	1
OT3	GGTGAGTC AGTGTGTG -5 AGTGAGG	159 ^c	2	CAGGGCCAGG AACACAGGAA	159 ^c	GGGAGGTATG TGCAGGGAGTG	160 ^c	DMSO	1	1	0
OT3	GGTGAGTG AGAGTGTG -6 TGTGTGG	160 ^c	2	TGCAGCCTGA GTGAGCAAGT GT	160 ^c	GCCCCAGGTGC TAAGCCCCCTC	160 ^c	DMSO	1	0	1
OT3	GGTGAGTG AGTGAGTG -7 AGTGAGG	160 ^c	2	TACAGCCTGG GTGATGGAGC	160 ^c	TGTGTATGGA CTTCCCCATTG T	160 ^c	betaína a 1M, TD	1	1	0
OT3	GGTGAGTG AGTGAGTG -8 AGTGAGG	160 ^c	2	GGCAGGCATT AAACTCATCA GGTCC	160 ^c	TCTCCCCAAG GTATCAGAGA GCT	160 ^c	DMSO	1	1	0
OT3	GGTGAGTG AGTGCCTG -9 CGGGTGG	161 ^c	2	GGGCCTCCCT GCTGGTTCTC	161 ^c	GCTGCCGTCC GAACCCAAGA	161 ^c	DMSO	0	1	1
OT3	GGTGAGTG TGTGTGTG -10 AGTGTGG	161 ^c	2	ACAAACCGAG GTGGACCAGA	161 ^c	ACTCCGAAAAT GCCCCGCACT	161 ^c	DMSO	1	1	0
OT3	GGTGAGTG TGTGTGTG -11 CATGTGG	161 ^c	2	AGGGGAGGG GACATTGCCCT	161 ^c	TTGAGAGGGTT CAGTGGTTGC	161 ^c	DMSO	1	0	1
OT3	GGTGAGTG AGTGTGTG -12 GTGTGG	161 ^c	2	CTAATGCCAAC GGCTGGGGG	162 ^c	AGCCAACGGC AGATGCAAT	162 ^c	DMSO	1	0	1
OT3	GGTGAGTG TGTGTGTG -13 CGTGGCG	162 ^c	2	GAGCGAAGTT AACCCACCGC	162 ^c	CACACATGCAC ATGCCCTG	162 ^c	68C, DMSO a 3%	2	0	0
OT3	GGTGAGTG TGTGTGTG -14 CGTGTGG	162 ^c	2	GCATGTGTCT AACTGGAGAC AATAGCA	162 ^c	TCCCCCATATC AACACACACA	162 ^c	DMSO	2	0	0
OT3	GGTGAGTG TGTGTGTG -15 CGTGTGG	162 ^c	2	GCCCCCTCCCG CCTTTGTGT	162 ^c	TGGGCAAAGG ACATGAAACAG ACA	163 ^c	DMSO	2	0	0
OT3	GGTGAGTG TGTGTGTG -16 CGTGTGG	163 ^c	2	GCCTCAGCTC TGCTCTTAAG CCC	163 ^c	ACGAACAGATC ATTTTTCATGG CTTCC	163 ^c	DMSO	2	0	0
OT3	GGTGAGTG AATGTGTG -17 CGTGAGG	163 ^c	2	CTCCAGAGCC TGGCTTACCA	163 ^c	CCCTCTCCGG AAGTGCCTTG	163 ^c	DMSO	0	1	1
OT3	TGTGGGTG AGTGTGTG -18 CGTGAGG	163 ^c	2	TCTGTCACCA CACAGTTACC ACC	163 ^c	GTTGCCTGGG GATGGGGTAT	163 ^c	DMSO	0	1	1
OT3	ACTGTGTG AGTGTGTG -19 CGTGAGG	164 ^c	3	GGGGACCCCTC AAGAGGCACT	164 ^c	GGGCATCAAA GGATGGGGAT	164 ^c	DMSO	2	0	1
OT3	AGAGAGTG AGTGTGTG -20 CATGAGG	164 ^c	3	TGTGGAGGGT GGGACCTGGT	164 ^c	ACAGTGAGGT GCGGTCTTG GG	164 ^c	DMSO	1	0	2
OT3	AGCGAGTG GGTGTGTG -21 CGTGGGG	164 ^c	3	CGGGGTGGCA GTGACGTCAA	164 ^c	GGTGCAGTCC AAGAGCCCCC	164 ^c	DMSO	0	0	3
OT3	AGGGAGTG ACTGTGTG -22 CGTGTGG	164 ^c	3	AGCTGAGGCA GAGTCCCCGA	165 ^c	GGGAGACAGA GCAGCGCCTC	165 ^c	DMSO	1	1	1
OT3	AGTGAGTG AGTGAGTG -23 AGTGAGG	165 ^c	3	ACCACCAAGAC CCCACCTCCA	165 ^c	AGGACGACTT GTGCCCCATTCA A	165 ^c	72C Recoser, DMSO a 3%	1	1	1
OT3	CATGAGTG AGTGTGTG -24 GGTGGGG	165 ^c	3	GGGTCAAGGAC GCAGGTCAGA	165 ^c	TCCACCCACC CACCCATCCT	165 ^c	72C Recoser, DMSO a 3%	2	0	1
OT3	CGTGAAGTG TGTGTATGC -25 GTGTGG	165 ^c	3	ACACTCTGGG CTAGGTGCTG GA	165 ^c	GCCCCCTCAC CACATGATGCT	166 ^c	DMSO	2	0	1

TABELA E										
OT3	GGACTGTG -26	AGTGTGTG CGTGAGG	166 ^c 3	GGGGCCATT CTCTGCTGCA	166 ^c TGGGGATCCTT GCTCATGGC	166 ^c DMSO	3	0	0	
OT3	GGTGTGTG -27	CCTGTGTG CGTGTGG	166 ^c 3	ACACACTGGC TCGCATTAC CA	166 ^c CCTGCACAG GCCAGGTGTT	166 ^c DMSO	2	1	0	
OT3	GTTTCATGA -28	GTGTGTGC GTGGGG	166 ^c 3	TGGGCACGTA GTAAACTGCA CCA	166 ^c CTCGCCCG TGACTGTAGG	166 ^c DMSO	0	3	1	
OT3	TGAGTGTG -29	AGTGTGTG CGTGGGG	167 ^c 3	TCAGCTGGC CTGGGCTTGG	167 ^c AGAGCACTGG TGAGCAGTCA GT	167 ^c DMSO	2	1	0	
OT3	TGCCAGTG -30	AGTGTGTG CGTGTGG	167 ^c 3	AGACACAGCC AGGGCCTCAG	167 ^c GGTGGGCGTG TGTGTGTACC	167 ^c 68C, DMSO a 3%	1	1	1	
OT3	TGGGTGTG -31	AGTGTGTG CGTGTGG	167 ^c 3	ACACTCTCAC ACACGCACCA A	167 ^c GAGAACGTAG GGCTGGCGGG	167 ^c 72C Recazer, DMSO a 3%	1	2	0	
OT3	TGTATGTGA -32	GTGTGTGC GTGTGG	167 ^c 3	ACTGCCTGCA TTTCCCCGGT	168 ^c TGGTGAGGGC TTCAGGGAGC	168 ^c DMSO	1	1	1	
OT3	TGTGAGAG -33	AGAGTGTG CGTGTGG	168 ^c 3	GCCAGGTTCA TTGACTGCC	168 ^c TCCTTCTACAC ATCGCGCC	168 ^c DMSO	2	1	0	
OT3	TGTGCCCTG -34	AGTGTGTG CGTGTGG	168 ^c 3	CGAGGGAGCC GAGTTCGTAA	168 ^c CTGACCTGGG GCTCTGGTAC	168 ^c DMSO	1	2	0	
OT3	TGTGTGTG -35	GTGTGTGC GTGTGG	168 ^c 3	TCCTCGGGAA GTCATGGCTT CA	168 ^c GCACTGAGCA ACCAGGAGCA C	169 ^c DMSO	2	1	0	
OT3	AGCGTGTG -36	AGTGTATG CGTGGGG	169 ^c 4	Não otimizado				1	0	3
OT3	ATTGAGTGT -37	GTGAGTGC GTGGGG	169 ^c 4	TAAACCGTT CCCCCGCCTC	169 ^c GCTCCCCTGC CAGGTGAACC	169 ^c DMSO	2	1	1	
OT3	CATGTGTG -38	GGTGTGTG CGTGTGG	169 ^c 4	CCTGCTGAGA CTCCAGGTCC	169 ^c CTGCGGAGTG GCTGGCTATA	169 ^c DMSO	2	0	2	
OT3	CCCGAGTG -39	TGTGTGTG CGTGTGG	169 ^c 4	CTCGGGACT GACAAGCCG	169 ^c GGAGCAGCTC TTCCAGGGCC	170 ^c DMSO	3	0	1	
OT3	CTGGAGTG -40	AGTGTGTGT GTGTGG	170 ^c 4	CCCCGACCAA AGCAGGAGCA	170 ^c CTGGCAGCCT CTGGATGGGG	170 ^c DMSO	1	2	1	
OT3	GTTTCATGA -41	GTGTGTGC GTGGGG	170 ^c 4	Não otimizado				0	3	1
OT3	TATGTGTG -42	GTGTGTGC GTGTGG	170 ^c 4	ATTCAGAGC CCCGGGAAA	170 ^c AGGCCCGGGT GTTATGGTTA	170 ^c DMSO	1	2	1	
OT3	TATGTGTG -43	GTGTGTGC GTGGGG	170 ^c 4	GCCAGTGGCT TAGTGTCTTG TGT	170 ^c TGACATATT CCTGGGCCAT GGGT	171 ^c DMSO	2	1	1	
OT3	TCTGTGTG -44	GTGTGTGC GTGGGG	171 ^c 4	TGCCAGAGA ACATGGGCCA GA	171 ^c CCATGCTGACA TCATATACTGG GAAGC	171 ^c DMSO	3	1	0	
OT3	TCTGTGTG -45	GTGTGTGC GTGTGG	171 ^c 4	GCGTGTCTCT GTGTGGTGC	171 ^c CCAGGCTGGG CACACAGTT	171 ^c DMSO	3	1	0	
OT3	TGAGCGTG -46	AGTGTGAG CGTGTGG	171 ^c 4	Não otimizado				2	2	0
OT3	TGTCTTGA -47	GTGTGTGC GTGTGG	171 ^c 4	TGCCAGTCC AATTTTCAGC AGCT	171 ^c AGGATGAGTC ATGTCCTTGT GGGG	172 ^c DMSO	2	2	0	
OT3	TTTGTGTG -48	GTGTGTGC GTGTGG	172 ^c 4	GGGTGAAAAT TTGGTACTGTT AGCTGT	172 ^c AATGACTCATT CCCTGGGTAT CTCCCA	172 ^c DMSO	2	2	0	
OT3	AAGGCCGTG -49	TGTGTGTG CGTGTGG	172 ^c 5	TGCCCATCA ATCACCTCGG C	172 ^c CAAGGTGGC AGGGCAGTGA	172 ^c DMSO	1	2	2	
OT3	AATTGCGTGT -50	GTGTGTGC GTGGGG	172 ^c 5	GCCTCCTCTG CCGCTGGTAA	172 ^c TGAGAGTCCCT GTTGCTCCACA CT	172 ^c DMSO	1	2	2	

TABELA E									
OT3	ATGGTGTG								
-51	TGTGTGTG	173 ^c 5		Não otimizado				2	2
	CGTGTGG								1
OT3	CACGTGTG								
-52	TGTGTGTG	173 ^c 5	GCCACCAAAA	173 ^c	ACATGCATCTG	173 ^c	DMSO	3	0
	CGTGTGG		TAGCCAGCGT		TGTGTGCGT				2
OT3	GAAATTGA								
-53	GTGTGTGC	173 ^c 5	ACAGACTGAC		TGTATCTTCT				
	GTGTGG		CCTTGAAATA	173 ^c	TGCCAATGGTT	173 ^c	DMSO	2	1
			ACCACT		TTCCC				2
OT3	TAAGTGTGT								
-54	GTGTGTGC	173 ^c 5	AGCCAAATT		TCCTGGAGAG				
	GTGTGG		CTAACACAGCA	173 ^c	CAGGCATTTT	173 ^c	DMSO	3	1
			GCAC		GT				1
OT3	TATATGTGT								
-55	GTGTGTGC	174 ^c 5	ACCTCCTTGT	174 ^c	GGGGGAAAGG	174 ^c	DMSO	2	1
	GTGGGG		GCTGCCTGGC		TAACCCCTGGG				2
OT3	TATCTGTGT								
-56	GTGTGTGC	174 ^c 5	CACAAAGCTC		TGATCCGATG				
	GTGTGG		TACCTTTCCA	174 ^c	TTGTTTCACAG	174 ^c	DMSO	3	1
			GTAGTGT		CT				1
OT3	TTTATGTGT								
-57	GTGTGTGC	174 ^c 5	TGTGGGATT		ACGCACAAAAA				
	GTGTGG		ACCTGCCTGG	174 ^c	TGCCCTTGTCA	174 ^c	DMSO	2	2
			C						1
OT3	TTTTTGTGT								
-58	GTGTGTGC	174 ^c 5	TGAGGCAGAC		GGCCGAGCAC				
	GTGGGG		CAGTCATCCA	175 ^c	AGTGTAGGGC	175 ^c	DMSO	2	3
			GC						0
OT3	AAAAATTGT								
-59	GTGTGTGC	175 ^c 6	ATTAGCTGGG	175 ^c	ACTGCATCTCA				
	GTGGGG		CGTGGGGAG		TCTCAGGCAG				
			CT						3
OT3	ACAATGTG								
-60	GTGTGTG	175 ^c 6	TGAAGCAGAA		TCAGCTTCACA				
	CGTGTGG		GGAGTGGAGA	175 ^c	TCTGTTTCAGT	175 ^c	DMSO	4	0
			AGGA		TCAGT				2
OT3	ATGTTGTG								
-61	TGTGTGTG	175 ^c 6	TGGTGGAGTG	175 ^c	AGAGCAGAAA				
	CGTGTGG		TGTGTGTGGT		GAGAGTGC				
			A		CC				2
OT3	CAAAATTGT								
-62	GTGTGTGC	176 ^c 6	GCCCCCTGTAC		TGCACAAGCC				
	GTGTGG		GTCCTGACAG	176 ^c	ACTTAGCCTCT	176 ^c	DMSO	3	1
			C		CT				2
OT3	CCCTGGTG								
-63	TGTGTGTG	176 ^c 6	AGCGCAGGTA	176 ^c	TCTCTGCC				
	CGTGTGG		AACAGGCCA		GTTCCTTGT				
									2
OT3	TCCCGTTGT								
-64	GTGTGTGC	176 ^c 6	ATGGGTC	176 ^c	ACAGCAGGAA				
	GTGGGG		GGTACCGACG		GGAGCCGCAG				
									1
OT3	TCTCTGTGT								
-65	GTGTGTGC	177 ^c 6	CGGGCGGGT	177 ^c	AGGAGGTCTC				
	GTGTGG		GGACAGATGA		GAGCCAGGGG				
			G						1
OT3	TTAAGGTG								
-66	GGTGTGTG	177 ^c 6	TCAACCTAGT		GTCTTATATACA				
	CGTGGGG		GAACACAGAC	177 ^c	GCCCCAACCC	177 ^c	DMSO	1	2
			CACTGA		TCATGT				3
OT3	TTATATTGT								
-67	GTGTGTGC	177 ^c 6	GCCAGGGCCA	177 ^c	TGTCTTTCTT				
	GTGGGG		GTGGATTGCT		AGTATGTCAGC				
					CGGA				0
OT3	TTGAGGAG								
-68	AGTGTGTG	177 ^c 6	GAGCCCCACC	178 ^c	GCCAGAGCTA				
	CGTGAGG		GGTCAGTCC		CCCACCGGCC				
									2
		178 ^c		178 ^c		178 ^c			
Alvo	GAGTCCGA								
4	GCAGAAGA	178 ^c 0	GGAGCAGCTG	178 ^c	GGGAAGGGGG				
	AGAAAGG		GTCAGAGGGG		ACACTGGGGA				
OT4	GAGTTAGA								
-1	GCAGAAGA	178 ^c 2	TCTCTCTTCA		ATCTGCACATG				
	AGAAAGG		ACTCATGACC	178 ^c	TATGTACAGGA	179 ^c	DMSO	0	1
			AGCT		GTCTAT				1
OT4	AAGTCAGA								
-2	GGAGAAGA	179 ^c 3	AAGACAGAGGAG	179 ^c	AGGGTGTACT				
	AGAAGGG	(SEQ ID NO:2229)	TGGGAAATCT		TGTTGAACTTT	179 ^c	DMSO	2	1
			CCC		GCA				1
OT4	AAGTCAGA								
-3	GGAGAGGA	179 ^c 3	GATGGCCCA	179 ^c	ACTTCGTAGAG				
	AGAAAGG		CTGAGCACGT		CCTTAAACATG	179 ^c	DMSO	1	0
					TGGC				2
OT4	AAGTCTGA								
-4	GCACAAGA	179 ^c 3	AGGATTAATG		TCAAACAAAGGT				
	AGAATGG		TTTAAAGTCAC	179 ^c	GCAGATACAG				
			TGGTGG		CA				
OT4	ACGTCTGA								
-5	GCAGAAGA	180 ^c 3	TCCAAGCCAC		TGCTCTGTGGA				
	AGAATGG		TGGTTCTCA	180 ^c	TCATATTG	180 ^c	DMSO	0	1
			GTCA		GGGA				2
OT4	GACTCCTA								
-6	GCAAAGA	180 ^c 3	ACTTTCAAGAG		CCCACGCTGA				
	AGAATGG		CTTGGGCAG	180 ^c	AGTGAATGG	180 ^c	DMSO	1	1
			GT		C				1

TABELA E											
OT4 -7	GAGACTGA GAAGAAGA AGAAAGG	180 ^c 3	CAAAGCATGC CTTTCAGCCG	180 ^c TTGGCACCT	GGCTCTTCGAT betaína a 1M, TD	180 ^c		1	1	1	
OT4 -8	GAGCCCGA GCAGAAGA AGGAGGG	180 ^c 3		Não otimizado				1	0	2	
OT4 -9	GAGCCTGA GCAGAAGG AGAAAGG	181 ^c 3	GGACTCCCTG CAGCTCCAGC	181 ^c AGGAACACAG GCCAGGCTGG	72C Recoder, 6% DMSO	181 ^c		0	0	3	
OT4 -10	GAGGCCGA GCAGAAGA AAGACGG	181 ^c 3	CCCTTTAGGC ACCTTCCCCA	181 ^c CCGACCTTCAT CCCTCCTGG	DMSO	181 ^c	DMSO	0	1	2	
OT4 -11	GAGTAAGA GAAGAAGA AGAAAGG	181 ^c 3	TGATTCTGCC TTAGAGTCCC AGGT	181 ^c TGGGCTCTGT GTC CCTACCCA	DMSO	181 ^c	DMSO	0	3	0	
OT4 -12	GAGTAGGA GGAGAAGA AGAAAGG	181 ^c 3		Não otimizado				2	1	0	
OT4 -13	GAGTCCGG GAAGGAGA AGAAAGG	182 ^c 3	AGGCAGGAGA GCAAGCAGGT	182 ^c TGACTGACCG CT	DMSO	182 ^c	DMSO	0	1	2	
OT4 -14	GATTCTAC CAGAAGAA GAATGG	182 ^c 3	CTCCCCATTG CGACCCGAGG	182 ^c AGAGGCATTG ACTTGGAGCA CCT	DMSO	182 ^c	DMSO	1	2	0	
OT4 -15	GCGACAGA GCAGAAGA AGAAAGG	182 ^c 3	CTGGAGGCCA GCAGGAAAGGC	182 ^c CCTCAGGGAG GGGGCCTGTAT	DMSO	182 ^c	DMSO	1	2	0	
OT4 -16	AAATCCAA CCAGAAGA AGAAAGG	182 ^c 4	ACTGTGGCG TTGTCCCCAC	183 ^c AGGTCTGGTC AGGGTTAACAG GA	DMSO	183 ^c	DMSO	1	0	3	
OT4 -17	AAGTCTGA GGACAAGA AGAATGG	183 ^c 4	GGCGCTCCCT TTTCCCTTTG T	183 ^c CGTCACCCATC GTCTCGTGGAA	DMSO	183 ^c	DMSO	2	0	2	
OT4 -18	AAGTTGGA GCAGGAGA AGAAAGG	183 ^c 4	TGCCATCTATA GCAGCCCCCT	183 ^c GCATCTTGCTA ACCGTACTTCT TCTGA	DMSO	183 ^c	DMSO	1	0	3	
OT4 -19	AATACAGA GCAGAAGA AGAATGG	183 ^c 4	GTGGAGACGC TAACACTGTG AGGT	183 ^c GCTCCTACAG C	DMSO	184 ^c	DMSO	1	2	1	
OT4 -20	AGGTACTA GCAGAAGA AGAAAGG	184 ^c 4	CCGAACTTCT GCTGAGCTTG ATGC	184 ^c CCAAGTCATG GGCAACAAGG GA	DMSO	184 ^c	DMSO	0	2	2	
OT4 -21	AGGTGCTA GCAGAAGA AGAAGGG	184 ^c 4		Não otimizado				1	1	2	
OT4 -22	AGGTGGGA GCAGAAGA AGAAGGG	184 ^c 4	TGCCCCCAAG ACCTTTCTCC	184 ^c ATGGCAGGCA GAGGAGGAAG	DMSO	184 ^c	DMSO	2	0	2	
OT4 -23	CAAACGGA GCAGAAGA AGAAAGG	184 ^c 4	GGGTGGGGC CATTGTGGGT T	184 ^c CTGGGGCCAG GGTTTCTGCC	DMSO	185 ^c	DMSO	3	0	1	
OT4 -24	CACTCTGA GGAGAAGA AGAAAGG	185 ^c 4	TGGAGAACAT GAGGGCTTG CAA	185 ^c TCCTTCTGTAG GCAATGGGAA CAA	DMSO	185 ^c	DMSO	3	0	1	
OT4 -25	CAGTCATG GCAGAAGA AGAAAGG	185 ^c 4	GCCACATGGT AGAAGTCGGC	185 ^c GGCAGATTTC CCCATGCTG	betaína a 1M, TD	185 ^c		1	2	1	
OT4 -26	CCGTCCCA GCAGTAGA AGAATGG	185 ^c 4	TGTACACCCC AAGTCCTCCC	185 ^c AAGGGGAGTG TGCAAGCCTC	DMSO	185 ^c	DMSO	3	1	0	
OT4 -27	GTCTCGA TCAGAAGAA GAAAGG	186 ^c 4	AGGTCTGGCT AGAGATCAG CA	186 ^c AGTCCAACACT CAGGTGAGAC CCT	DMSO	186 ^c	DMSO	3	1	0	
OT4 -28	TAATCCAAT CAGAAGAA GAAGGG	186 ^c 4	CCAAGAGGAC CCAGCTGTG GA	186 ^c GGGTATGGAA TTCTGGATTAG CAGAGC	DMSO	186 ^c	DMSO	0	2	2	
OT4 -29	TATACGGA GCAGAAGA AGAATGG	186 ^c 4	ACCATCTCTTC ATTGATGAGT CCCAA	186 ^c ACACTGTGAGT ATGCTTGGCGT	DMSO	186 ^c	DMSO	2	2	0	
OT4 -30	ACTTCCCTG CAGAAGAA GAAAGG	186 ^c 5	GGCTGCGGG GAGATGAGCT C	187 ^c TCGGATGCTT TCCACAGGGC T	DMSO	187 ^c	DMSO	2	2	1	
OT4 -31	AGGACTGG GCAGAAGA AGAAGGG	187 ^c 5	TCTTCCAGGA GGGCAGCTCC	187 ^c CCAATCCTGAG CTCCTACAAAGG CT	DMSO	187 ^c	DMSO	1	0	4	

TABELA E										
OT4 -32	AGGTTGGA GAAGAAGA AGAACGG	187 ^c 5	GAGCTGCACT GGATGGCACT	TGCTGGTTAAG GGGTTTTG GA	187 ^c	DMSO	1	1	3	
OT4 -33	AGTCAGA GCAGGAGA AGAATGG	187 ^c 5	TCTGGGAAGG TGAGGAGGCC A	TGGGGGACAA TGAAAAGCAA TGA	187 ^c	DMSO	0	2	3	
OT4 -34	ATGACACA GCAGAAGA AGAACGG	188 ^c 5	CTTGCTCCCA GCCTGACCCC	AGCCCTTGCC ATGCAGGACC	188 ^c	DMSO	3	1	1	
OT4 -35	ATGACAGA GAAGAAGA AGAAAGG	188 ^c 5	GGGATTTTAT CTGTGGGTG CGAA	AACCACAGATG TACCTCTCAAAG CT	188 ^c	DMSO	2	2	1	
OT4 -36	CCGCCCCCT GCAGAAGA AGAACCG	188 ^c 5	ACCCATCAGG ACCGCAGCAC	TCTGGAACCTG GGAGGCGGA	188 ^c	72C Recorer, DMSO a 3%	3	1	1	
OT4 -37	GCAGGAGA GCAGAAGA AGAAAGG	189 ^c 5	CGTCCCTCAC AGCCAGCCTC	CCTCCTGGG CCTGGGGTTC	189 ^c	DMSO	1	3	1	
OT4 -38	GTTCAGA GCAGAAGA AGAATGG	189 ^c 5	CCCTCTGCAA GGTGGAGTCT CC	AGATGTTCTGT CCCCAGGCCT	189 ^c	DMSO	1	3	1	
OT4 -39	GTITAGAAG CAGAAGAA GAAAGG	189 ^c 5	GGCTTCCACT GCTGAAGGCC T	TGCCGCTCCA CATACCTCTCC	189 ^c	DMSO	2	1	2	
OT4 -40	TATGGCAA GCAGAAGA AGAAAGG	189 ^c 5	AGCATTGCCT GTCGGGTGAT GT	AGCACCTATTG GACACTGGTC TCT	190 ^c	DMSO	1	3	1	
OT4 -41	TGGTGGGA TCAGAAGAA GAAAGG	190 ^c 5	TCTAGAGCAG GGGCACAAATG C	TGGAGATGGA GCCTGGTGGG A	190 ^c	DMSO	2	2	1	
OT4 -42	ACCCACGG GCAGAAGA AGAACGG	190 ^c 6	GGTCTCAGAA AATGGAGAGA AAGCACG	CCACAGAAA CCTGGCCCT	190 ^c	DMSO	1	2	3	
OT4 -43	ACTCTGTAT CAGAAGAA GAAGGG	190 ^c 6	GGTTGCTGAT ACCAAAACGT TTGCT	TGGGTCCTCTC CACCTCTGCA	191 ^c	DMSO	0	3	3	
OT4 -44	ACTGATGA GCAGAAGA AGAAAGG	191 ^c 6	ACTCTCCTAA GTACTGATAT GGCTGT	CAGAACATTC TCTGTTGCCA	191 ^c	DMSO	0	4	2	
OT4 -45	ATTTTAGTG CAGAAGAA GAAAGG	191 ^c 6	Não otimizado				2	2	2	
OT4 -46	ATTTTAGTG CAGAAGAA GAAAGG	191 ^c 6	Não otimizado				2	2	2	
OT4 -47	CCATGCCA GCAGAAGA AGAACGG	191 ^c 6	CAATGCCCTGC AGTCCTCAGG A	TCCCAAGAGAA AACTCTGTCT GACA	191 ^c	DMSO	4	1	1	
OT4 -48	CCATTACA GCAGAAGA AGAACGG	191 ^c 6	GCATTGGCTG CCCAGGGAAA	TGGCTGTGCT GGGCTGTGTT	192 ^c	DMSO	2	2	2	
OT4 -49	CGAGGCGG GCAGAAGA AGAACGG	192 ^c 6	CCACAAGCCT CAGCCTACCC G	ACAGGTGCCA AACACTGCCT	192 ^c	DMSO	2	1	3	
OT4 -50	TCATTGCAGCAGAAGAA TCATTGTAGCAGAAGAACGG (SEQ ID NO:2230)			TCATTGCAGCAGAAGAA TCATTGTAGCAGAAGAACGG (SEQ ID NO:2230)	GCCTCTTGC AATGAGACT CTTT	CGATCAGTCC CCTGGCGTCC	192 ^c	DMSO	2 / 1	2 / 3
OT4 -51	TCTCCAGG GCAGAAGA AGAACGG			TCTCCAGG GCAGAAGA AGAACGG	TCCCAGAAC TGCCCTCGCA	AGGGGTTTCC AGGCACATGG G	192 ^c	DMSO	0	4
193 ^c										
Alvo 5	GTCATCTTA GTCATTACC TGAGG	193 ^c 0	TCCTAAAAATC AGTTTGAGAT TTACTTCC	AAAGTGTAGC CAACATACAGA AGTCAGGA	193 ^c	DMSO				
OT5 -1	GGTATCTAA TTACCTGTGG (SEQ ID NO:2231) GGTATCTAAGTC ATACCTGTGG (SEQ ID NO:2232)			GGTATCTAA TTACCTGTGG (SEQ ID NO:2231) GGTATCTAAGTC ATACCTGTGG (SEQ ID NO:2232)	ACATCTGGGG AAAGCAAAAG TCAACA	TGTCTGAGTAT CTAGGCTAAAG GTGTT	193 ^c	DMSO	1 / 2	1
OT5 -2	GTAATATTA GTCATTACC GGTGG	194 ^c 3	ACGATCTTC TTCATTTCCCT GTACA	AGTGCTTTGTG AACTGAAAAGC AAACAA	194 ^c	DMSO	0	3	0	

TABELA E												
OT5 -3	GTAATCTGA GTCATTTCC TGAAA	194 ^c 3	GCACCTGGT GCTGCTAAAT GCC	GGGCAACTGA ACAGGCATGA ATGG	194 ^c	DMSO	1	2	0			
OT5 -4	GTCATCCTA GTCATTAC TGAAA	194 ^c 3	AACTGTCCTG CATCCCCGCC	GGTCACCTG GATCCACCCA	194 ^c	DMSO	1	1	1			
OT5 -5	GTCATCCTA GTGCTTAC TGAGG	194 ^c 3	Não otimizado	195 ^c	195 ^c		1	1	1			
OT5 -6	GTCATCTGA GGCATTAA CTGGGG	195 ^c 3	CATCACCCCT CACCAGGCC	ACCACTGCTG CAGGCTCCAG	195 ^c	72C Recozer, DMSO a 3%	0	3	0			
OT5 -7	AATATGTTA GTCATTAC TGAGG	195 ^c 4	Não otimizado				2	0	2			
OT5 -8	ATAAACGTA GTCATTAC TGAGG	195 ^c 4	CCTGACCGT GGTCCCAGC	TGGTGCCTGG TGTGTGTTGG	195 ^c	72C Recozer, DMSO a 3%	1	2	1			
OT5 -9	ATCATCATC GTCATTATC TGAGG	195 ^c 4	TGGGAACATT GGAGAAAGTT CCTGA	CCATGTGACTA CTGGGCTGCC C	196 ^c	DMSO	1	1	2			
OT5 -10	ATCATTTTA CTCATTACT TGTGG	196 ^c 4	AGCCTTGGCA AGCAACTCCC T	GGTCTCTCTC TCAGAAAAGAA AGAGG	196 ^c	DMSO	1	0	3			
OT5 -11	ATCATTTTA GTCATCTCC TGTGG	196 ^c 4	GGCAGCGGAC TTCAGAGCCA	GCCAGAGGCT CTCAGCAGTC C	196 ^c	DMSO	1	0	3			
OT5 -12	CACAGCTTA GTCATCACC TGAGG	196 ^c 4	CCAGCCTGGT CAATATGGCA	ACTGTGCCCA GCCCATATT	197 ^c	DMSO	2	1	1			
OT5 -13	CCCAGCTTA GTCATTAGC TGTGG	197 ^c 4	ATGCCAACAC TCGAGGGGCC	CGGGTTGTGG CACCGGGTTA	197 ^c	DMSO	2	1	1			
OT5 -14	CTCACCTTT GTCATTTCC TGAGG	197 ^c 4	TTGCTCTAGT GGGGAGGGG G	AGAGTTCAGG CATGAAAAGAA GCAACA	197 ^c	DMSO	3	0	1			
OT5 -15	CTCATTTTA TTCATTGCC TGAGG	197 ^c 4	AGCTGAAGAT AGCAGTGT AAGCCT	TGCAATTGAG GGGCTCTCTC A	197 ^c	DMSO	1	1	2			
OT5 -16	CTCTCCCTA GTCACTACC TGAGG	198 ^c 4	AGTCACTGGA GTAAGCTGC CT	TGCCAGGCCA AAGTTTAGT GTGT	198 ^c	DMSO	2	0	2			
OT5 -17	CTTATCTCT GTCATTAC TGAGG	198 ^c 4	GGGTCTCCCT CAGTGCCTG	TGTGTGGTAG GGAGCAAAC GACA	198 ^c	DMSO	2	0	2			
OT5 -18	GACAGCTC CGTCATTAC CTGGGG	198 ^c 4	TGGGGCTGT TAAGAGGCAC A	TGACCAACAC ACCCCCACG	198 ^c	DMSO	1	2	1			
OT5 -19	GCCACCTC AGTCATTAG CTGGGG	198 ^c 4	TCAAACAGA TTGACCAAGG CCAAT	TGTGTGTTAA GCTGCACCCC AGG	199 ^c	DMSO	1	0	3			
OT5 -20	GGAATCTTA CTCATTACT TGAGG	199 ^c 4	TCTGGCACCA GGACTGATTG TACA	GCACGCAGCT GACTCCCAGA	199 ^c	DMSO	1	2	1			
OT5 -21	GTGGCCTC AGTCATTAC CTGCGG	199 ^c 4	Não otimizado				1	0	3			
OT5 -22	GTGTTTTA GTGATTACC TGAGG	199 ^c 4	AGCATCTGT ATACCCCTACC TGTCT	ACCAAGGCTG CCACAGAGTC	199 ^c	DMSO	1	0	3			
OT5 -23	TACATCTTA GTCCTCAC CTGTGG	199 ^c 4	TAGTCTTGTG CCCAGGCTG	CTCGGCCCCCT GAGAGTTCAT	200 ^c	DMSO	1	2	1			
OT5 -24	TCCATCTCA CTCATTAC TGAGG	200 ^c 4	TCCATCTCACTCA TTACCTGAGG (SEQ ID NO:2233) TCCATCTCACTCA TTACCTGATG (SEQ ID NO:2234)	TTACCTGAGG (SEQ ID NO:2233) CTGCAACCA GGCCCTTACC	200 ^c	GAGCAGCAGC AAAGCCACCG	200 ^c	DMSO	1	1	2	
OT5 -25	TTCATCTCA GTCAACAC CTGGGG	200 ^c 4	GCCTGGAGAG CAAGCCTGGG	AGCCGAGACA ATCTGCCCG	200 ^c	DMSO	1	1	2			
OT5 -26	TITATATTA GTGATTACC TGTGG	200 ^c 4	TITATATTAGTGTAT TACCTGCGG (SEQ ID NO:2235)	AGTGAACCAA ACAAGCAGCA GTCTGA	200 ^c	GGCAGGTCTG ACCAGTGGG	201 ^c	Sem TD	DMSO	1	2	1

TABELA E

OT5 -27	AACGTGTA AGTCATTAC CTGAGG	201' 5	AGGCTCAGAG AGGAAGCAA TGGA	TGAGTAGACA GAAATGTTACC GGTGT	201' DMSO	3	0	2
OT5 -28	AAGATCAC AGTCATTAC CTGGGG	201' 5	TCAGAGATGT TAAAGCCTG GTGGG	AGTGAACCAA GGGAATGGGG GA	201' DMSO	3	0	2
OT5 -29	AGAATATTA GTCCTTACCA TGAGG	201' 5	TGTGCTTCT GGGGTAGTGG CA	CACCTCAGCC CTGTAGTCCTG G	201' DMSO	0	4	1
OT5 -30	AGCAGATT AGTGAATTAC CTGGGG	202' 5	CCATTGGGTG ACTGAATGCA CA	GCCACTGTCC CCAGCCTATT	202' betaina a 1M, TD	1	3	1
OT5 -31	AGTAGCTTA GTGATTACCA TGAGG	202' 5	ACCAAGAAC TGAAAAGGAA ACCC	TGAGATGGCAT ACGATTACCC A	202' DMSO	1	2	2
OT5 -32	CACGGCTT ACTCATTAC CTGGGG	202' 5	AGGGTGGGA CTGAAAGGAG CT	TGGCATCACTC AGAGATTGGAA CACAA	202' DMSO	3	1	1
OT5 -33	CATATGTTA GGCATTAC CTGGGG	202' 5	ACCAAGTCTG TGTGACCTTG GA	TCCTATGGGA GGGGAGGCTT CT	203' DMSO	3	1	1
OT5 -34	CATTCTTA GTCATTTCC TGAGG	203' 5	CCAGGTGTGG TGGTCATGA CC	GCATACGCCA TGATAATGAG CC	203' 68C, DMSO a 3%	4	0	1
OT5 -35	TGCAGCTA ACTCATTAC CTGCGG	203' 5	CAGGCGCTGG GTTCTTAGCC T	CCTCCTGGG CCCCATGGTG	203' DMSO	2	3	0
OT5 -36	TTCCTTTA GTTATTACCA TGAGG	203' 5	TGGGGTCCAA GATGCCCCCT TGAGG	TGAAACTGCTT GATGAGGTGT GGA	204' DMSO	1	2	2
OT5 -37	AACTTGAA AGTCATTAC CTGTGG	204' 6	GCTGGGCTTG GTGGTATATG C	ACTTGCAAAGC TGATAACTGAC TGA	204' DMSO	5	0	1
OT5 -38	AAAGTCAC AGTCATTAC CTGGGG	204' 6	AGTTGGTGTG ACTGACAATG GGA	CGCAGCGCAC GAGTTCATCA	204' DMSO	3	0	3
OT5 -39	AATGTCTTC ATCATTACC TGAGG	204' 6	AGAGGAGGCA CAATTCAACC CCT	GGCTGGGGAG GCCTCACAAAT	204' DMSO	1	1	4
OT5 -40	AGATGCTT GGTCATTAC CTGTGG	205' 6	GGGAAAGTT GGGAAAGTC GCA	AGGACAAGCT ACCCACACCC	205' DMSO	1	3	2
OT5 -41	AGTAGATTAA GTTATTACCA TGAGG	205' 6	TGGTGCTATCA AAGGGTTGCT TCT	TCATTCAGCA CGCCGGGAG	205' DMSO	0	3	3
OT5 -42	AGTAGGTTA GTAATTACCA TGAGG	205' 6	CCCAGGCTGC CCATCACACT	TGGAGTAAGTA TACCTGGGG ACCT	205' DMSO	1	3	2
OT5 -43	CAAATGAG AGTCATTAC CTGAGG	205' 6	TCAGTCCCCC TGGGTCTCA	TGTGCAAATAC CTAGCACGGT GC	206' DMSO	4	2	0
OT5 -44	CATGTCCTGA ATCATTACCA TGAGG	206' 6	AGCACTCCCT TTTGAATTTC GTGCT	ACTGAAGTCCA GCCTCTTCCAT TTCA	206' DMSO	2	1	3
OT5 -45	CCTGACTT GGTCATTAC CTGTGG	206' 6	GAAACCGTC CCTGGTGCCA	GGGGAGTAGA GGGTAGTGT GCC	206' DMSO	2	0	4
OT5 -46	CGTGCATTA GTCATTACCA TGAGG	206' 6	TTGCGGGTCC CTGTGGAGTC	AGGTGCCGTG TTGTGCCCAA	207' DMSO	1	2	3
Alvo 6	GGAATCCC TTCTGCAGC ACCTGG	207' 0	GCCCTACATC TGCTCTCCCT CCA	GGGCCGGGAA AGAGTTGCTG	207' DMSO			
OT6 -1	GGAACCCC GTCTGCAG CACCAAGG	207' 2	TTGGAGTGTG GCCCGGGTTG	ACCTCTCTTTC TCTGCCTCACT GT	207' DMSO	0	1	1
OT6 -2	GGAACACC TTCTGCAGC TCAGG	207' 3	CACACCATGC TGATCCAGGC	GCAGTACGGA AGCACGAAAC	207' DMSO	1	1	1
OT6 -3	GGAAGCTC TGCTGCAG CACCTGG	208' 3	CTCCAGGGCT CGCTGTCCAC	CTGGGCTCTG CTGGTTCCCC	208' DMSO	0	2	1
OT6 -4	GGAATATCT TCTGCAGC CCCAGG	208' 3	CTGTGGTAGC CGTGGCCAGG	CCCCATACCCAC CTCTCCGGGA	208' DMSO	0	2	1

TABELA E											
OT6 -5	GGAATCAC TTTACAGC ACCAGG	208 ^f 3	GGTGGCGGG ACTTGAATGA G	208 ^f	CCAGCGTGT CCAAGGGAT	208 ^f	betaína a 1M, TD	0	1	2	
OT6 -6	GGAATCCC CTCTCCAG CCCCCTGG	208 ^f 3	GGAATCCCCTTC CAGCCCCCTGG (SEQ ID NO:2236) GGAATCCCCTTC CAGCCTCTGG(SE Q ID NO:2237)	CCAGAGGTGG GGCCCTGTGA	209 ^c	TTTCCACACTC AGTTCTGCAG GA	209 ^c DMSO	1	1	1 / 2	
OT6 -7	GGAATCTCT TCTTCAGCA TCTGG	209 ^f 3	GGAATCTCTTC TGGCATCTGG(SE Q ID NO:2238)	TGTGACTGGT TGTCTGTCTT CCT	209 ^c	GCAGTGTTTG TGGTGATGGG CA	209 ^c betaína a 1M, TD	0	1	5	
OT6 -8	GGAATTGCT TCTGCAGC GCCAGG	209 ^f 3		CTGGCCAAGG GGTGAGTGGG	209 ^c	TGGGACCCCA GCAGCCAATG	209 ^c DMSO	1	0	2	
OT6 -9	GGACTCCC CTCTGCAG CAGCTGG	209 ^f 3		ACGGTGTGCT GGCTGCTCTT	209 ^c	ACAGTGCTGA CCGTGCTGGG	210 ^c DMSO	1	1	1	
OT6 -10	GGAGTCCC TCCTACAG CACCAGG	210 ^f 3		TGGTTGGGC CTCAGGGATG G	210 ^c	TGCCCTCCACA AAAATGTCTAC CT	210 ^c DMSO	0	0	3	
OT6 -11	GGAGTCCC TCCTACAG CACCAGG	210 ^f 3		TGGTTGGGC CTCAGGGATG G	210 ^c	ACCCCTTATCC CAGAACCCAT GA	210 ^c DMSO	0	0	3	
OT6 -12	GGCATCCA TTCTGCAGC CCCTGG	210 ^f 3		TCCAAGTCAG CGATGAGGGC T	210 ^c	TGGGAGCTGT TCCTTTTGGC CA	210 ^c DMSO	0	3	0	
OT6 -13	GGCTTCCC TTCTGCAGC CCCAAGG	211 ^f 3		CACCCCTCTC AGCTCCCCAA	211 ^c	GCTAGAGGGT CTGCTGCCTT	211 ^c DMSO	1	2	0	
OT6 -14	TGAATCCCA TCTCCAGCA CCAGG	211 ^f 3		AGACCCCTTG GCCAAGCACA	211 ^c	CTTGCCTCAC CCCGCCTCC	211 ^c DMSO	2	1	0	
OT6 -15	AAAATACCT TCTGCAGTA CCAGG	211 ^f 4		ACATGTGGGA GGCGGACAGA	211 ^c	TCTCACTTTGC TGTTACCGATG TCG	211 ^c DMSO	0	1	3	
OT6 -16	AAAATCCCT TCTTCAACA CCTGG	211 ^f 4		GGACGACTGT GCCTGGGACA	212 ^c	AGTGCCTCAGA TGTTGTAACT GCT	212 ^c 72C Recorer, DMSO a 3%	0	1	3	
OT6 -17	ACACTCCCT CCTGCAGC ACCTGG	212 ^f 4		GGAGAGCTCA GCGCCAGGTC	212 ^c	CAGCGTGGCC CGTGGGAATA	212 ^c DMSO	1	1	2	
OT6 -18	ACCATCCCT CCTGCAGC ACCAGG	212 ^f 4		GCTGAAGTGC TCTGGGGTGC T	212 ^c	ACCCCACTGT GGATGAATTG GTACC	212 ^c DMSO	1	1	2	
OT6 -19	AGAGGCC CTCTGCAG CACCAGG	212 ^f 4		TCGGGGTGCA CATGCCATC	212 ^c	TTGCCTCGCA GGGGAAAGCAG	213 ^c DMSO	0	1	3	
OT6 -20	AGGATCCC TTGTGCAG CTCCCTGG	213 ^f 4		CTCGTGGGAG GCCAACACCT	213 ^c	AGCCACCAAC ACATACCAGG CT	213 ^c DMSO	2	0	2	
OT6 -21	CCACTCCCT TCTGCAGC ACCCGG	213 ^f 4		GCATGCCCTT AATCCCCGGCT	213 ^c	AGGATTTCAGA TGATGGGGC T	213 ^c DMSO	2	1	1	
OT6 -22	GAAGGCC TTCAGCAG CACCTGG	213 ^f 4		CGCCCCAGCA CAAAGTGCAT	213 ^c	GCAAATTCTG CACCTACTCTA GCCCT	213 ^c DMSO	1	1	2	
OT6 -23	GATATCCCT TCTGTATCA CCTGG	214 ^f 4		AGCTCACAA AATTGGGGT AACAGT	214 ^c	GCAGTCACCC TTCACTGCCTG T	214 ^c DMSO	1	1	2	
OT6 -24	GGGTCGCG TTCTGCAGC ACCTGG	214 ^f 4		AAACTGGGCT GGGCTCCGG	214 ^c	GGGGCTAAGG CATTGTCAGAC CC	214 ^c DMSO	2	0	2	
OT6 -25	GTCTCCCT TCTGCAGC ACCAGG	214 ^f 4		GCAGGTAGGC AGTCTGGGGC	214 ^c	TCTCCTGCCTC AGCCTCCCA	214 ^c betaína a 1M, TD	1	2	1	
OT6 -26	GTCTCCCT TCTGCAGC ACCAGG	214 ^f 4		GCAGGTAGGC AGTCTGGGGC	215 ^c	TCTCCTGCCTC AGCCTCCCA	215 ^c betaína a 1M, TD	1	2	1	
OT6 -27	GTCTCCCT TCTGCAGC ACCAGG	215 ^f 4		GCAGGTAGGC AGTCTGGGGC	215 ^c	TCTCCTGCCTC AGCCTCCCA	215 ^c betaína a 1M, TD	1	2	1	
OT6 -28	TCATTCCCG TCTGCAGC ACCCGG	215 ^f 4		GCTCTGGGGT AGAAGGAGGC	215 ^c	GGCTGTCAA CCAACCAACC	215 ^c DMSO	2	2	0	

TABELA E												
OT6	TGCACCCC TCCTGCAG CACCAGG	215 ^f 4	TGACATGTT TGTGCTGGC	215 ^c	AAATCCTGCAG CCTCCCCTT	216 ^c	DMSO	0	2	2		
OT6	TGCATACC CTCTGCAG CACCAGG	216 ^f 4	TCCTGGTGAG ATCGTCCACA GGA	216 ^c	TCCTCCCCACT CAGCCCTCCC	216 ^c	DMSO	0	3	1		
OT6	TGCATGGC TTCTGCAGC ACCAGG	216 ^f 4	TCCTAATCCAA GTCCTTTGTT AGACA	216 ^c	AGGGACCAGC CACTACCCCTC A	216 ^c	DMSO	2	2	0		
OT6	AATATTCCC TCTGCAGC ACCAGG	216 ^f 5	GGGACACCAG TTCCTTCAT	216 ^c	GGGGGAGATT GGAGTTCCCC	216 ^c	DMSO	1	0	4		
OT6	ACCAATTCT TCTGCAGC ACCTGG	217 ^f 5	ACACCACTAT CAAGGCAGAG TAGGT	217 ^c	TCTGCCTGGG GTGCTTCCC	217 ^c	DMSO	1	1	3		
OT6	AGCTCCCA TTCTGCAGC ACCCGG	217 ^f 5	CTGGGAGCGG AGGGAAGTGC	217 ^c	GCCCCGACAG ATGAGGCCCTC	217 ^c	DMSO	1	2	2		
OT6	CAGATT CCT GCTGCAGC ACCCGG	217 ^f 5 (SEQ ID NO:2239)	CAGATTACTGCTG CAGCACCGGG	217 ^c	CGGGTCTCGG AATGCCTCCA	217 ^c	ACCCAGGAATT GCCACCCCC	217 ^c	DMSO	1	2	3
OT6	CCAAGAGC TTCTGCAGC ACCTGG	217 ^f 5	TTGCTGTGGT CCCGGTGGT	218 ^c	GCAGACACTA GAGCCCGGCC	218 ^c	DMSO	3	2	0		
OT6	CCCAGCCC TGCTGCAG CACCCGG	218 ^f 5	GGTGTGGTGA CAGGTGGGT	218 ^c	ACCTGCGTCTC TGTGCTGCA	218 ^c	DMSO	2	3	0		
OT6	CCCCCTCCC TCCTGCAG CACCGGG	218 ^f 5	CTCCCAGGAC AGTGCTCGC	218 ^c	CCTGGCCCCA TGCTGCCCTG	218 ^c	DMSO	2	2	1		
OT6	CTACTGACT TCTGCAGC ACCTGG	218 ^f 5	TGCGTAGGTT TTGCCTCTGT GA	218 ^c	AGGGAAATGAT TTTTTCCACCC CCT	219 ^c	DMSO	2	3	0		
OT6	CTCCTCCCT CCTGCAGC ACCTGG	219 ^f 5	CTCCGCAGCC ACCGTTGTA	219 ^c	TGCATTGACGT ACGATGGCTC A	219 ^c	DMSO	1	3	1		
OT6	TCTGTCCCT CCTGCAGC ACCTGG	219 ^f 5	ACCTGCAGCA TGAACCTCG CA	219 ^c	ACCTGAGCAA CATGACTCACCC TGG	219 ^c	DMSO	2	1	2		
OT6	ACACAAACTTC CAGCACCTGG ACCTGG(SEQ ID NO:2240)			TCTCCAGTTTC TTGCTCTCAT GG	219 ^c	ACCATTTGTA ACCCAGTC betaina a 1M, TD	219 ^c	3 / 2	3	1		
OT6	ACTGTCATT TCTGCAGC ACCTGG	220 ^f 6	TGGGGTGGTG GTCTTGAACTC CA	220 ^c	TCAGCTATAAC CTGGGACTTGT GCT	220 ^c	DMSO	2	1	3		
OT6	ACTTTATCT TCTGCAGC ACCTGG	220 ^f 6	AGCAGCCAGT CCAGTGTCT G	220 ^c	CCCTTTCATCG AGAACCCAG GG	220 ^c	DMSO	3	1	2		
OT6	ATCCCTTCT TCTGCAGC ACCTGG	220 ^f 6	TGGACGCTGC TGGGAGGAGA	220 ^c	GAGGGTCTCGG GCTGCTCGTG	220 ^c	DMSO	0	3	3		
OT6	CACCAACG TTCTGCAGC ACCAGG	220 ^f 6	AGGTTGCAC TCTGTTGCC GG	221 ^c	TGGGGTGATT GGTTGCCAGG T	221 ^c	DMSO	3	2	1		
OT6	CATGTGGC TTCTGCAGC ACCTGG	221 ^f 6	TCTTCCTTGC CAGGCAGCAC A	221 ^c	TGCAGGAATA CAGGGTATGA GGAGT	221 ^c	DMSO	4	0	2		
OT6	CATTTCCTT TCTGCAGC ACCTGG	221 ^f 6	GGACGCCCTAC TGCTGGACC	221 ^c	CCCCCTGGCAG CCCATGGTAC	221 ^c	DMSO	3	0	3		
OT6	CTCTGTCTC TCTGCAGC ACCTGG	221 ^f 6	AGGCAGTCAT CGCCTTGCTA	221 ^c	GGTCCCACCTT CCCCTACAA	222 ^c	DMSO	2	3	1		
OT6	CTGTACCCCT CCTGCAGC ACCAGG	222 ^f 6	Não otimizado					3	1	2		
OT6	TTGAGGCC GTCTGCAG CACCGGG	222 ^f 6	CCCCAGCCCC CACCAAGTTTC	222 ^c	CAGCCCAGGC CACAGCTCA	222 ^c	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 *BamHI* codificado pelo modelo doador ssODN, 200 ng de produtos de PCR purificados foram digeridos com *BamHI* 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'-GGCGAGGGCGATGCCACCTAcGG-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 transfetadas 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 transfetados 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 *CLTA*), 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 *CLTA* (**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 a 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 *BamHI* codificado em ssODNs homólogos endógenos foram comparadas para estes genes. Em ambos os sítios, introdução de um sítio *BamHI* 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 a 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 àquelas 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 (%) ± s.e.m.		Alvo truncado	SEQ ID NO:	Frequência de mutação indel (%) ± s.e.m.		Gene
			U2OS.EGFP	HEK293			U2OS.EGFP	HEK293	
T1	GGGTGGGGGGAGTTGCTCCIGG	2242.	23,69 ± 1,99	6,98 ± 1,33	GTGGGGGGAGTTGCTCCIGG	2243.	23,93 ± 4,37	8,34 ± 0,01	VEGFA
OT1-3	GG <u>A</u> GGGAGGGAGTTGCTCCIGG	2244.	17,25 ± 2,97	7,26 ± 0,62	<u>A</u> TGG <u>A</u> GGGAGTTGCTCCIGG	2245.	N.D.	N.D.	IGDCC3
OT1-4	GG <u>G</u> GGGGAGGGAGTTGCTCCIGG	2246.	6,23 ± 0,20	2,66 ± 0,30	<u>G</u> AGGG <u>G</u> GGGAGTTGCTCCIGG	2247.	N.D.	N.D.	LOC116437
OT1-6	<u>C</u> GGGGGGAGGGAGTTGCTCCIGG	2248.	3,73 ± 0,23	1,41 ± 0,07	<u>G</u> GGG <u>A</u> GGGAGTTGCTCCIGG	2249.	N.D.	N.D.	CACNA2D
OT1-11	GG <u>G</u> GGGGAGGGAGTTGCTCCIGG	2250.	10,4 ± 0,7	3,61 ± 0,02	<u>G</u> GA <u>G</u> GGGGAGTTGCTCCIGG	2251.	N.D.	N.D.	
T3	GGTGAGTGAGTGTCGCGTGIGG	2252.	54,08 ± 1,02	22,97 ± 0,17	GAGTGAGTGTCGCGTGIGG	2253.	50,49 ± 1,25	20,05 ± 0,01	VEGFA
OT3-1	GGTGAGTGAGTGTC <u>T</u> GTGaGG	2254.	6,16 ± 0,98	6,02 ± 0,11	GAGTGAGTGTC <u>T</u> GTGaGG	2255.	N.D.	N.D.	(abParts)
OT3-2	<u>A</u> GTGAGTGAGTGTC <u>T</u> GTGgGG	2256.	19,64 ± 1,06	11,29 ± 0,27	GAGTGAGTGTC <u>T</u> GTGgGG	2257.	5,52 ± 0,25	3,41 ± 0,07	MAX
OT3-4	<u>G</u> CTGAGTGAGTG <u>T</u> ATGCGTGIGG	2258.	7,95 ± 0,11	4,50 ± 0,02	GAGTGAGTG <u>T</u> ATGCGTGIGG	2259.	1,69 ± 0,26	1,27 ± 0,10	
OT3-9	GGTGAGTGAGTG <u>C</u> GTGCG <u>G</u> IGG	2260.	N.D.	1,09 ± 0,17	GAGTGAGTG <u>C</u> GTGCG <u>G</u> IGG	2261.	N.D.	N.D.	TPCN2
OT3-17	<u>G</u> ITGAGTGAA <u>T</u> GTGCGTGIGG	2262.	1,85 ± 0,08	N.D.	GAGTGAA <u>T</u> GTGCGTGIGG	2263.	N.D.	N.D.	SLIT1
OT3-18	<u>I</u> GTG <u>G</u> GTGAGTGTCGCGTGIGG	2264.	6,16 ± 0,56	6,27 ± 0,09	<u>G</u> GTGAGTGTCGCGTGIGG	2265.	N.D.	N.D.	COMDA
OT3-20	<u>A</u> GA <u>G</u> AGTGAGTGTCG <u>A</u> TGIGG	2266.	10,47 ± 1,08	4,38 ± 0,58	GAGTGAGTGTCG <u>A</u> TGIGG	2267.	N.D.	N.D.	
T4	GAGTCCGAGCAGAAGAAGAAGG	2268.	41,56 ± 0,20	12,65 ± 0,31	GTCCGAGCAGAAGAAGAAGG	2269.	43,01 ± 0,87	17,25 ± 0,64	EMX1
OT4-1	GAGT <u>A</u> AGCAGAAGAAGAAAGG	2270.	19,26 ± 0,73	4,14 ± 0,66	<u>G</u> TT <u>A</u> AGCAGAAGAAGAAAGG	2271.	N.D.	N.D.	HCN1
OT-4_Hsu31	GAGT <u>C</u> TA <u>G</u> CAAGAAGAAGAA <u>G</u> AG	2272.	4,37 ± 0,58	N.D.	<u>G</u> T <u>C</u> TA <u>G</u> CAAGAAGAAGAA <u>G</u> AG	2273.	N.D.	N.D.	MFAP1

Frequências de mutação foram medidas pelo ensaio de T7EI. Médias de medições em duplicita 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 regiao de complementaridade

Tabela 3C Frequências de mutação indel em sítios fora do alvo e 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 1 de VEGFA	<u>GT</u> GGGGGGAGTTGCTCCtGG	2274.	0 (sobre o alvo)	23,93 ±4,37	8,34 ± 0,01	
	<u>GT</u> GGGGGGAGTTGC <u>CCC</u> aGG	2275.	1	Não detectado	Não detectado	
	<u>GT</u> GGGGGG <u>I</u> GTTTGCTCCcGG	2276.	1	Não detectado	Não detectado	
	<u>GT</u> GGG <u>I</u> GGAGTTGCT <u>A</u> ctGG	2277.	2	Não detectado	Não detectado	
	<u>GT</u> GGGGGGAG <u>C</u> TTTGCTCCtGG	2278.	2	Não detectado	Não detectado	
	<u>GT</u> GGG <u>I</u> GG <u>C</u> GT <u>T</u> TGCTCCaGG	2279.	2	Não detectado	Não detectado	
	<u>GT</u> GG <u>A</u> GGGAG <u>C</u> TTGCTCCtGG	2280.	2	6.88 ± 0.19	Não detectado	
	<u>GT</u> GGG <u>I</u> GGAG <u>T</u> TTGCT <u>A</u> CaGG	2281.	2	Não detectado	Não detectado	
	<u>GG</u> GGGGGG <u>C</u> AGTTTGCTCCtGG	2282.	2	Não detectado	Não detectado	
	<u>GT</u> GT <u>GGGGAA</u> TTTGCTCCaGG	2283.	2	Não detectado	Não detectado	
	<u>C</u> T <u>G</u> CT <u>GGGAG</u> TTTGCTCCtGG	2284.	3	Não detectado	Não detectado	
	<u>TT</u> IGGG <u>A</u> G <u>G</u> TTTGCTCCaGG	2285.	3	Não detectado	Não detectado	
	<u>C</u> T <u>G</u> AGGG <u>C</u> AG <u>T</u> TTGCTCCaGG	2286.	3	Não detectado	Não detectado	
	<u>G</u> T <u>A</u> AGGG <u>A</u> AG <u>T</u> TTGCTCCtGG	2287.	3	Não detectado	Não detectado	
	<u>GG</u> GGGT <u>AG</u> AG <u>T</u> TTGCTCCaGG	2288.	3	Não detectado	Não detectado	
	<u>GG</u> GT <u>GGGGAC</u> TTTGCTCCaGG	2289.	3	Não detectado	Não detectado	
	<u>GG</u> GG <u>G</u> AG <u>C</u> AG <u>T</u> TTGCTCCaGG	2290.	3	Não detectado	Não detectado	
	<u>T</u> T <u>GGGG</u> <u>I</u> TAG <u>T</u> TTGCTCCtGG	2291.	3	Não detectado	Não detectado	
	<u>T</u> T <u>G</u> AGGGGG <u>A</u> G <u>T</u> CTGCTCCaGG	2292.	3	Não detectado	Não detectado	
	<u>C</u> T <u>GGGG</u> <u>I</u> GA <u>I</u> TTGCTCCtGG	2293.	3	Não detectado	Não detectado	
	<u>G</u> A <u>G</u> AGGGGG <u>A</u> G <u>T</u> TTGCTCCtGG	2294.	3	Não detectado	Não detectado	
	<u>T</u> T <u>GGGGGGAG</u> TTTGCTCCaGG	2295.	3	Não detectado	Não detectado	
	<u>T</u> T <u>C</u> GGGGGG <u>A</u> G <u>T</u> TTGCTCC <u>C</u> GG	2296.	3	Não detectado	Não detectado	
	<u>C</u> T <u>CGGGGGAG</u> TTTGCT <u>CC</u> aGG	2297.	3	Não detectado	Não detectado	
	<u>GT</u> GT <u>GGGGAG</u> T <u>C</u> TGCTCCaGG	2298.	3	Não detectado	Não detectado	
	<u>G</u> A <u>G</u> GGGGGG <u>C</u> AG <u>T</u> TTGCTCCaGG	2299.	3	Não detectado	Não detectado	
	<u>G</u> A <u>G</u> GGGGGG <u>A</u> G <u>T</u> TTTGT <u>CC</u> aGG	2300.	3	Não detectado	Não detectado	
	<u>GT</u> GG <u>C</u> GGAG <u>T</u> TTGCT <u>G</u> tGG	2301.	3	Não detectado	Não detectado	
	<u>GT</u> C <u>GGGGGGAG</u> T <u>GGG</u> CTCCaGG	2302.	3	Não detectado	Não detectado	
	<u>G</u> A <u>G</u> GGGGGG <u>A</u> G <u>T</u> G <u>T</u> TC <u>C</u> GG	2303.	3	Não detectado	Não detectado	
	<u>GT</u> GG <u>I</u> GG <u>G</u> AG <u>T</u> TTG <u>I</u> T <u>CC</u> iGG	2304.	3	Não detectado	Não detectado	
	<u>GT</u> GGGGGG <u>G</u> T <u>GG</u> CTGCTCCaGG	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
				50,49 ± 1,25	20,05 ± 0,01
Sítio 3 de VEGFA	<u>GAGTGAGTGTGCGGTGtGG</u>	2306.	0 (sobre o alvo)	Não detectado	Não detectado
	<u>CAGTGAGTGTGCGGTGtGG</u>	2307.	1	Não detectado	Não detectado
	<u>TGTGAGTGTGCGGTGgGG</u>	2308.	1	Não detectado	Não detectado
	<u>TGTGAGTGTGCGGTGaaGG</u>	2309.	1	Não detectado	Não detectado
	<u>TGTGAGTGTGCGGTGtGG</u>	2310.	1	Não detectado	Não detectado
	<u>GAGTGTGTGTGCGGTGtGG</u>	2311.	1	Não detectado	Não detectado
	<u>GAGTGGGTGTGCGGTGgGG</u>	2312.	1	Não detectado	Não detectado
	<u>GAGTGA<u>CT</u>GTGCGGTGtGG</u>	2313.	1	Não detectado	Não detectado
	<u>GAGTGA<u>GT</u>GTGCGGTGgGG</u>	2314.	1	Não detectado	Não detectado
	<u>GAGTGA<u>GT</u>GTGCGGTGtGG</u>	2315.	1	Não detectado	Não detectado
	<u>GAGTGA<u>GT</u>GTGCGGTGtGG</u>	2316.	1	Não detectado	Não detectado
	<u>GAGTGA<u>GT</u>GTGCGGTGgGG</u>	2317.	1	Não detectado	Não detectado
	<u>GAGTGA<u>GT</u>GTGCGGTGtGG</u>	2318.	1	Não detectado	Não detectado
	<u>GAGTGA<u>GT</u>GTGCGGTGCGG</u>	2319.	1	Não detectado	Não detectado
	<u>CTGTGAGTGTGCGGTGaaGG</u>	2320.	2	Não detectado	Não detectado
	<u>ATGTGAGTGTGCGGTGtGG</u>	2321.	2	Não detectado	Não detectado
	<u>CCCTGAGTGTGCGGTGtGG</u>	2322.	2	Não detectado	Não detectado
	<u>TGTG<u>GT</u>GTGCGGTGtGG</u>	2323.	2	Não detectado	Não detectado
	<u>TGTG<u>GG</u>GTGCGGTGtGG</u>	2324.	2	Não detectado	Não detectado
	<u>GGGTG<u>GT</u>GTGCGGTGtGG</u>	2325.	2	Não detectado	Não detectado
	<u>TGTG<u>GT</u>GTGCGGTGtGG</u>	2326.	2	Não detectado	Não detectado
	<u>TGTG<u>GG</u>GTGCGGTGtGG</u>	2327.	2	Não detectado	Não detectado
	<u>TGTG<u>GT</u>GTGCGGTGtGG</u>	2328.	2	Não detectado	Não detectado
	<u>GAGAGAG<u>GT</u>GTGCGGTGtGG</u>	2329.		Não detectado	Não detectado
	<u>GAGTGT<u>GT</u>GTGCGGTGtGG</u>	2330.	2	Não detectado	Não detectado
	<u>TGTGAGTGTG<u>GT</u>GTGtGG</u>	2331.	2	Não detectado	Não detectado
	<u>GAGTGT<u>GT</u>GTGCGGTGtGG</u>	2332.	2	Não detectado	Não detectado
	<u>GAGTC<u>AGT</u>GTGCGGTGaaGG</u>	2333.	2	Não detectado	Não detectado
	<u>GAGTGT<u>GT</u>GTGCGGTGtGG</u>	2334.	2	Não detectado	Não detectado
	<u>GAGTGT<u>GT</u>GTGCG<u>CA</u>TGtGG</u>	2335.	2	Não detectado	Não detectado
	<u>GAGTGA<u>AG</u>GTGCG<u>GT</u>tGG</u>	2336.	2	Não detectado	Não detectado
	<u>GAGTGA<u>GT</u>GTGCG<u>AG</u>aaGG</u>	2337.	2	Não detectado	Não detectado
Sítio 1 de EMX1	<u>GTCCGAGCAGAAAGAAGAAgGG</u>	2338.	0 (sobre o alvo)	43,01 ± 0,87	17,25 ± 0,64
	<u>GTC<u>GT</u>GAGCAGAAAGAAGAA<u>At</u>GG</u>	2339.	1	Não detectado	Não detectado
	<u>GTCC<u>CA</u>GAG<u>CA</u>AAGAAGAA<u>At</u>GG</u>	2340.	2	Não detectado	Não detectado
	<u>GTCC<u>GG</u>GAG<u>GG</u>AAGAAGAA<u>Ag</u>GG</u>	2341.	2	Não detectado	Não detectado
	<u>GTC<u>AG</u>GG<u>AG</u>AAGAAGAA<u>Ag</u>GG</u>	2342.	2	Não detectado	Não detectado
	<u>GAC<u>AG</u>GAG<u>CA</u>AAGAAGAA<u>Ag</u>GG</u>	2343.	2	Não detectado	Não detectado
	<u>GTGG<u>GG</u>GAG<u>CA</u>AAGAAGAA<u>Ag</u>GG</u>	2344.	2	Não detectado	Não detectado
	<u>GT<u>AC</u>IAG<u>CA</u>AAGAAGAA<u>Ag</u>GG</u>	2345.	2	Não detectado	Não detectado
	<u>GTC<u>GT</u>GAG<u>CA</u>AAGAAGAA<u>At</u>GG</u>	2346.	2	Não detectado	Não detectado
	<u>GT<u>GC</u>TAG<u>CA</u>AAGAAGAA<u>Ag</u>GG</u>	2347.	2	Não detectado	Não detectado
	<u>IAC<u>AG</u>GAG<u>CA</u>AAGAAGAA<u>Ag</u>GG</u>	2348.	3	Não detectado	Não detectado
	<u>IACGG<u>GG</u>GAG<u>CA</u>AAGAAGAA<u>At</u>GG</u>	2349.	3	Não detectado	Não detectado
	<u>AA<u>CG</u>GAG<u>CA</u>AAGAAGAA<u>Ag</u>GG</u>	2350.	3	Não detectado	Não detectado
	<u>GAC<u>AC</u>GAG<u>CA</u>AAGAAGAA<u>Ag</u>GG</u>	2351.	3	Não detectado	Não detectado
	<u>CT<u>CG</u>GAT<u>CA</u>AAGAAGAA<u>Ag</u>GG</u>	2352.	3	Não detectado	Não detectado
	<u>G<u>AC</u>T<u>GG</u>GAG<u>CA</u>AAGAAGAA<u>Ag</u>GG</u>	2353.	3	Não detectado	Não detectado
	<u>IT<u>CC</u>C<u>GT</u>GAG<u>CA</u>AAGAAGAA<u>Ag</u>GG</u>	2354.	3	Não detectado	Não detectado
	<u>IT<u>CC</u>C<u>AC</u>GAG<u>CA</u>AAGAAGAA<u>At</u>GG</u>	2355.	3	Não detectado	Não detectado
	<u>CT<u>CT</u>I<u>AG</u>G<u>GG</u>GAG<u>CA</u>AAGAAGAA<u>Ag</u>GG</u>	2356.	3	Não detectado	Não detectado
	<u>AT<u>CC</u>A<u>AT</u>CAG<u>AA</u>GAG<u>AA</u>Ag<u>GG</u></u>	2357.	3	Não detectado	Não detectado
	<u>G<u>CC</u><u>CC</u><u>GT</u>GAG<u>CA</u>AAGAAGAA<u>Ac</u>GG</u>	2358.	3	Não detectado	Não detectado
	<u>AT<u>CC</u>A<u>AC</u>CAG<u>AA</u>GAG<u>AA</u>Ag<u>AA</u>GG</u>	2359.	3	Não detectado	Não detectado
	<u>G<u>AC</u>T<u>GA</u>G<u>GA</u>AAG<u>AA</u>AG<u>AA</u>Ag<u>AA</u>GG</u>	2360.	3	Não detectado	Não detectado
	<u>GT<u>GG</u>GA<u>TC</u>A<u>GA</u>A<u>GA</u>A<u>GA</u>Ag<u>AA</u>GG</u>	2361.	3	Não detectado	Não detectado
	<u>G<u>AC</u>A<u>G</u>A<u>GA</u>A<u>GA</u>A<u>GA</u>Ag<u>AA</u>Ag<u>AA</u>GG</u>	2362.	3	Não detectado	Não detectado

Tabela 3C Frequências de mutação indel em sítios fora do alvo de endogenous genes in human cells							potenciais of tru-RGNs		
Alvo ID	Sítio Alvo + PAM	SEQ ID NO:	Número de desemparelhamentos	Frequência de mutação Indel (%)					
				± s.e.m.					
				Células U2OS.EGFP	Células HEK293				
	GTC <u>A</u> GGCAGAAGAAGAA <u>a</u> GG	2363.	3	Não detectado	Não detectado				
	GTT <u>G</u> GAGAAGAAGAA <u>Ag</u> GG	2364.	3	Não detectado	Não detectado				
	GTA <u>A</u> AGAGAAGAAGAA <u>Ag</u> GG	2365.	3	Não detectado	Não detectado				
	CTCC <u>I</u> AGCAAAGAAGAA <u>t</u> GG	2366.	3	Não detectado	Não detectado				
	TTC <u>A</u> GGCAGGAGAAGAA <u>t</u> GG	2367.	3	Não detectado	Não detectado				
	GTT <u>G</u> GACAGGAGAAGAA <u>Ag</u> GG	2368.	3	Não detectado	Não detectado				
	GCC <u>I</u> GAGCAGAAG <u>Ag</u> GG	2369.	3	Não detectado	Não detectado				
	GTC <u>I</u> GAGGA <u>CA</u> AGAAGAA <u>t</u> GG	2370.	3	Não detectado	Não detectado				
	GTCC <u>CG</u> GA <u>AG</u> GAGAAGAA <u>Ag</u> GG	2371.	3	Não detectado	Não detectado				
	GGCC <u>CG</u> AGCAGAAGAA <u>AG</u> AcGG	2372.	3	Não detectado	Não detectado				
	GTC <u>CT</u> AGCAGGAGAAGAA <u>Ag</u> AG	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	Seqüência de sítio fora do alvo	S#	tru-RGN			Controle		
			Indel	WT	Freq.	Indel	WT	Freq
Sítio 1 de VEGFA	GTGGGGGGAGTTTGCC <u>Ca</u> GG	2374.	1500	225640	0,66%	3	135451	0,00%
	GTGGGGGG <u>I</u> GTTTGCTCC <u>c</u> GG	2375.	1552	152386	1,01%	0	86206	0,00%
	GTGG <u>G</u> GGAGTTGCT <u>A</u> cGG	2376.	1	471818	0,00%	0	199581	0,00%
	GTGG <u>G</u> GGAGTTGCT <u>A</u> CaGG	2377.	0	337298	0,00%	1	211547	0,00%
	GTGG <u>G</u> GG <u>C</u> GGTTGCTCC <u>c</u> GG	2378.	2	210174	0,00%	1	105531	0,00%
	GT <u>G</u> GGGGGA <u>A</u> TTTGCTCC <u>c</u> GG	2379.	673	715547	0,09%	1	387097	0,00%
	GTGGGGGGAG <u>C</u> TT <u>I</u> CTCC <u>i</u> GG	2380.	5	107757	0,00%	1	58735	0,00%
	GGGGGG <u>C</u> AGTTGCTCC <u>i</u> GG	2381.	1914	566548	0,34%	3	297083	0,00%
Sítio 3 de VEGFA	GT <u>G</u> TGAGTG <u>T</u> GCG <u>T</u> G <u>GG</u>	2382.	58	324881	0,02%	9	122216	0,01%
	GT <u>G</u> TGAGTG <u>T</u> GCG <u>T</u> G <u>GG</u>	2383.	532	194914	0,27%	11	73644	0,01%
	GAG <u>T</u> GG <u>G</u> GTG <u>T</u> GCG <u>T</u> G <u>GG</u>	2384.	70	237029	0,03%	10	178258	0,01%
	GAG <u>T</u> GG <u>A</u> CTG <u>T</u> GCG <u>T</u> G <u>GG</u>	2385.	6	391894	0,00%	0	239460	0,00%
	GAG <u>T</u> GG <u>A</u> G <u>T</u> G <u>T</u> G <u>G</u> G <u>GG</u>	2386.	15	160140	0,01%	10	123324	0,01%
	GT <u>G</u> TGAGTG <u>T</u> GCG <u>T</u> G <u>GG</u>	2387.	19	138687	0,01%	1	196271	0,00%
	CA <u>G</u> TGAGTG <u>T</u> GCG <u>T</u> G <u>GG</u>	2388.	78	546865	0,01%	41	355953	0,01%
	GT <u>G</u> TGAGTG <u>T</u> GCG <u>T</u> G <u>GG</u>	2389.	128	377451	0,03%	56	133978	0,04%
	GAG <u>T</u> GT <u>G</u> GTG <u>T</u> GCG <u>T</u> G <u>GG</u>	2390.	913	263028	0,35%	78	178979	0,04%
	GAG <u>T</u> GT <u>G</u> GTG <u>T</u> G <u>T</u> G <u>GG</u>	2391.	40	106933	0,04%	36	58812	0,06%
	GAG <u>T</u> GT <u>G</u> GTG <u>T</u> G <u>T</u> G <u>GG</u>	2392.	681	762999	0,09%	63	222451	0,03%
	GAG <u>T</u> GT <u>G</u> GTG <u>T</u> G <u>T</u> G <u>GG</u>	2393.	331	220289	0,15%	100	113911	0,09%
	GAG <u>T</u> GT <u>G</u> GTG <u>T</u> G <u>T</u> G <u>GG</u>	2394.	0	35725	0,00%	8	186495	0,00%
	GAG <u>T</u> GT <u>G</u> GTG <u>T</u> G <u>G</u> <u>GG</u>	2395.	94	246893	0,04%	16	107623	0,01%
Sítio 1 de EMX1	GTC <u>A</u> GAG <u>G</u> GAGAAGAA <u>Ag</u> GG	2396.	0	201483	0,00%	4	148416	0,00%
	GTC <u>A</u> GAG <u>G</u> GAGAAGAA <u>Ag</u> GG	2397.	10	545662	0,00%	5	390884	0,00%
	GT <u>C</u> TGAG <u>C</u> A <u>AG</u> AA <u>GA</u> AA <u>t</u> GG	2398.	2	274212	0,00%	0	193837	0,00%
	GTC <u>I</u> GAG <u>C</u> A <u>AG</u> AA <u>GA</u> AA <u>t</u> GG	2399.	440	375646	0,12%	10	256181	0,00%
	G <u>A</u> C <u>A</u> GAG <u>C</u> A <u>AG</u> AA <u>GA</u> AA <u>Ag</u> GG	2400.	2	212472	0,00%	1	158860	0,00%
	GT <u>A</u> CT <u>A</u> G <u>C</u> A <u>AG</u> AA <u>GA</u> AA <u>Ag</u> GG	2401.	152	229209	0,07%	103	157717	0,07%
	GT <u>G</u> GG <u>A</u> G <u>C</u> A <u>AG</u> AA <u>GA</u> AA <u>Ag</u> GG	2402.	50	207401	0,02%	36	111183	0,03%
	GT <u>C</u> CC <u>C</u> A <u>G</u> C <u>A</u> G <u>C</u> A <u>AG</u> AA <u>At</u> GG	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 VEGFA	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., Reyne, 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 transcrecional; 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)GUUUUAGAGCUAGAAAUAGCAAGGUAAAAUAAG GCUAGUCCG (SEQ ID NO: 1),

(X₁₇₋₁₈)GUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUU

AAAAUAAGGUAGUCCGUUAUC (SEQ ID NO: 2),
(X₁₇₋₁₈)GUUUUAGAGCUAUGCUGUUUUGGAAACAAAACA
GCAUAGCAAGUUAAAAUAAGGUAGUCCGUUAUC (SEQ ID NO:
3),
(X₁₇₋₁₈)GUUUUAGAGCUAGAAAAGCAAGUUAAAAUAAG
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 GGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCC GUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 8), ou uma porção ativa da mesma; UAGCAAGUUAAAAUAAGGCUAGGUCCGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC (SEQ ID NO: 2405) ou uma porção ativa da mesma;

AGCAUAGCAAGUUAAAAUAAGGCUAGGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGC (SEQ ID NO: 2407) ou uma porção ativa da mesma;

CAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 2409) ou uma porção ativa da mesma;

UAGCAAGUUAAAAUAAGGCUAGGUCCGUUAUCAACUUGAAAAAGU G (SEQ ID NO: 2410) ou uma porção ativa da mesma;

UAGCAAGUUAAAAUAAGGCUAGGUCCGUUAUCA (SEQ ID NO: 2411) ou uma porção ativa da mesma; ou UAGCAAGUUAAAAUAAGGCUAGGUCCG (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 GGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCC GUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 8); o cRNA é (X₁₇₋₁₈)GUUUUAGAGCUA (SEQ ID NO: 2404) e o tracrRNA é UAGCAAGUUAAAAUAAGGCUAGGUCCGUUAUCAACUUGAAAAAGU GGCACCGAGUCGGUGC (SEQ ID NO: 2405); ou o cRNA é (X₁₇₋₁₈)GUUUUAGAGCUAUGC (SEQ ID NO: 2408) e o tracrRNA é AGCAUAGCAAGUUAAAAUAAGGCUAGGUCCGUUAUCAACUUGAAA

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

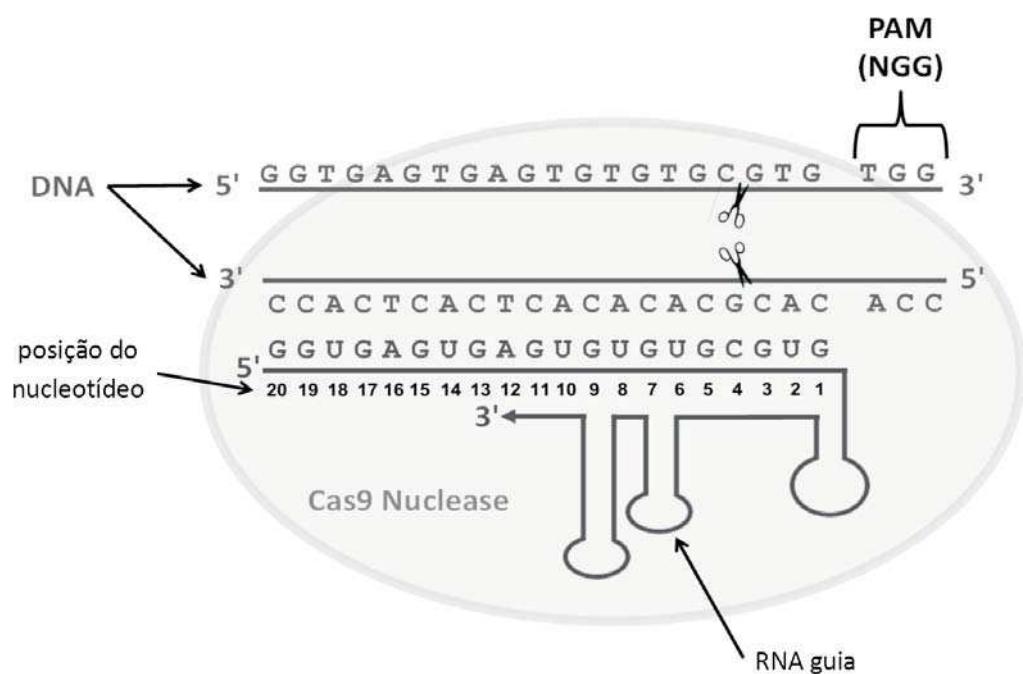


Figura 1

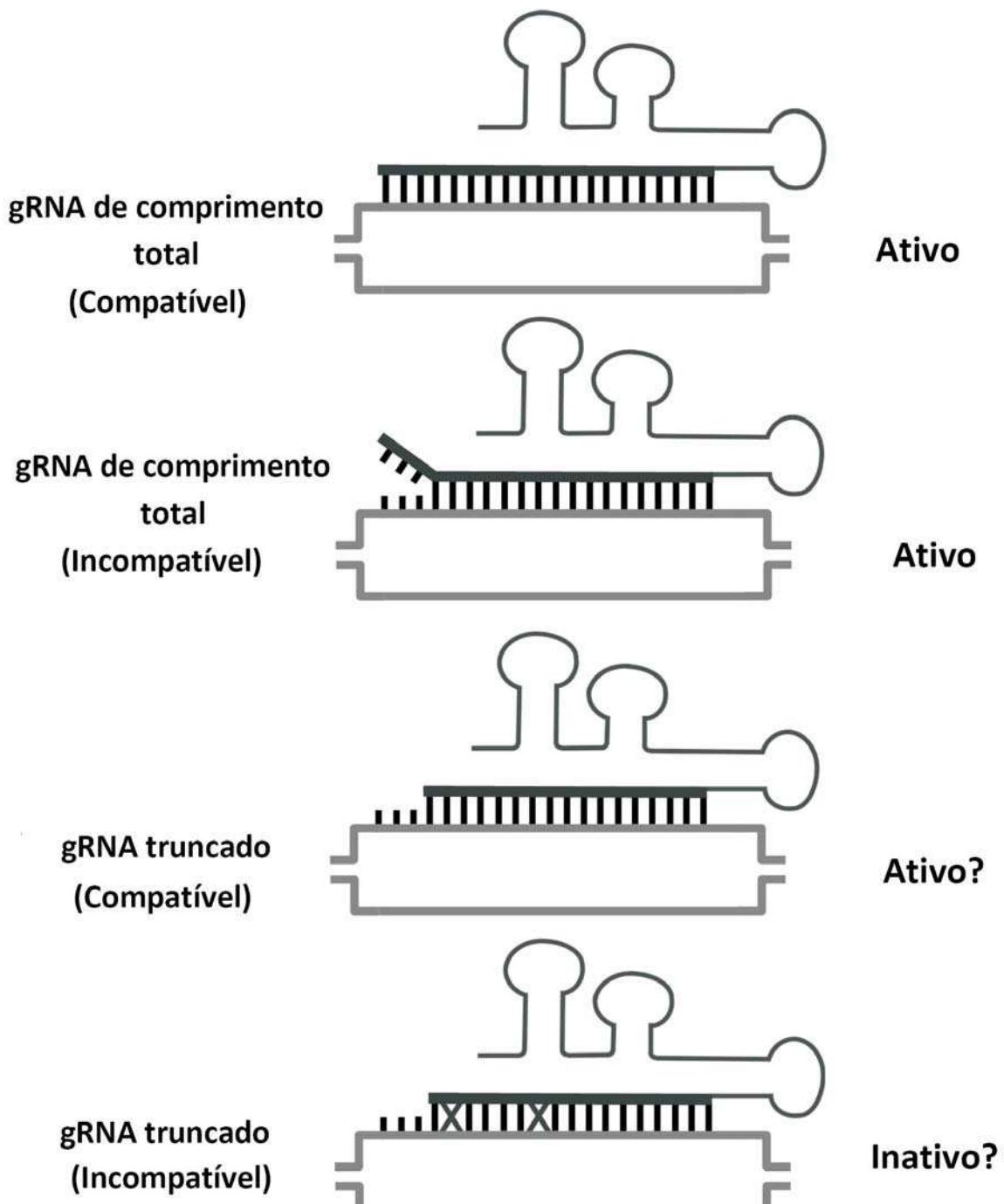


Figura 2A

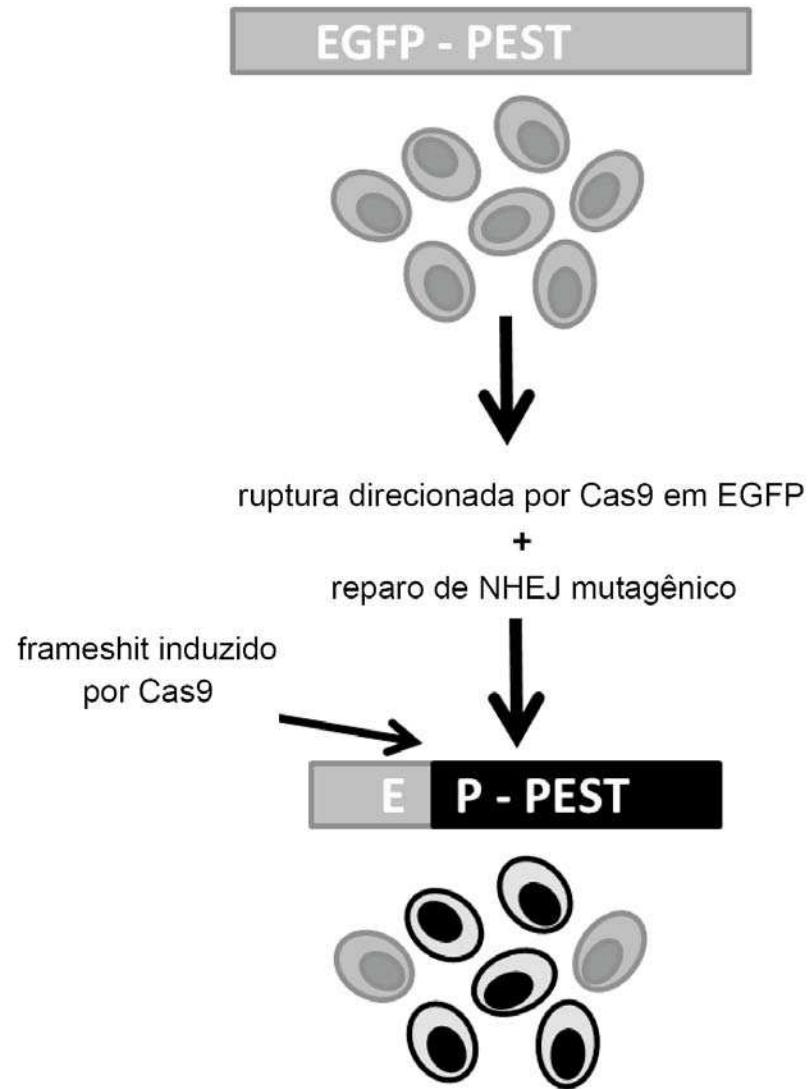


Figura 2B

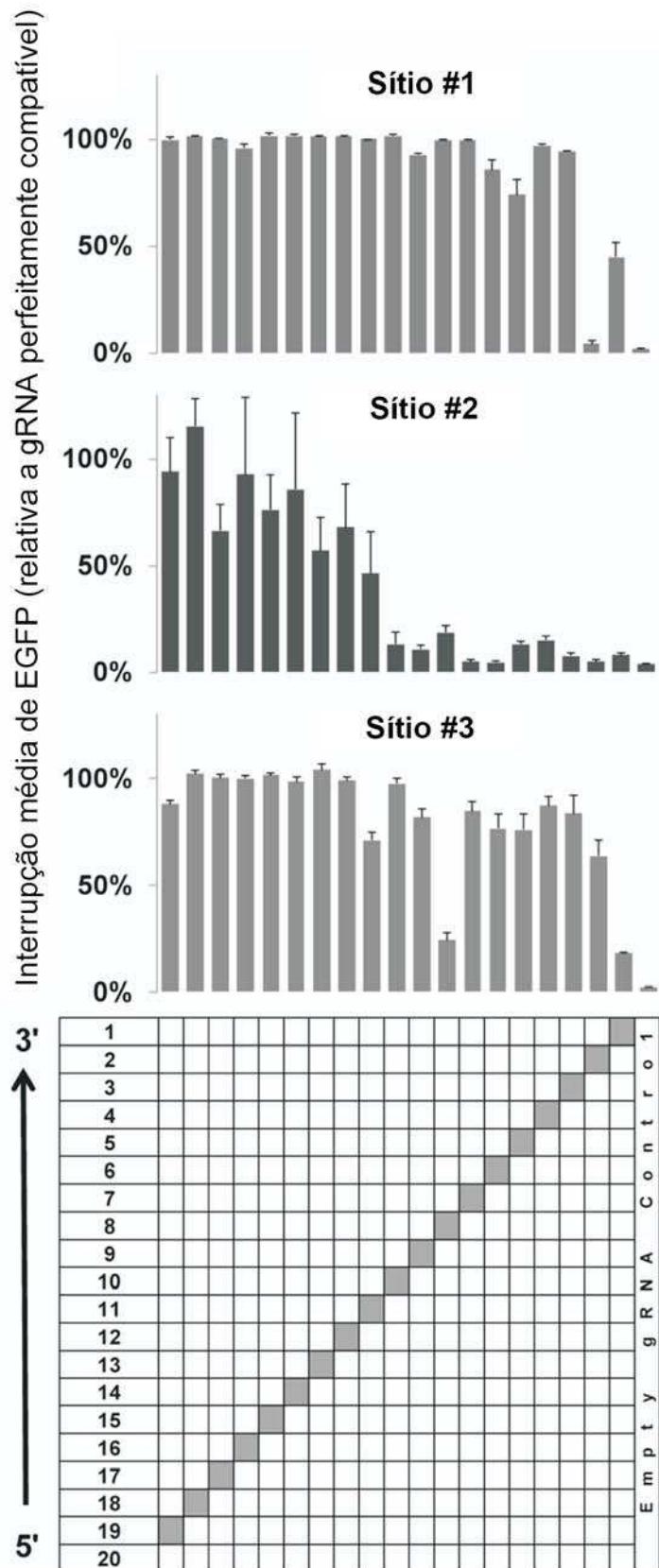


FIG. 2C

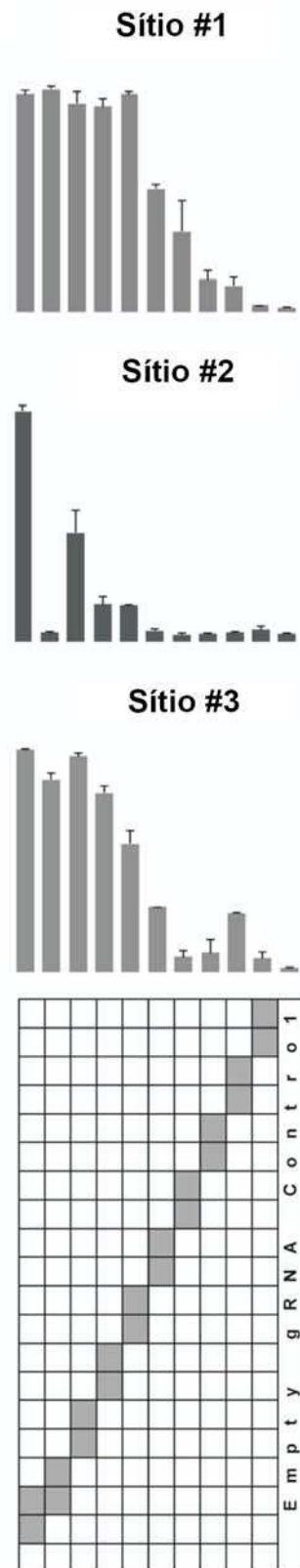
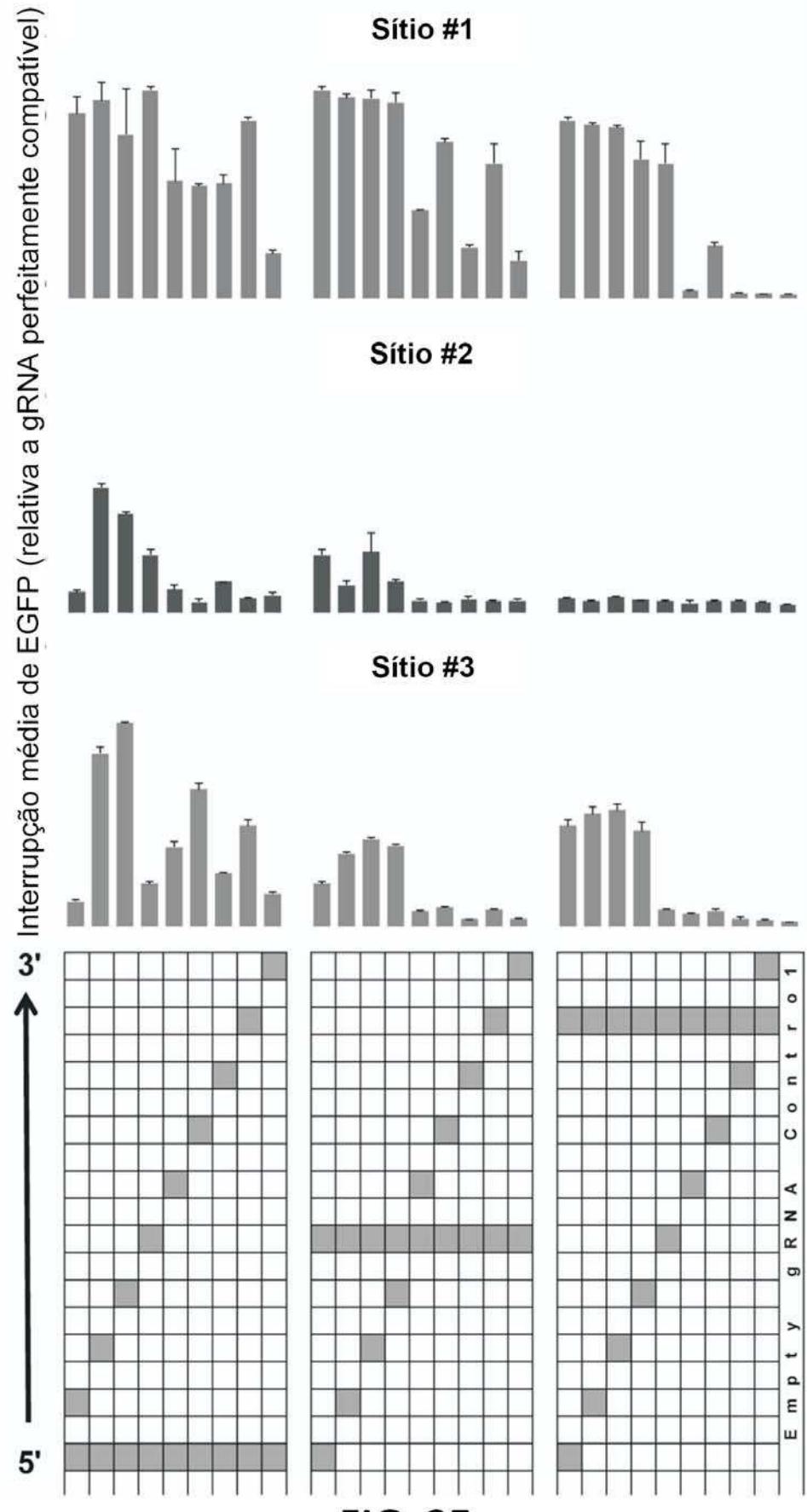
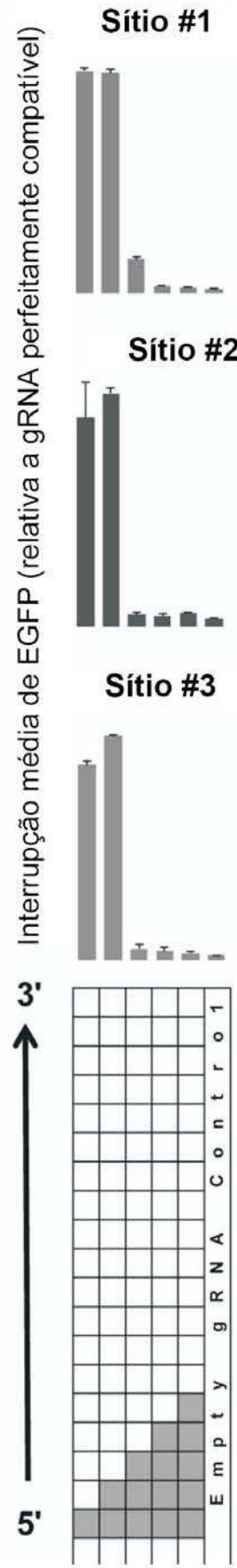
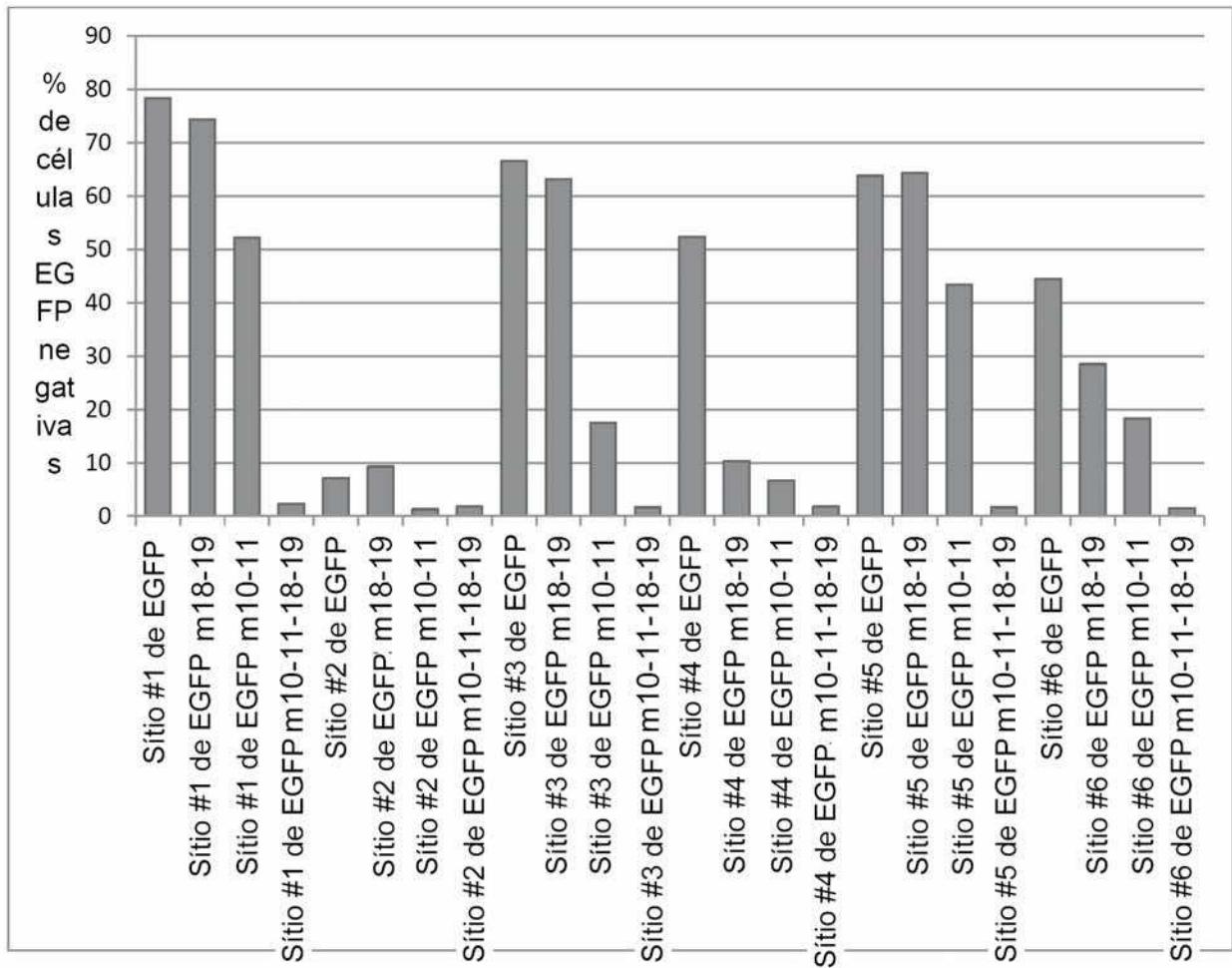
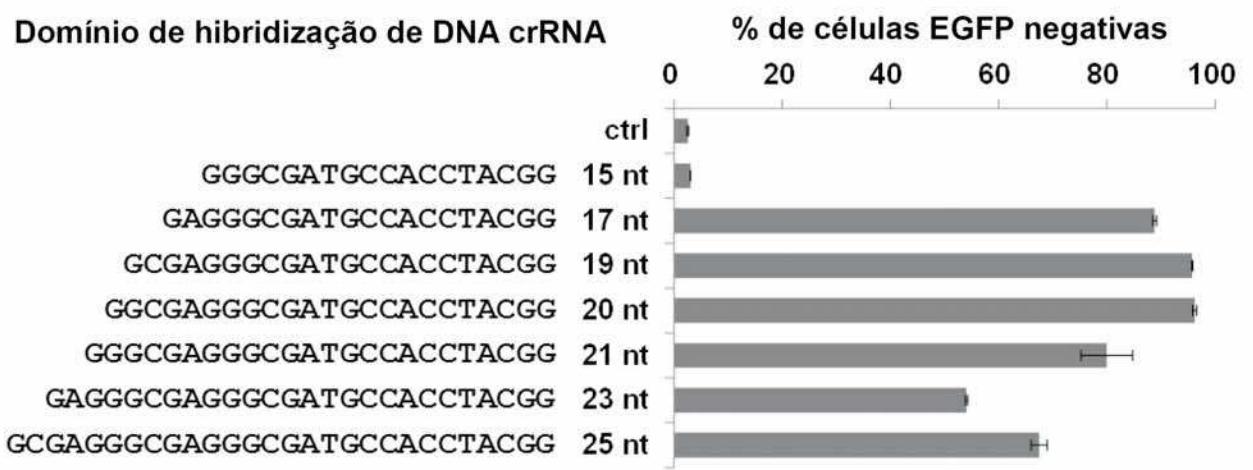
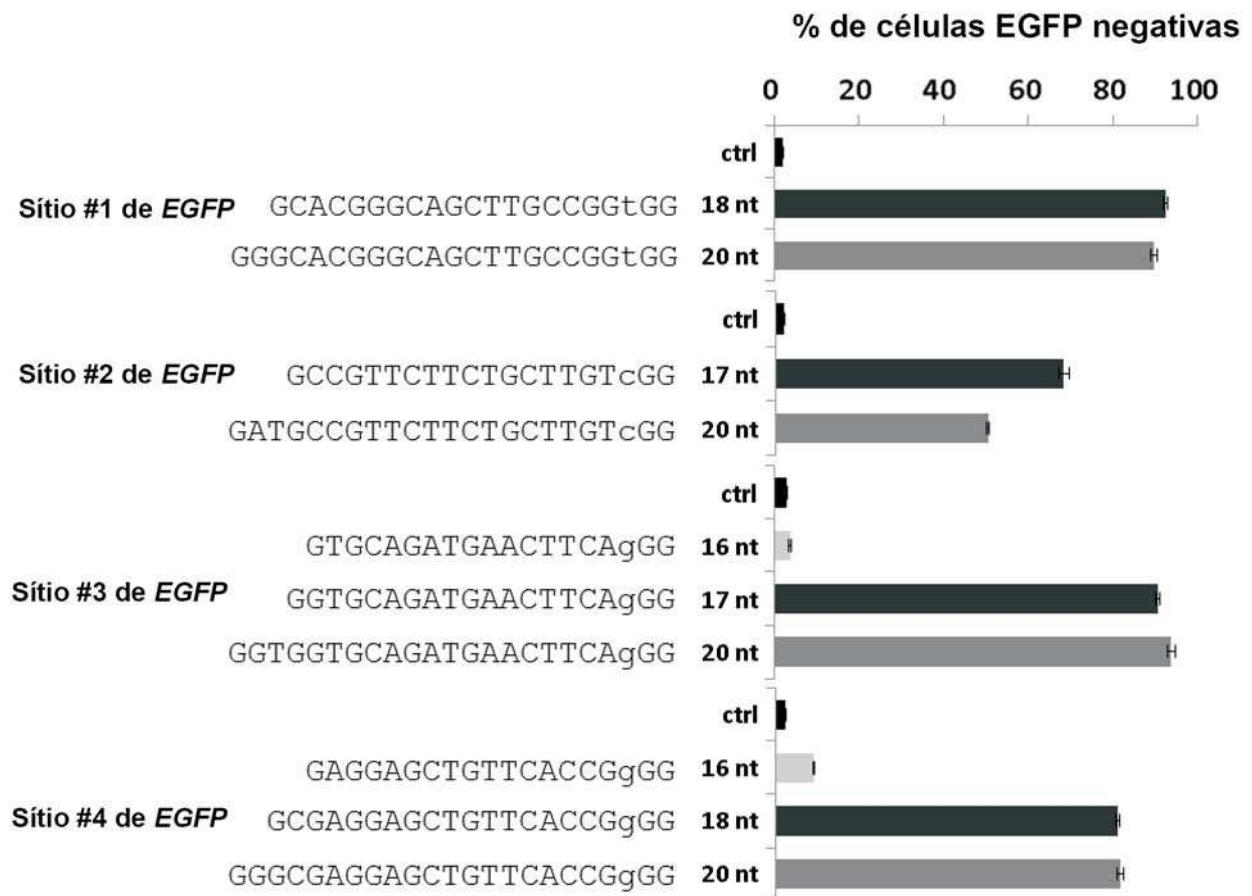


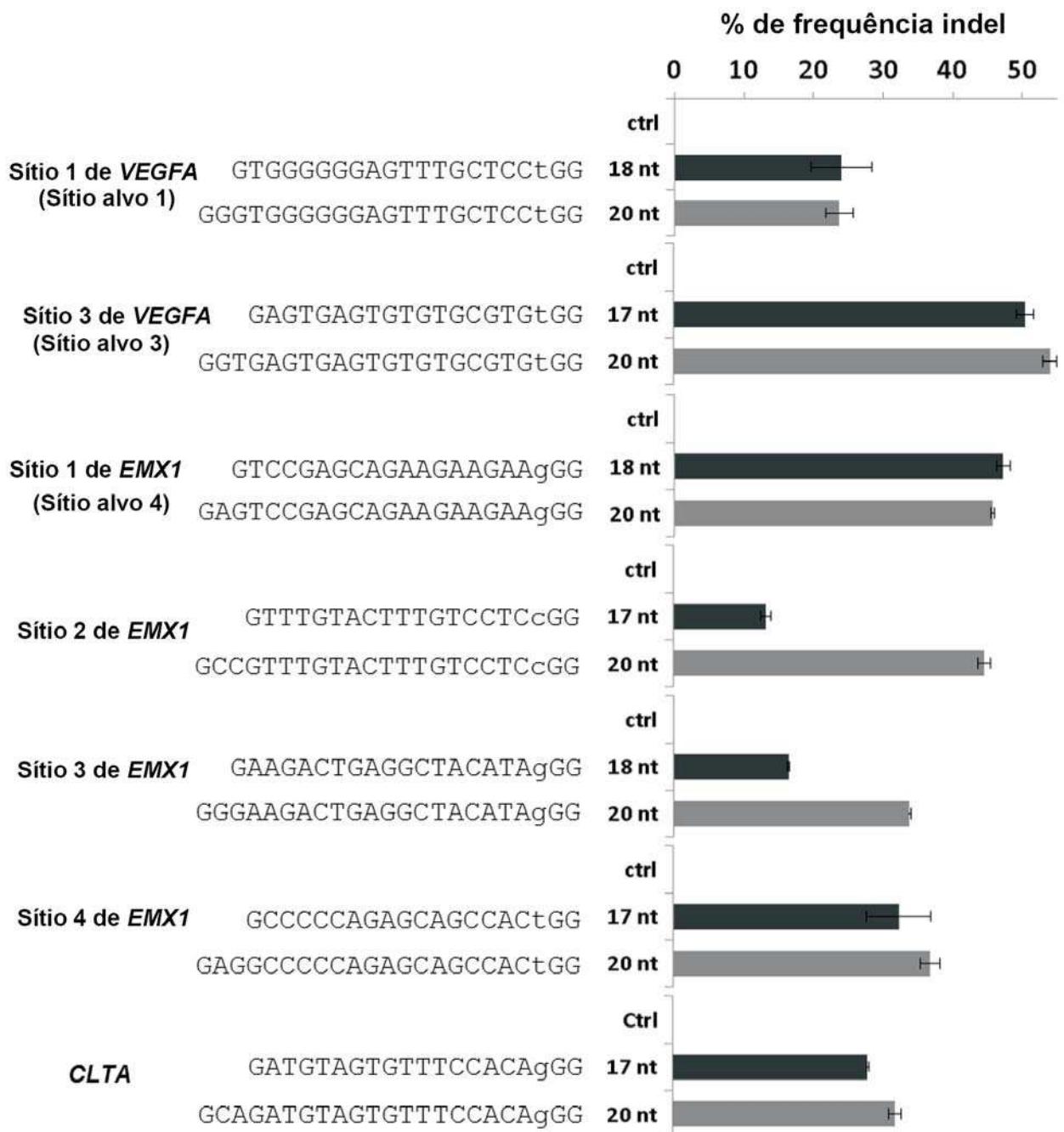
FIG. 2D

**FIG. 2E**

**FIG. 2F**

**FIG. 2G****FIG. 2H**

**FIG. 3A**

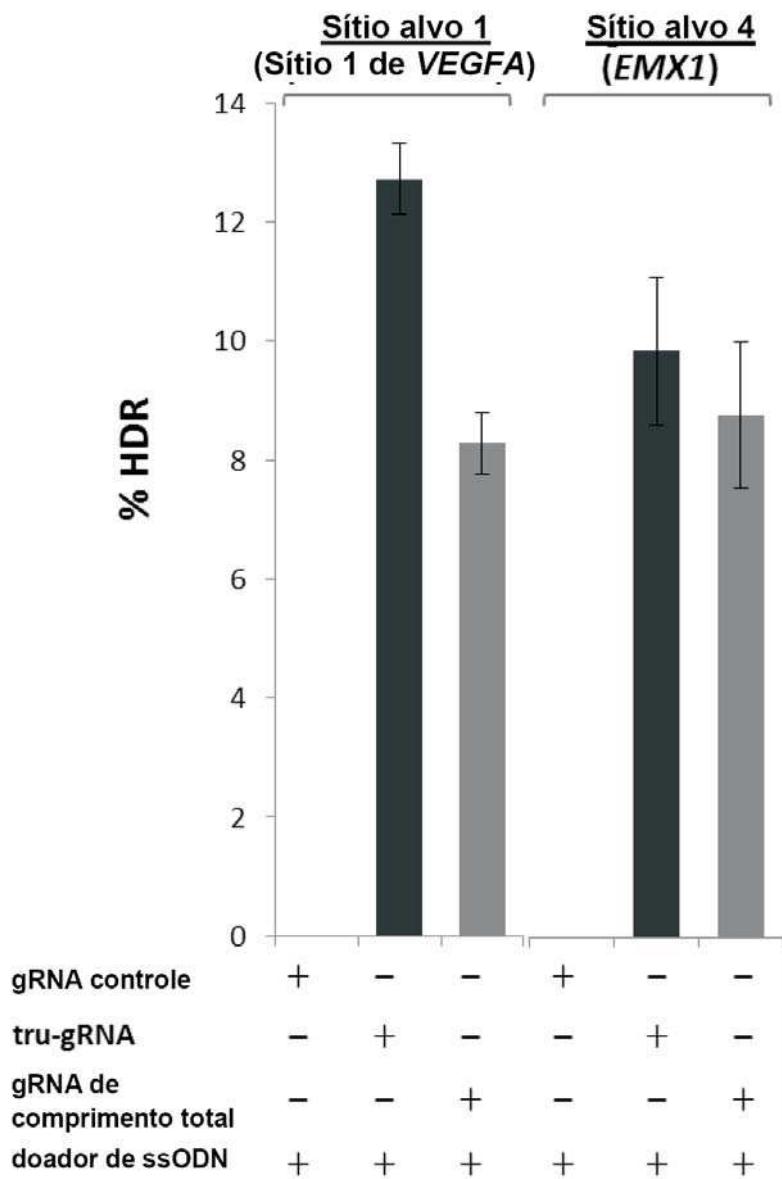
**FIG. 3B**

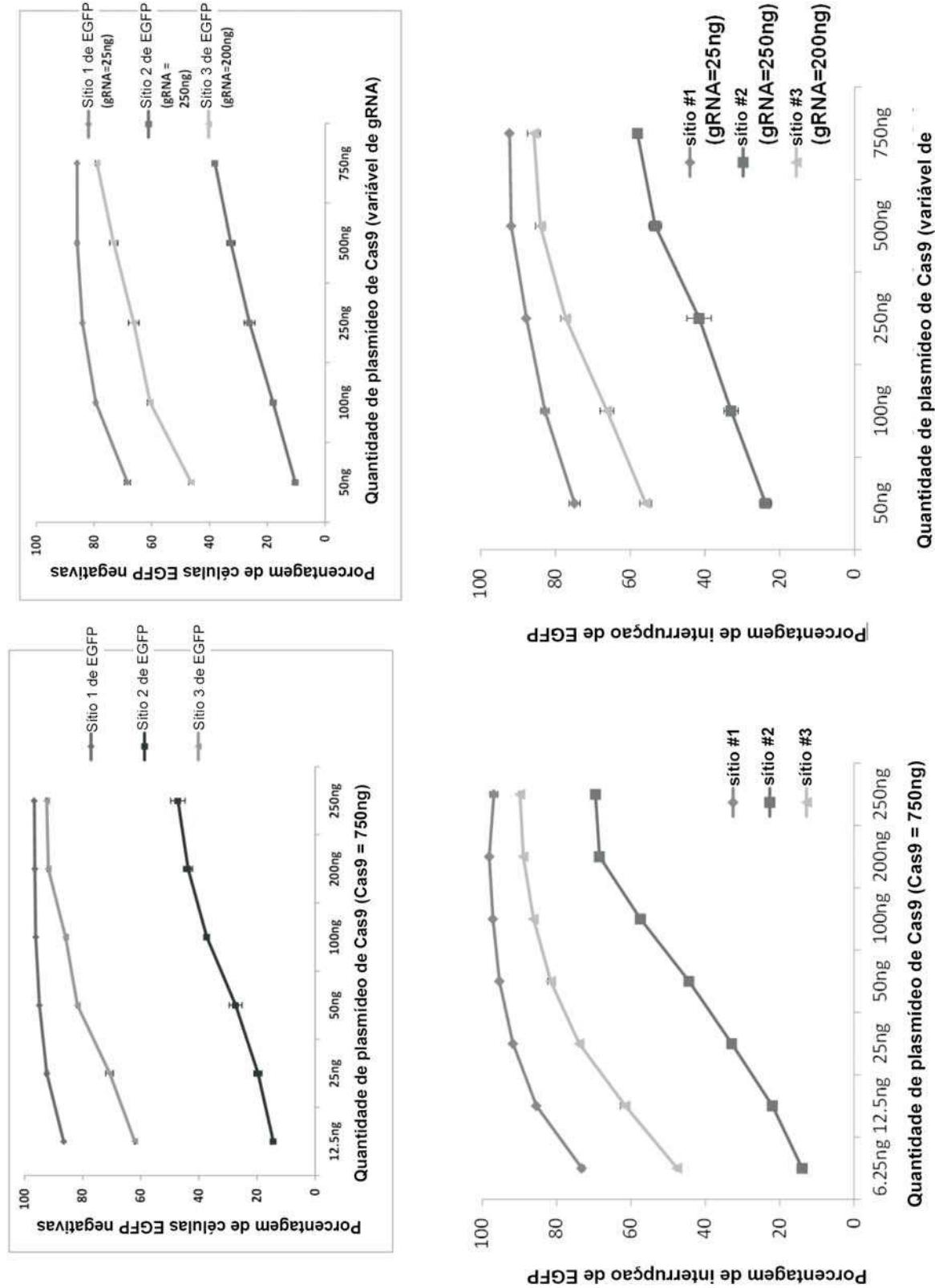
gRNA truncado de *EMX1*

gRNA de comprimento total de *EMX1*

GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGCTCCCATCACATCAACCGGTGG	tipo selvagem	x 35
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAG-----	>	Δ202
<-----	>	Δ115
GAA-----	>	Δ94
<-----	>	Δ78
GAAGCTGGAGG-----	>	Δ72
GAAGCTGGA-----	GG	Δ56
GAAGCTGGAGGAGGAAGGGCCTGA-----	GTGG	Δ39
GAAGCTGGAGGAGGAG-----	GAAGGGCTCCCATCACATCAACCGGTGG	Δ26 x 2
GAAGCTGGAGGAGGAAGGGCCTGAGT-----	CCATCACATCAACCGGTGG	Δ22
GAAGCTGGAGGAGGAAGGGCCTGAG-----	TCCCATCACATCAACCGGTGG	Δ21 x 3
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAG-----	CATCACATCAACCGGTGG	Δ18
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGA-----	GCTCCCATCACATCAACCGGTGG	Δ14
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGC-----	AGAAGGGCTCCCATCACATCAACCGGTGG	Δ6 x 3
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGC-----	AGAAGAAGGGCTCCCATCACATCAACCGGTGG	Δ3 x 3
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGA-----	AAGAAGGGCTCCCATCACATCAACCGGTGG	Δ2 x 2
GAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAACAGAAGGGCTCCCATCACATCAACCGGT	+ 2	

FIG. 3C

**FIG. 3D**



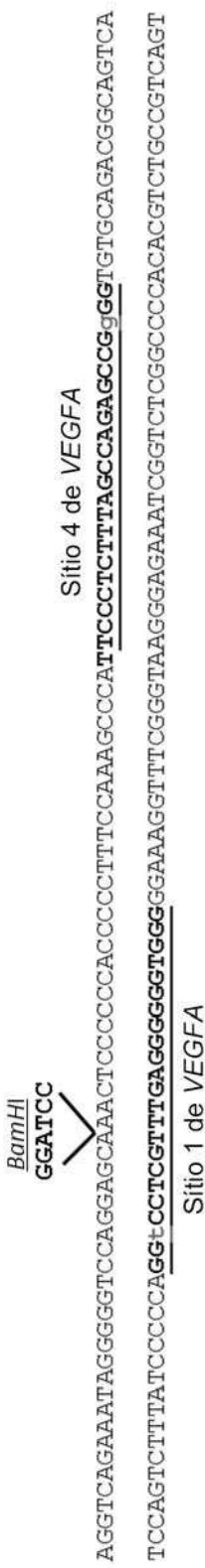
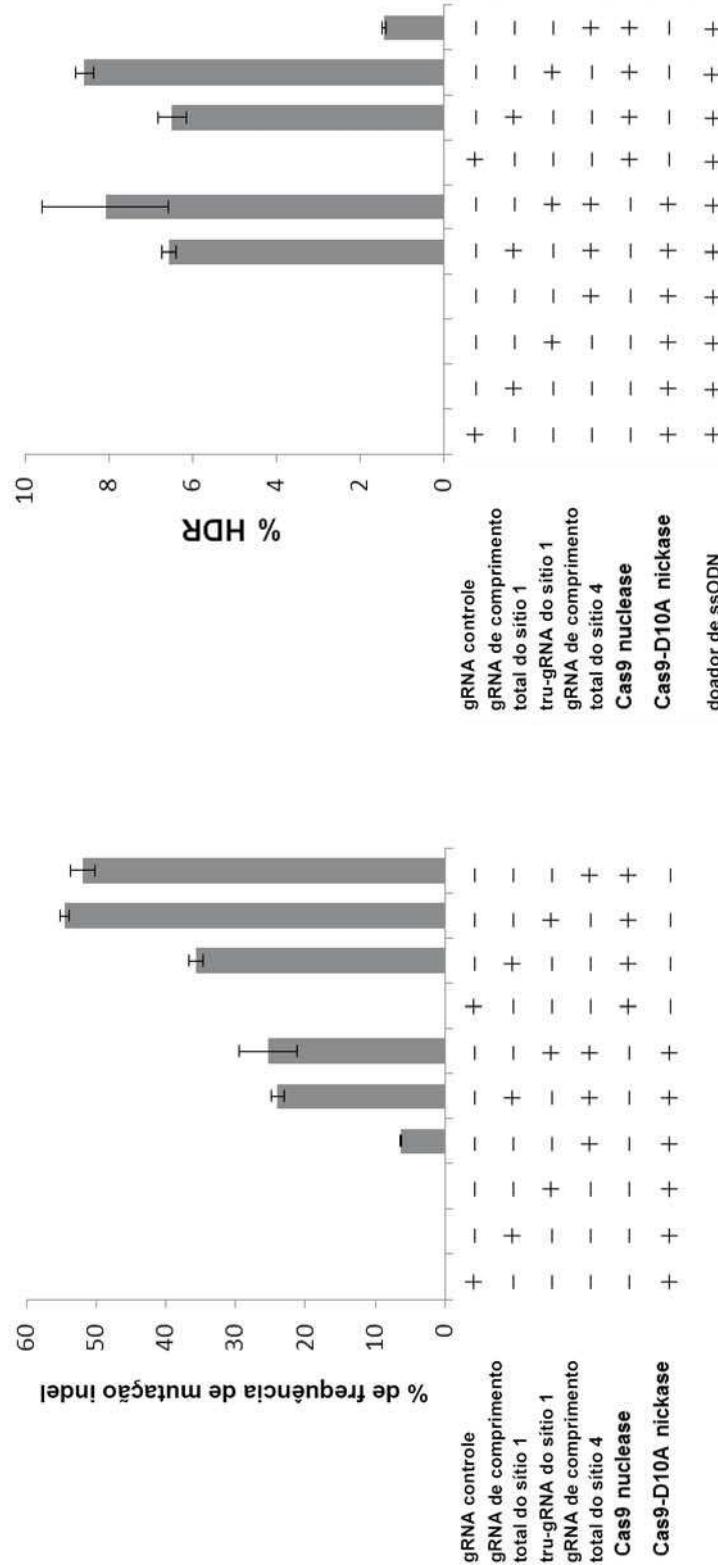
**FIG. 4A****FIG. 4B****FIG. 4C**

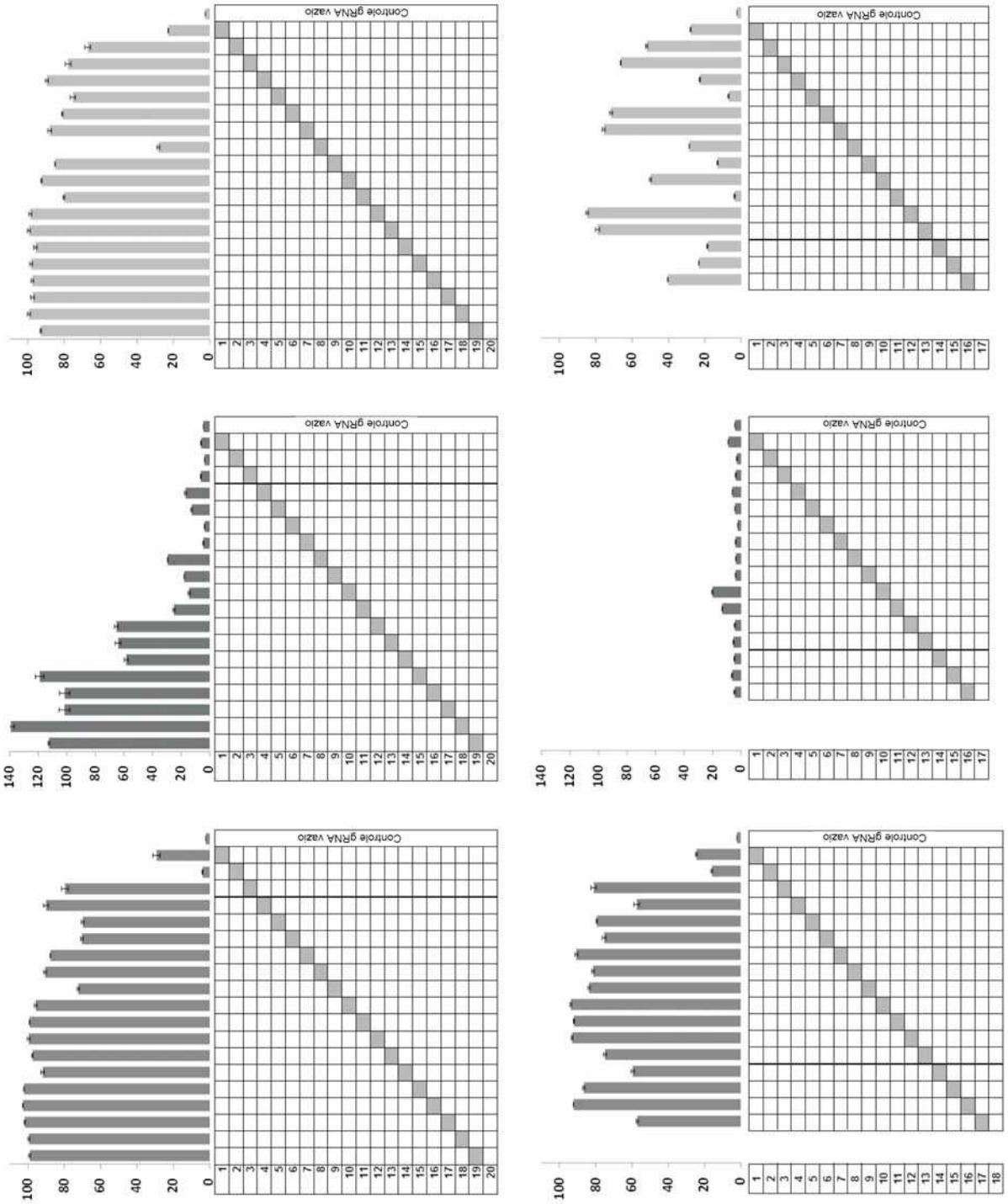
Figura 5A

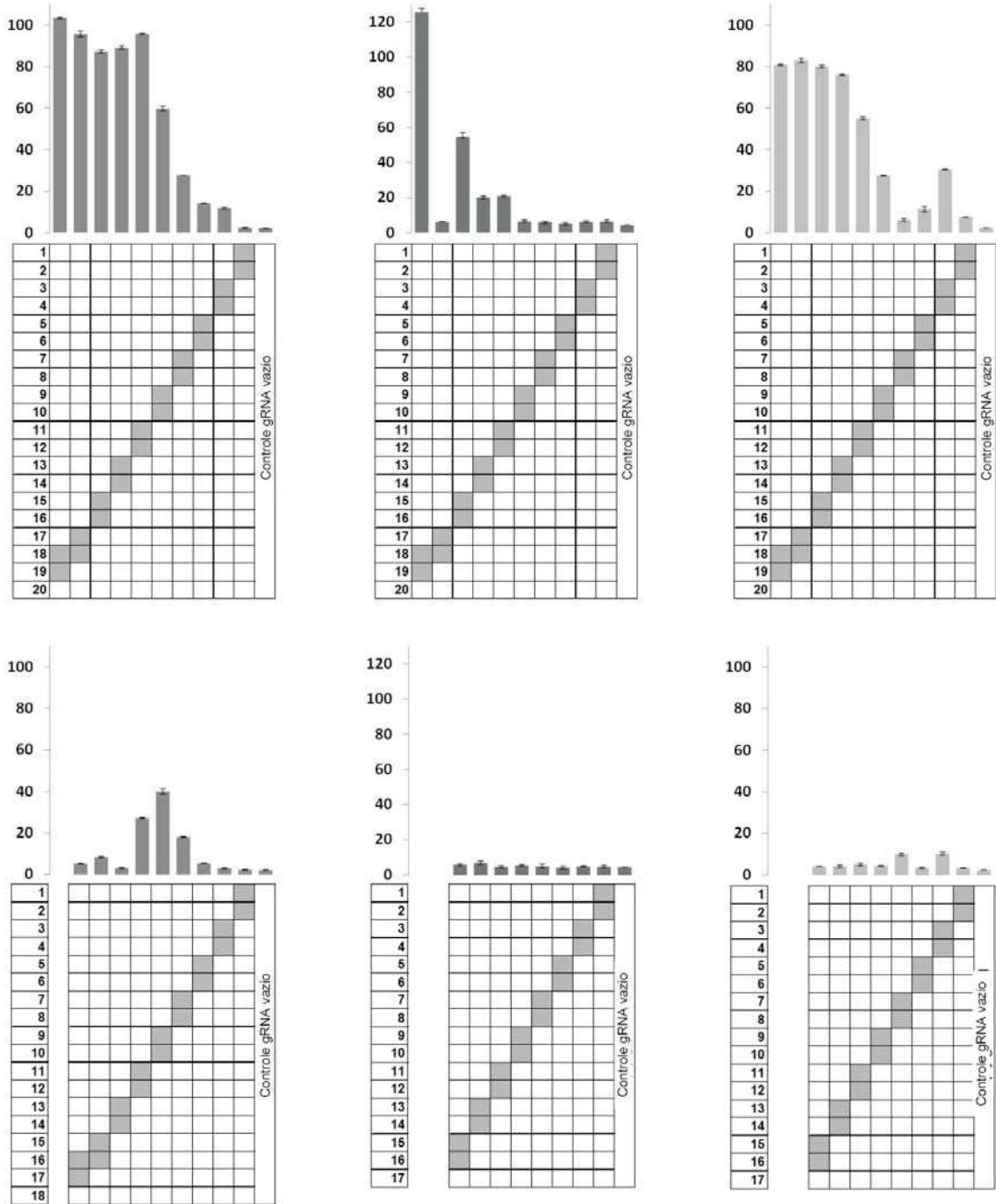
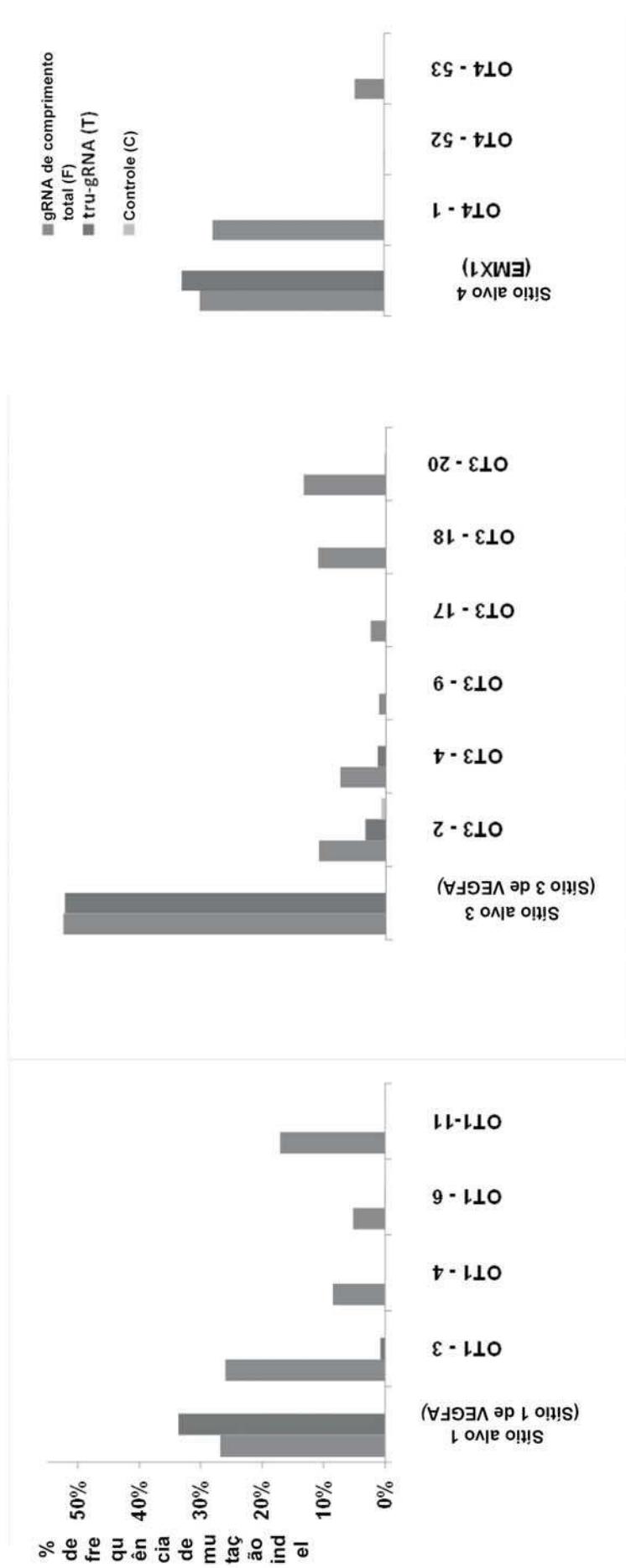
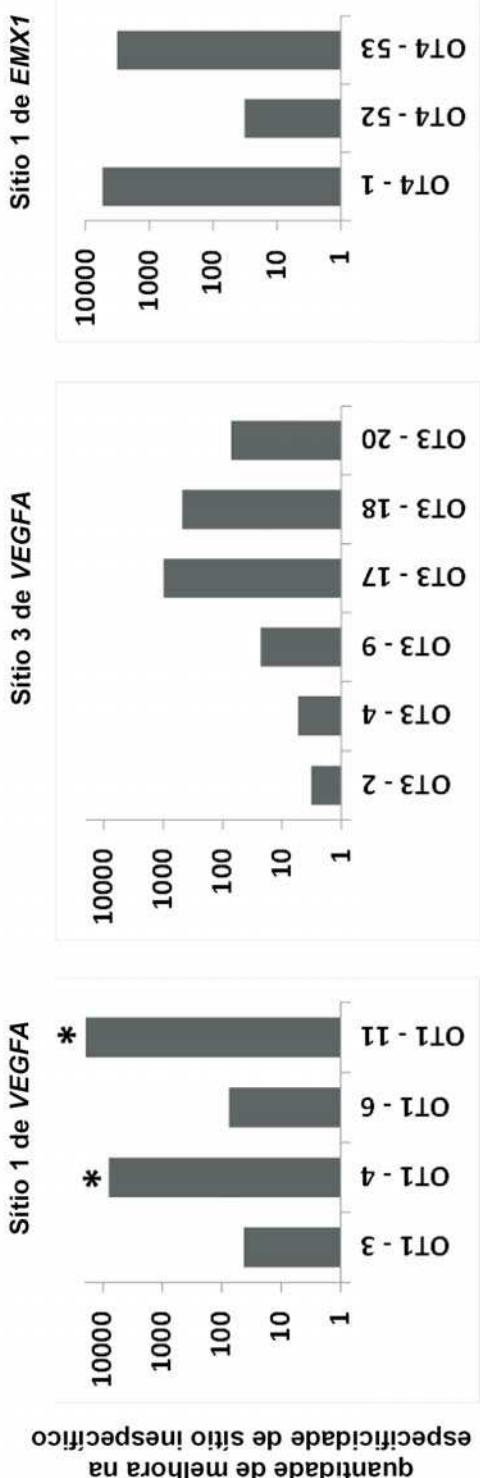
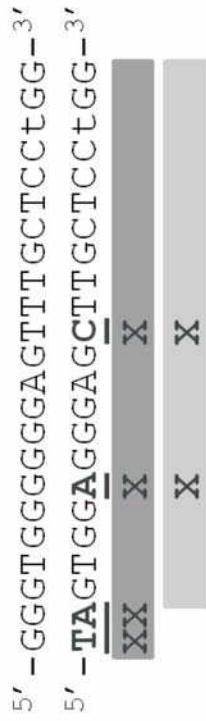
Fig. 5B

Fig. 6A



**FIG. 6B**

Sítio 1 de VEGFA (no alvo):

Sítio inespecífico:

Alvo de gRNA de comprimento total:

Alvo de tru-gRNA:

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

FIG. 6C

Figura 7A

gRNA de comprimento total do Sítio 1 de VEGFA

TAGCTTTGGGAGGTCAAGAAATAGGGGTCCAGAGCAAACCCCCCACCCTTCCAAAGCCC Tipo Selvagem x87

```

TAGCTGTT-----> Δ205
<-----CCTTCCAAAGCCC Δ71
<-----CCAAAGCCC Δ67
<-----CCCCACCCCTTCCAAAGCCC Δ55
TAGCTG-----TTCCAAAGCCC Δ49
TAG-----CCCTTCCAAAGCCC Δ49
TAGCTTTGGGAGGTCAAGA-----Δ47
TAGCTGTT-----CCCCTTCCAAAGCCC Δ43
TAGCTTTGGGAGGTCAAGAAAT-----AGCCC Δ39
TAGCTTTGGGTGG-----CCTTCCAAAGCCC Δ38
TAGCTTTGGGAGGTCAAGAAATAGG-----CAAAGCCC Δ33
TAGCTTTGGGAGGT-----CACCCCTTCCAAAGCCC Δ32
TAGCTTTCTGA-----CCTCCCACCCCTTCCAAAGCCC Δ30
TAGCTTTGGGAGGTCAAGAAATAGGGGTCCAG-----ANCCC Δ28
TAGCTTTGGGAGGTCAAGAAAT-----ACCCCTTCCAAAGCCC Δ26
TAGCTTTGGGAGGTCAAGAAATAGGGGT-----CCCTTCCAAAGCCC Δ22
TAGCTTTGGGAGGTCAAG-----AAACTCCCCCACCCCTTCCAAAGCCC Δ19 x2
TAGCTTTGGGAGGTCAAGAAATAGGGGT-----CCACCCCTTCCAAAGCCC Δ17
TAGCTTTGGGAGGTCAAGAAATAGGGGT-----CCCACCCCTTCCAAAGCCC Δ16
TAGCTTTGGGAGGTCAAGAAATAGGG-----TCCCCCNCCCCCTTCCAAANCCC Δ15
TAGCTTTGGGAGGTCAAG-----AGCAAACCCCCCACCCCTTCCAAAGCCC Δ15
TAGCTTTGGGAGGTCAAGAAAT-----AAACTCCCCCACCCCTTCCAAAGCCC Δ15
NAGCTGNTTGGGAGGNNA-----NNGCAAACCCCCCACCCCTTCCAAANCCC Δ14
TAGCTTTGGGAGGTCAAGAAATAGGGGT-----CCCCCACCCCTTCCAAAGCCC Δ14
TAGCTTTGGGAGGTCAAGAAATAGGGGTCCA-----CCCACCCCTTCCAAAGCCC Δ13
TAGCTTTGGGAGGTCAAGAAAT-----AGCAAACCCCCCACCCCTTCCAAAGCCC Δ12
TAGCTTTGGGAGGTCAAGAAATAGGGGTCCAGGA-----CCCCCACCCCTTCCAAAGCCC Δ9
TAGCTTTGGGAGGTCAAGAAATAGGGGTCC-----AACTCCCCCACCCCTTCCAAAGCCC Δ7
TAGCTTTGGGAGGTCAAGAAATAGGGGTCT-----CACCCCTTCCACCCCTTCCAAAGCCC Δ6
TAGCTTTGGGAGGTCAAGAAATAGGGGTCCAGG-----ACTCCCCCACCCCTTCCAAAGCCC Δ5 x2
TAGCTTTGGGAGGTCAAGAAATAGGGGTCCAGG-----AACTCCCCCACCCCTTCCAAAGCCC Δ4 x2
TAGCTTTGGGAGGTCAAGAAATAGGGGTCC-----AGCAAACCCCCCACCCCTTCCAAAGCCC Δ3
TAGCTTTGGGAGGTCAAGAAATAGGGGTCCAG-----GCAAACCCCCCACCCCTTCCAAAGCCC Δ2 x3

TAGCTTTGGGAGGTCAAGAAATAGGGGTCCAGGAGTGCTAGATCTCATCTAAAGACATTCTGA +191
GAAAAATGTATCTGTTTCTTCAGAAGAAATTACACTTAATAGATATTATGGTAACTAAAGTAAG
GCAGATAATTTGGCCATCAGCTTATATTGTGGGATAATCTCTTTGCTGACCTTGAAAGNTGTG
GCATATTACAACAAGTAGGAAAATTGCAAACCCCCCACCCCTTCCAAAGCCC

```

Figura 7B

gRNA truncado do Sítio1 de VEGFA

TGGGAGGTCA~~GAAATAGGGGTCC~~~~GGAGCAA~~ACTCCCCC~~ACCCC~~TTCCAAGCCCATTCCCTC Tipo Selvagem x85

TGGGAGGTCA~~GAAATAGGGGTCC~~~~CAGGA~~-----> Δ144
 <-----CACCC~~CTT~~CCAAGCCCATTCCCTC Δ112
 <-----CAA~~ACTCCC~~CCC~~ACCC~~TTCCAAGCCCATTCCCTC Δ111
 <-----AAGCCCATTCCCTC Δ80
 <-----TCCCTC Δ80
 <-----TTCCAAGCCCATTCCCTC Δ61
 TGGGAGGTCA-----> Δ61
 <-----CCCAC~~CCC~~TTCCAAGCCCATTCCCTC Δ56
 <-----CCCAC~~CCC~~TTCCAAGCCCATTCCCTC Δ44
 TGGG-----CCC~~TTT~~CCAAGCCCATTCCCTC Δ39
 TGGGAG-----CCCC~~TTT~~CCAAGCCCATTCCCTC Δ36
 TGGGAGGTCA~~GAAA~~-----CACCC~~CC~~TTCCAAGCCCATTCCCTC Δ26
 TGGGAGGTCA~~GAAATAGGGGT~~-----CCC~~CC~~TTCCAAGCCCATTCCCTC Δ20
 TGGGAGGTCA~~GAA~~-----AAACTCCCCC~~ACCC~~TTCCAAGCCCATTCCCTC Δ17
 TGGGAGGTCA~~GAAATAGGGGT~~-----CCACCC~~CC~~TTCCAAGCCCATTCCCTC Δ17
 TGGGAGGTCA~~GAAATAGGGGT~~-----CCCAC~~CC~~TTCCAAGCCCATTCCCTC Δ16
 TGGGAGGTCA~~GAAATAGG~~-----TCCCCC~~ACCC~~TTCCAAGCCCATTCCCTC Δ16
 TGNGAGGTCA~~GAGA~~-----AGCAA~~ACTCCC~~CCC~~ACCC~~TTCCAAGCCCATTCCCTC Δ15
 TGGGAGGTCA~~GAAATAGGGGT~~-----CCCAC~~CC~~TTCCAAGCCCATTCCCTC Δ15
 TGGGAGGTCA~~GAAATAGGGG~~-----TCCCCC~~ACCC~~TTCCAAGCCCATTCCCTC Δ13 x2
 TGGGAGGTCA~~GAAATAGGGGTCC~~~~CAGGA~~-----CACCC~~CC~~TTCCAAGCCCATTCCCTC Δ12
 TGGGAGGTCA~~GAAATAGG~~-----GCAA~~ACTCCC~~CCC~~ACCC~~TTCCAAGCCCATTCCCTC Δ11 x2
 TGGGAGGTCA~~GAAATAGGGGTCC~~~~CAGG~~-----CCCAC~~CC~~TTCCAAGCCCATTCCCTC Δ9
 TGGGAGGTCA~~GAAATAGGGGTCC~~~~CAGG~~-----NCCCC~~ACCC~~TTCCAAGCCCATTCCCTC Δ8
 TGGGAGGTCA~~GAAATAGGGGTCC~~~~CAG~~-----CTGGCCCC~~ACCC~~TTCCAAGCCCATTCCCTC Δ7
 TGGGAGGTCA~~GAAATAGGGGT~~-----CAA~~ACTCCC~~CCC~~ACCC~~TTCCAAGCCCATTCCCTC Δ6 x5
 TGGGAGGTCA~~NAAATAGGGGTCC~~~~CAG~~-----ACTCCCCC~~ACCC~~TTCCAAGCCCATTCCCTC Δ6
 TGGGAGGTCA~~GAAATAGGGGTCC~~~~CAGG~~-----ACTCCCCC~~ACCC~~TTCCAAGCCCATTCCCTC Δ5
 TGGGAGGTCA~~GAAATAGGGGTCC~~-----AGCAA~~ACTCCC~~CCC~~ACCC~~TTNNNNNNCCATTNNNTC Δ3
 TGNGAGGTCA~~GAAATAGGGGTCC~~~~CAGGA~~-----AAACTCCCCC~~ACCC~~TTCCAAGCCCATTCCCTC Δ2

TGGGAGGTCA-----CCCC~~TTT~~CCAAGCCC~~ACCC~~TTCCAAGCC Δ9 (+14 Δ23)
 TGGGAGGTCA~~GAAATAGGGGTCC~~~~CAGGAA~~AGCAA~~ACTCCC~~CCC~~ACCC~~TTCCAAGCCCATTCCC +2

Figura 7C

gRNA de comprimento total do Sítio3 de VEGFA

Figura 7D

gRNA truncado do Sítio3 de VEGFA

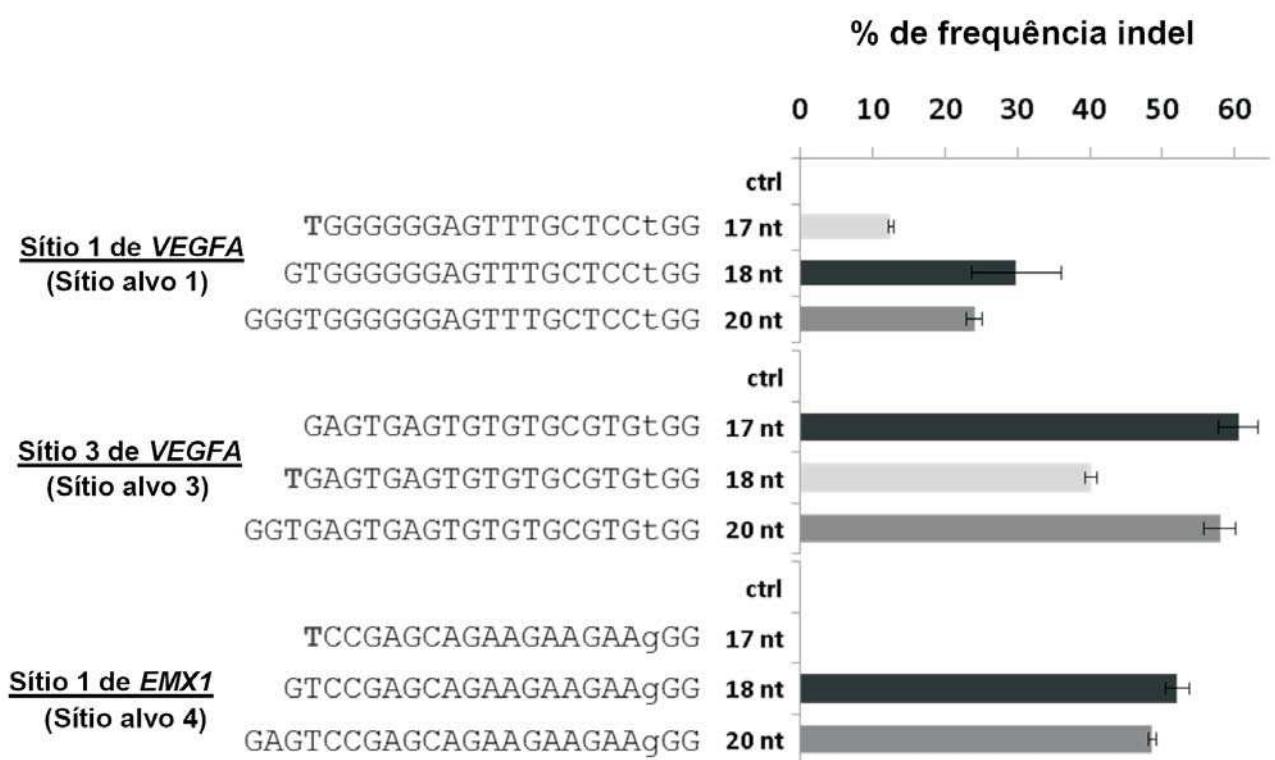
FIG. 7E

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

AGACAGGACATTCTGACACCCCAGGAGCAAACCTCCCTCCATCCCACAAATCCGTCTTAGATGTGCA Tipo Selvagem x18
 <----- AACTCCCTCCATCCCACAAATCCGTCTTAGATGTGCA Δ53
 AGACAGGACATTCTGACACC----- CCATCCCACAAATCCGTCTTAGATGTGCA Δ17
 AGACAGGACATTCTGACACCCCAGGA----- CCCACAAATCCGTCTTAGATGTGCA Δ15
 AGACAGGACATTCTGACACCCCAG----- GCACAAACTCCCTCCATCCCACAAATCCGTCTTAGATGTGCA Δ2
 AGACAGGACATTCTGACACCCCAGGAGGGCAAACCTCCCTCCATCCCACAAATCCGTCTTAGATGTGCA +2

OT1-6

GAGAGAGGCTCCCATCACGGGGGAGGGAGTTGCTCCTGGGAACCTGTGATCCCCACAGGAAACA Tipo Selvagem x87
 GAGAGAGGCTCCCATCACGGGGG----- AGGGAACA Δ35 x3
 GAGAGAGGCTCCCATCACGGGGG----- GGGGAACCTGTGATCCCCACAGGAAACA Δ14 x1

OT1-11

TGGACTCTACCCACTGAATGCCAGGAGCAAACCTCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Tipo Selvagem x27
 TGGACTCTACCCACTG----- AATGTCTC Δ43
 TGGACTCTA----- CCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ25
 TGGACTCTAC----- CCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ24
 TGGACTCTACCCACTGAATG----- CCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ15
 TGGACTCTACCCACTGAATG----- CCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ15
 TGGACTCTACCCACTGAAT----- GCACAAACTCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ7
 TGGACTCTACCCACTGAATGCCAGGA----- TTCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ6
 TGGACTCTACCCACTGAATGCCCTGG----- CATCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ6
 TGGACTCTACCCACTGAATGCCAGG----- AAACCTCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC Δ3

TGGACTCTACCCACTGAATGCCAGGATGGAAGATAATTTTCCATAGACCAGGGTGGGGAATGGTTCGGGAT
 GATTCAAGCACATCACATTATTGTGCACTTATTCTATTACTATTATATTGTAAATGTATACTAAAAATAATTA
 TACAACTCACCATAATGTAGAACCGAGTGGGAGCCGCAAACCTCCCTCCCCGAGTTGTGACAGCAAAAATGTCTC
 A +158

TGGACTCTACCCACTGAATGCCAGGCAAACCTCCCTCCCCGAGTTGTGACAGCAAAAATGTCTGGCCTAATGTC
 TGGCAAAAATGTCTCAAGACATTGCCAAATGTCCCCT +23 (Δ2 +25)

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

ACCCACCTCCCTATCCTCAAAACTTGGCCAGAGGCGGGTGGAGGGGCCCTAGGAGCGCCTTGGTG Tipo
Selvagem x30

```

ACCCAC-----GCCCTTGGTG Δ51
ACCCACCTCC-----TATCCTAGGAGCGCCTTGGTG Δ36
ACCCACCTCCCTATCCTCAAAACTT-----CCCTTGGTG Δ33
ACCCACCTCCCTATCCTCAAAACTTGGCCAGAG-----TAGGAGCGCCTTGGTG Δ18
ACCCACCTCCCTATCCTCAAAACTT-----GGGGTGGAGGGGCCCTAGGAGCGCCTTGGTG Δ10
ACCCACCTCCCTATCCTCAAAACTT-----GGCGGGTTGGAGGGGCCCTAGGAGCGCCTTGGTG Δ7
ACCCACCTCCCTATCCTCAAAACTTGGCCAGA-----GTGGAGGGGCCCTAGGAGCGCCTTGGTG Δ6
ACCCACCTCCCTATCCTCAAAACTT-----CTGGCAGTCTGTCAGTGCCTTACACTTCTACAAGGGGGCTCTCCCTGCATTCTGA +21 (Δ40,  
+61)
ACCCACCTCCCTATCCTCAAAACTTGGCCAGAGGGCGCCTCCCCCAGGAAGTGCTCCGGCCAGCCCAGGGTAACAC
CGCTAGCCCTGCCCTCTGGGACCATAGCCCAGGGACCCAGACTCTGGCCACGCTCATCCCACCGCGGGGTGG
AGGGGCCCTAGGAGCGCCTTGGTG +108

```

OT2-15

TGACTGCGGTGCCAACATGTGGCAGATGCCAGAGGCAGGGGTGTGGGGGTACTTGTGGCGTT Tipo
Selvagem X71

```

TGACTGTC-----GGTGTGGGGGTACTTGTGGCGTT Δ33
TGACTGCGGTGCCAACATGTGGCAGATGCCAGAG-----GGGCGTT Δ23
TGACTGCGGTGCCAACATGTGGCAGA-----TGTGGGGGTACTTGTGGCGTT Δ15
TGACTGCGGTGCCAACATGT-----GGCGGGGTGTGGGGGTACTTGTGGCGTT Δ14
TGACTGCGGTGCCAACATGTGGCAGATG-----GAGGCAGGGGTGTGGGGGTACTTGTGGCGTT Δ4
TGACTGCGGTGCCAACATGTGGCAGATGCCAGA-----GGGTGTGGGGGTACTTGTGGCGTT Δ4
TGACTGCGGTGCCAACATGTGGCAGATGCCAGAG-----GGGTGTGGGGGTACTTGTGGCGTT Δ2
TGACTGCGGTGCCAACATGTGGCAGATGCCAGAGTTGCAGGGGTGTGGGGGTACTTGTGGCGTT +2
TGACTGCGGTGCCAACATGTGGCAGATGCCAGAGCTGCAGGGGTGTGGGGGTACTTGTGGCGTT +2

```

OT2-24

ACAAGATGACTATGCCCTGGGCCCATCCTCCCTCCCCACCCACCCGCCTCAGGCTTGAAGA Tipo
Selvagem x8

```

ACAAGA-----> Δ121
<-----CTCAGGCTTGAAGA Δ82
ACAAGATG-----> Δ80
<-----Δ79
ACAAGATGACTATGCCCTGGGCCCATCCTCC-----> Δ42
ACAAGATGACTATGCCCTGGG-----CTTGAAGA Δ34
ACAAGATGACTATGCCCTGGGCCCATCCTCCCTCCC-----CAGGCTTGAAGA Δ14
ACAAGATGACTATGCCCTGGGCCCATCCTCCCTCC-----CCTCAGGCTTGAAGA Δ12
ACAAGATGACTATGCCCTGGGCCCATCCTCCCTCCCC-----GCAGGCTTGAAGA Δ11
ACAAGATGACTATGCCCTGGGCCCATCCTCCCTCCCCAC-----CTTGAAGA Δ11
ACAAGATGACTATGCCCTGGGCCCATCCTCCCTCCC-----GCTCAGGCTTGAAGA Δ11

```

ACAAGATGACTATGCCCTGGGCCGCCTCAAGTGATCCAGCTGCCCTGGCTCCAAAAGTGCTAGCAGTACA
GATGTGAGCCTCCATGCCTGGCTATTGCAACATCCCCTCTGTGAAGCAGGGTTCTGCAGTGACAGCAAGA
AGAGCACAGGGCAAAAAACTTGTCCCTTAGAAAGGATCTACCTTTAGGCTGAGAATGGCA +76 (Δ43
+119)

Fig. 8C

Alvo 3 (Sítio 3 de VEGFA) :

OT3-2

OT3-9

```

GTGTTGGGATGCGGGAGTGGGTGAGTGAGTGCCTGCCGGTGGCGATGCAAGCGTGTGCGAATGCGTG x173
GTGTTGGGATGCGGGA-----> Δ80
GTGTTGGGATGC-----GCGTG Δ50
GTGTTGGGATGCGGGAGTGGGTGAGTGA-----GTGGCGATGCAAGCGTGTGCGAATGCGTG Δ10

```

GTGTTGGGATGCAGGGAGTGGGTGAGTGAGTGCAAGTGCAGGGTGGCGATGCAAGCTGTGCGAATGCGTG +2

OT3-18

TTTCAAAGACAGTAGATCTAAATGTCCTCACGCACACACTACCCACACATAAAAGGTGGTAAC TG Tipo
Selvagem x27

```

TTTCAAAGACAGTAGATCT----- TAAAAGTGGTAACTG Δ32
TTTCAAAGACAGTAGATCTAAATGT----- CATAAAAGTGGTAACTG Δ23
TTTCAAAGACAGTAGATCTAAATGTCC-----CACATAAAAGTGGTAACTG Δ18 x4
TTTCAAAGACAGTAGATCTAAATGTCC-----CCACACATAAAAGTGGTAACTG Δ15
TTTCAAAGACAGTAGATCTAAATGT-----CTCACCCACACATAAAAGTGGTAACTG Δ12
TTTCAAAGACAGTAGATCTAAATGTCCCTCA-----CACACACTCACCCACACATAAAAGTGGTAACTG Δ2

```

TTTCAAAGACAGTAGATCTAAATGTCCTCACAGGCTGGAGTACAGTGGCATGATATCAGCTCACTGCAATCTCG
GGCTCCGGGTTCAAGCCATGCACACACTCACCCACACATAAAAGGTGGTAAC +63

Alvo 4 (*EMX1*):

OT4-1

ACCTGTACATCTGCACAAGATTGCCTTACTCCATGCCTTCTTCTGCTCTAACTCTGACAATC Tipos
Selvagem x20

```

----- ATC Δ64
ACCTGTACATCTGCACAAGATTGCCTTACTCC----- ACAATC Δ28
ACCTGTACATCTGCACAAGATTGCCTTACTCCAT----- ACTCTGACAATC Δ20
ACCTGTACATCTGCACAAGATTGC----- CTTCTGCTCTAACCTGACAATC Δ20
ACCTGTACATCTGCACAAGATTGCCTTACTCCATGCCTTCT----- CAATC Δ19
ACCTGTACATCTGCACAAGATTGCCTTACTCCA----- TGCTCTAACCTGACAATC Δ14
TCCTGTACATCTGCACAAGATTGCCTTACTCC----- CTTCTCTGCTCTAACCTGACAATC Δ8

```

Fig. 9A**Alvo 1 (Sítio 1 de VEGFA) :****OT1-3**

TCAGACAGGACATTCTGACACCCCAGGAGCAAACCTCCATCCCACAAATCCGTCTTAGATGTG Tipo Selvagem
x41

TCAGACAGGACATT-----CAAACCTCCATCCCACAAATCCGTCTTAGATGTG Δ15

TCAGACAGGACATTCT-----GAGCAAACCTCCATCCCACAAATCCGTCTTAGATGTG Δ10

TCAGACAGGACATTCTGAC-----GCAAACCTCCATCCCACAAATCCGTCTTAGATGTG Δ9

TCAGACAGGACATCTGACAC-----GCAAACCTCCATCCCACAAATCCGTCTTAGATGTG Δ7

TCAGACAGGACATTCTGACACCCCAGG-----GCAAACCTCCATCCCACAAATCCGTCTTAGATGTG Δ2 x6

TCAGACAGGACATTCTGACACCCCAGGA~~T~~GTCCTCCTCCATCCCACAAATCCGTCTTAGATGTG +2

TCAGNCAGGACATTNNNGNCACCCCAGGAAACNNNGAGTTTCGNTNCNNNGANNGTCAGACCCAGNAGCAAACCTCCATC
CCACAAATCCGTCTTAGATGTG +38

TCAGACAGGACATTNTGACACCCCAGGAGTNTGCACNTCAGTTTCTTANTATGTNGNNNNGGGCANGNACAAANNTTN
GCAAACCTCCATCCCACAAATCCGTCTTAGATGTG +54

TCAGACAGGACATTCTGACACCCCAGGATGTTGAGTCAGAGTCTCTTTGTACCCAGGGCTGGAGTGCAGTGGAA
ACCTGTGCCCTTGATATCCTCTTGAGGTAAAGAGTCATCATGGATCANCNNATAAGCAAACCTCCATCC
+116

TCAGACAGGACATTCTGACACCCCAGGATAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACG
CCTGGTATCTTATAGCCTGTCGGGTTTCGCCACCTCTAGGGAGTTCTGAGCTCTGACGCCGGCGTGTAGGAGATAGC
AGGCCGTTAATGACCATCCCAGCGAATTCTCACTGTGAGATGAGGAAGTGAGCTCAGGGAGGCTGAGTGTCCCAGGCCT
GTTGCCAGATGAGGCCACGCTGAGACTGTGCAAACCTCCATCCCACAAATCCGTCTTAGATGTG +247

OT1-6

TGGAGAGAGGCTCCATCACGGGGAGGGAGTTGCTCCTGGGAACCTGTGATCCCCACAGGAAAC Tipo Selvagem
x88

TGGAGAGAGGCTCCATCACGGGGAGGGAGTTG-----CCTGTGATCCCCACAGGAAAC Δ11

TGGAGAGAGGCTCCATCACGGGGAGGGAGTT-----GGGAACCTGTGATCCCCACAGGAAAC Δ7

OT1-11

AGCATCGCTGGACTCTACCCACTGAATGCCAGGAGCAAACCTCCCTCCCCGAGTTGTGACAGCAAAC Tipo Selvagem
x84

AGCATCGCTGGACTCTACCCACTGAATGCCAGGA-----CCCGAGTTGTGACAGCAAAC Δ14

AGCATCGCTGGACTCTACCCACTGA-----GCAAACCTCCCTCCCCGAGTTGTGACAGCAAAC Δ9

AGCATCGCTGGACTCTACCCACTGAATGCCAG-----GCAAACCTCCCTCCCCGAGTTGTGACAGCAAAC Δ2

AGCATCGCTGGACTCTACCCACTGAATGCCAGGAGTCAGACGATTGAATGTATCAACTTGGCACATTGCCATCAACTGGT
GAGTGTCTAAAAATATCCATTGCTGTGATCAGTAATGCCACAGGGTGACCATTAAAGGACAGAGTCCATGTTTATCCATCC
TTAGCAAACCTCCCTCCCCGAGTTGTGACAGCAAAC +133

AGCATCGCTGGACTCTACCCACTGAATGCCAGAGCCCTTCTCCCTCTTCCAGAGGGCTGCCAGAGATCAGGTT
GGAGGGCTCTTGTCTTATGCCATTCTCCCCCAGGCACCTGGAGGAGGGCACTGTTGAGTGTGCAAGTCTTCTC
TGTTACTGTTGGCAAACCTCCCTCCCCGAGTTGTGACAGCAAAC +142

AGCATCGCTGGACTCTACCCACTGAATGCCAGGATGTTTGTGCGACGGANTCTCACTCTGTCGNCCGGGCTGGAGTGC
NNGGCACANTCCTCANCTGACTGCNATGTCCGCCTCCCGGATTCAAGTGTGATTCTCTGCCCCAGCCTCCCGAGTAGCTGGG
ATTATAGGTGCCTGCCACCATGCCCTGGCTAATTTTTTTTTAAATGGAGTCTCACTCTGTTGCCCGAGTTGTGA
CAGCAAAC +186

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

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

CACCTCCCTATCCTCAAAACTTGGCCAGAGACTTANACCTAACACTATGAGACTGCTACNAGAGAACATCANAAAAACTTTCAGGACATTCTCTGGNGGGTGGAGGGGCCCTANGAGCGCCTGNNGGA +74

CACCTCCCTATCCTCAAAACTTGGCCAGAGTCTACAGATTATAAAATATTACAGTTAATCATGACACATATTGTTTATTTCAAATATTTTCTAGTTAAACCCACCATTATATAACCAATTATATTGATATTATTAAAATTTTGTTATTAAACACCCACCAAATCATTTACAGCGGGTGGAGGGGCCCTAGGAGCGCCTTGGTGGGA +150

OT2-15

GTCGGTCCCCACATGTGGCAGATGCCAGAGGCAGGGTGTGGGGGTACTTGTGGCGTTTGGG Tipo Selvagem x79
 GTCGGTCCCCCAC-----GCAGGGTGTGGGGGTACTTGTGGCGTTTGGG Δ19
 GTCGGTCCCCACATGTGGC-----GCAGGGTGTGGGGGTACTTGTGGCGTTTGGG Δ12
 GTCGGTCCCCACATGTGGCAGATGCCAGA-----GGTGTGGGGGTACTTGTGGCGTTTGGG Δ5
 GTCGGTCCCCACATGTGGCAGATGCCA-----GGGTGTGGGGGTACTTGTGGCGTTTGGG Δ5
 GTCGGTCCCCACATGTGGCAGATGCCAGA-----GGGTGTGGGGGTACTTGTGGCGTTTGGG Δ4
 GTCGGTCCCCACATGTGGCAGATGCC-----GCAGGGTGTGGGGGTACTTGTGGCGTTTGGG Δ4
 GTCGGTCCCCACATGTGGCAGATGCC---AGGCAGGGTGTGGGGGTACTTGTGGCGTTTGGG Δ2 x2

OT2-24

GTCCCTCTGGCCCCATCCTCCCCTCCCCACCCACCCGCCCTCAGGCTTGAAGAGGAAGAGAGCA
 GTCCCTCTGGCCCC-----ANAGNANNANNANN Δ36
 GTCCCTCTGGCCCCATCCTCCCCTCCCC-----TGAAGAGGAAGAGAGCG Δ19
 GTCCCTCTNNNCNNCNT-----CCNCNCCTCAGGCTTGAAGAGGAAGAGAGCG Δ17
 GTCCCTCTGGCCCCATCCTCCCCTC-----CCTCAGGCTTGAAGAGGAAGAGAGCG Δ13
 GTCCCTCTGGCCCCATCCTCCCCTCCCC-----CTCAGGCTTGAAGAGGAAGAGAGCG Δ10
 GTCCCTNNNNNNNCTCCTCCCC-----NNCCNCNCTCAGGCTTGAAGAGGAAGAGAGCG Δ8
 GTCCCTCTGGCCCCATCCTCCCCTCCCCACCCA-----CTCAGGCTTGAAGAGGAAGAGAGCG Δ6
 GTCCCTCTGGCCCCATCCTCCCCTCCCCACCCA-----CCTCAGGCTTGAAGAGGAAGAGAGCG Δ5

GTCCCTCTGGCCCCATCCTCCCCTCCCCACCCACCCGCCATCGTACGTGTCCCTACTAAGCTGCAATTGGCAT
 CTTCAGCTAAGTCGAAGTTCGACCTCAGGCTTGAAGAGGAAGAGAGCG +58

Fig. 9C

Alvo 3 (Sítio 3 de VEGFA):

OT3-2*

OT3-9

TGGAGGTGTTGGGATGC~~GGG~~GAGTGGGTGAGTGA~~GTGCGTGC~~~~GGGTGGC~~GATGCAAGCGTGTGCGAAT Tipo Selvagem x101
 TGGAGGTGTTGGGATGC~~GGG~~GAGTGG-----GTGCGTGC~~GGGTGGC~~GATGCAAGCGTGTGCGAAT Δ8
 TGGAGGTGTTGGGATGC~~GGG~~GAGTGGGTGA-----GTGCGGGTGGC~~GATGCAAGCGTGTGCGAAT~~ Δ8

OT3-18

```

CAAAGACAGTAGATCTTAAATGTCCTCACGCACACACTCACCCACACATAAAAGGTGGTAAGTGTGT  Tipo
Selvagem x64
CAAAGACAGTAGATCTTAA-----GCACACACTCACCCACACATAAAAGGTGGTAAGTGTGT Δ10
CAAAGACAGTAGATCTTAAATGTCCTCACG-----TCACCCACACATAAAAGGTGGTAAGTGTGT Δ7
CAAAGACAGTAGATCTTAAATGTC-----GCACACACTCACCCACACATAAAAGGTGGTAAGTGTGT Δ5
CAAAGACAGTAGATCTTAAATGTCCTCACGCCGACNATTGCATGAAGAATCTGCTTAGGGTTAGGCCTTTGCGC
TGCTTCGCGATGTACGGGCCAGATAGCACACACTCACCCACACATAAAAGGTGGTAAGTGTGT +71

```

CAAAGACAGTAGATCTTAAATGTCCTCAC GCAAATT TTTGGTTCATGATATGGCTGGCGTGATGCTTT
CATTTGTA AAAATTGCTGTTCTTTGACAATTTAAGTGACTGTTTCATTGACTACAAGTTGAAAATAAAATAA
TTAAGAAAAAAATTCCAATGACTGTGCTGTGGTGGGCACACACTCACCCACACATAAAAGGTGGTAACTGTGT
+157

CAAAGACAGTAGATCTTAAATGTCCTCACGTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAAAC
GATTTATCACCTCANAAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAACCTACTTGGCTCTTGC
CCATATGATAAAGTCCGTGGGCACTTCATTGAGGGTGAATCAAATCCGGACAACCTGGATGTCGAGCACAC
ACTCACCCACACATAAAAGGTGGTAACGTGT +190
CAAAGACAGTAGATCTTAAATGTCCTCACG-----
TTATTTAGAGACAGAGTCTCACTCTGTTGCCAGGCTGGGGTGCAGTGGTACGAACCTGGCTCACTGCAACCTCC
GTCTCCTGGGCTCAAGTGATTATCCTGCCTCAGCCTCCCAGTAGCTGGGATTACAGGGTCCCCACCACACACCC
GGCTAATTTTGATTTCAAGTAGAGCTGGGGTTCACCATGTTGCCAGCCTGTTCTGGCACACACTCACCCA
CACATAAAAGGTGGTAACGTGT +211 (Δ 16 +227)

Alvo 4 (EMX1):

OT4-1

GATTGCCTTACTCCATGCCCTTCTTCTGCTCTAACCTGACAAATCTGTCTGCCATGCCATAA Tipo
Selvagem x74
GATTGCCTTACTC-----CTTCTTCTGCTCTAACCTGACAAATCTGTCTGCCATGCCATAA Δ9
GATTGCCTTACTCCATGCC-----TTCTGCTCTAACCTGACAAATCTGTCTGCCATGCCATAA Δ6 x2
GATTGCCTTACTCCATGC-----TCTTCTGCTCTAACCTGACAAATCTGTCTGCCATGCCATAA Δ6
GATTGCCTTACTCCATGCC-----TTCTTCTGCTCTAACCTGACAAATCTGTCTGCCATGCCATAA Δ3 x3