

SISTEMAS DE EDIÇÃO GENÔMICA – ZFN, TALENs E CRISPR/Cas9

Matheus Rodrigues Ferreira

Biomédico – Faculdades Integradas de Três Lagoas – FITL/AEMS

Erlí de Souza Bento

Bióloga – UFMS; Mestre em Agronomia (Genética e Melhoramento de Plantas) – UNESP
Docente das Faculdades Integradas de Três Lagoas - FITL/AEMS.

Catarina Akiko Miyamoto

Farmacêutica-Bioquímica – USP; Doutora em Ciências Biológicas (Bioquímica) – USP;
Pós-doutorado – *Weill Medical College of Cornell University*;
Docente das Faculdades Integradas de Três Lagoas – FITL/AEMS

RESUMO

Durante o processo de replicação do DNA, pode ocorrer introdução de pares de bases com pareamentos errôneos e surgimento de mutações gênicas. Além disso, fatores externos podem ocasionar clivagens do DNA genômico. As células possuem sistemas de reparos para tais eventos. Entre os mesmos incluem-se reparações por homologia (HDR) e por não homologia (NHEJ) quando ocorrem clivagens de DNA-dupla fita. Com o domínio destes mecanismos, sistemas artificiais de reparo têm sido desenvolvidos para a edição genômica. Para tal, deve-se realizar segmentação precisa de uma sequência específica de DNA e criação de um sítio específico de dupla fita para nucleases. Dentre os sistemas artificiais de edição genômica, incluem-se ZFN, TALENs e CRISPR/Cas9. Os dois primeiros utilizam proteínas quimeras com domínios de ligação ao DNA fusionadas a nucleases sítio-específicas. Sistema ZFN – domínio dedo de zinco/ nuclease específica, TALENs – ativador de transcrição efetor / nuclease específica. O sistema CRISPR/Cas9 utiliza RNAs guias para direcionar a enzima Cas9 a sítios específicos de interesse.

Palavras-Chave: reparo de DNA; reparações por homologia; HDR; reparações por não homologia; NHEJ.

1 INTRODUÇÃO

O metabolismo do DNA compreende os processos que mantêm cópias fiéis de DNA (replicação) e os que afetam a estrutura inerente da informação (reparo e recombinação) (NELSON; COX, 2008).

Durante o processo de replicação pode ocorrer introdução de pares de bases com pareamentos errôneos. Além disso, vários fatores externos podem induzir mutações gênicas ou cromossômicas. Esses processos podem gerar clivagens de DNA-dupla fita (DSB – *Double-Strand Break*), que são potencialmente letais às células (NELSON; COX, 2008). Estas, por sua vez, possuem sistemas de reparos para tais situações; dentre os mesmos incluem-se reparações por homologia (HDR – *homology-directed repair*) e por extremidades de junções não

homólogas (NHEJ – *non homology usend-joining*) (CHANDRASEGARAN; CARROL, 2016; MAEDER; GERSBACH, 2016; TAKATA et al., 1998).

Estes sistemas de reparo podem ser utilizados em processos de edição genômica para introdução de sequências desejáveis ou eliminação das indesejáveis. HDR, para a produção de transgênicos ou adição homóloga de um DNA doador tipo selvagem (correção de genes), enquanto NHEJ pode eliminar possíveis sequências indesejáveis (genes *knockout*) (CHANDRASEGARAN; CARROL, 2016; LAUFER; SINGH, 2015; GILLES; AVEROF, 2014).

Os objetivos deste trabalho são descrever os sistemas de edição acima citados, assim como mostrar as vantagens e desvantagens de cada um.

2 METODOLOGIA

A metodologia empregada para a realização deste trabalho foi levantamento bibliográfico em artigos científicos nacionais e internacionais indexados em plataformas de pesquisas, como *Lilacs*, *Pubmed* e *Scielo*. Dentre as palavras-chaves utilizadas, incluíram-se CRISPR/Cas9, edição genômica, ZFN e TALEN.

A compilação dos dados preconizou estudos recentes, no entanto não se excluiu publicações antigas com material relevante.

3 SISTEMA DE REPARO DE DNA

Os sistemas de reparo de DNA podem corrigir eventuais pareamentos errôneos e mutações gênicas que ocorrem durante a replicação, e aberrações cromossômicas e clivagens de DNA advindas por fatores externos (NELSON; COX, 2008). Uma das principais vias de reparação, HDR ou NHEJ, podem reparar as últimas (MAEDER; GERSBACH, 2016).

HDR se baseia na recombinação homóloga das cadeias simples do final clivado do DNA molde e uma sequência dupla fita contendo a(s) mudança(s) desejada(s) e subsequente reparo da ruptura, de maneira molde dependente. A via NHEJ liga duas terminações de DSBs por um processo independente de homologia de sequência de DNA, e pode produzir várias composições de nucleotídeos (TAKATA et al., 1998). Devido a esta característica, esta via é propensa a erros e

muitas vezes resultam em inserções ou deleções (*InDelmut*) no sítio da ruptura (MAEDER; GERSBACH, 2016).

4 EDIÇÃO GENÔMICA

Edição genômica é uma ferramenta da biologia molecular utilizada para a realização de mudanças permanentes na sequência do DNA genômico. Para tal, deve-se realizar segmentação precisa de uma sequência específica de DNA e criação de um sítio específico de dupla fita para nucleases (AGROTIS e KETTLER, 2015).

Recentemente, tem sido descrito o desenvolvimento de metodologias para estas finalidades (AGROTIS; KETTLER, 2015). Dentre as mesmas, incluem-se nuclease dedo de zinco (ZFN – *zinc finger nuclease*) (GILLES; AVEROF, 2014), nucleases baseadas como ativadores de transcrição (TALENs – *transcription activator-like effector-based nucleases*) e o sistema de curtas repetições palindrômicas regularmente interespaçadas/genes associados a CRISPR (CRISPR/Cas9 – *clustered regularly interspaced short palindromic repeats/CRISPR-associated nuclease 9*) (BARRANGOU; MARRAFFINI, 2014).

4.1 Sistema Nuclease Dedo de Zinco (ZFN)

Os ZFNs são compostos por uma proteína quimérica domínio dedo de zinco/nuclease. A interação proteína-DNA pelo domínio ZF possibilita a clivagem do DNA em um sítio específico adjacente ao do de ligação (LAUREANO, 2015; MINCZUK et al., 2008).

O domínio ZF, constituído por cerca de trinta resíduos de aminoácido sem configuração $\beta\beta\alpha$ conservada, contém o motivo Cys2-His2, um dos mais comuns de ligação ao DNA encontrado em eucariotos (CHANDRASENGARAN; CARROL, 2015; GAJ et al., 2013). Esta estrutura é estabilizada por um íon Zn^{2+} , e reconhece 3-4 pares de bases (pb) do sulco maior da dupla hélice, por inserção da α -hélice, com níveis variáveis de seletividade (GAJ et al., 2013). A estrutura cristalina sugere que os aminoácidos das posições -1, +1,+2, +3, +5 e +6 da α -hélice podem ser alterados enquanto os demais constituem o esqueleto consenso que dá

especificidade à sequência (CHANDRASENGARAN; CARROL, 2015).

Matrizes não naturais com mais de três domínios ZF tornam possível as interações dos mesmos com sequências específicas de 9-18 pbs de DNA. Uma vez que 18 pbs podem conferir especificidade dentro de 68 bilhões de pbs do DNA; este método foi o primeiro desenvolvido para reconhecer sequências alvo específicas do genoma humano (BEERLI et al., 2000).

4.2 Nucleases Baseadas Como Ativadores de Transcrição (TALENs)

TALENs correspondem à fusão de um ativador de transcrição efetor e uma nuclease específica. As mesmas correspondem a domínios de repetições de 33-35 aminoácidos altamente conservados, com exceção dos resíduos das posições 12 e 13. Estes são os responsáveis pela especificidade de interação da proteína com um par de bases de DNA, portanto são hipervariáveis, e são denominados repetições variáveis de dois resíduos (RVDs) (LAUREANO, 2015; MA; LIU, 2015; DENG et al., 2012; MAK et al., 2012).

Os RVDs frequentemente encontrados His/Asp (HD), Asn/Gly (NG) e Asn/Ile (NI) reconhecem citosina (C), timina (T) e adenina (A), respectivamente. Podem-se construir proteínas TALs com diferentes RVDs para especificidade de diferentes alvos de DNA. No entanto, não se sabe como as repetições TAL reconhecem sequências específicas de DNA (CERMAK et al., 2011).

Um dedo de zinco com 30 aminoácidos reconhece três pares de bases de DNA e TALENs com 34 aminoácidos especificam somente um único par de bases. O grande tamanho e a natureza repetitiva das matrizes ZFs e TALEs representam obstáculos para a entrega destas proteínas *in vivo*. Estes tamanhos de proteínas podem proibir a entrega de ambos por um vetor viral que tem capacidade limitada de empacotamento (MAEDER; GERSBACH, 2016).

4.3 CRISPR/Cas9

O sistema CRISPR/Cas corresponde a um mecanismo de defesa adaptativo contra fagos e consiste de uma nuclease guiada por RNA (SCHIML et al., 2014). Foi inicialmente descoberto em *E. coli* (TAKASE et al., 1987), mas atualmente sabe-se

que está presente em bactérias e arqueas. O mesmo detecta e degrada DNAs exógenos de plasmídeos e de bacteriófagos (FINERAN; CHARPENTIER, 2012).

Dentre as nucleases guiada por RNA, a mais recente e utilizada é a Cas9, proveniente de *Streptococcus pyogenes* (SCHIML et al., 2014).

Há três tipos de sistemas CRISPR/Cas identificados até o momento; o de tipo II é o mais extensivamente estudado (SAVIC; SCHWANK, 2015).

O mecanismo molecular do sistema CRISPR/Cas9 é simples e eficiente para indução de DSB sequência específica. Uma curta sequência de CRISPR RNA (crRNA) liga-se diretamente a um sítio de reconhecimento de 20 nucleotídeos sobre o DNA, o protoespaçador. A sequência 'NGG' a jusante do protoespaçador [denominada motivo adjacente ao protoespaçador – PAM (do inglês – *protospacer-adjacent motif*)] é crucial para esta ligação. Um segundo RNA, denominado crRNA transativador (tracrRNA), liga-se a crRNA, e assim a proteína Cas9 é recrutada ao complexo. Cas9 é uma endonuclease que contém dois domínios de nucleases; cada um cliva uma das duas fitas de DNA três pares de bases a montante de PAM (SCHIML et al., 2014).

CRISPR/Cas9 tipo selvagem (TS) cliva a sequência genômica invasora por uma RNA endonuclease (Cas9). A sequência invasora é capturada e inserida no genoma do hospedeiro entre regiões CRISPR. Logo após a infecção do DNA exógeno, as sequências nas regiões CRISPR são transcritas a uma única sequência de RNA denominada CRISPR RNA precursor (pré-crRNA). No sistema CRISPR/Cas9, o pré-crRNA encontra-se ligado a outros RNAs denominados RNAs transativadores. Uma vez ligados, o pré-crRNA é processado, pela RNase III, a um único duplex crRNA:tracrRNA, e juntos formam o complexo com a endonuclease (DING et al., 2016; HUMPHREY; KASINSKI, 2015).

O duplex crRNA:tracrRNA pode ser reduzido artificialmente em uma única sequência de RNA para facilitar o manuseio deste sistema. Esta condensação denominada RNA guia de sequências (sgRNA) direciona Cas9 a sequências específicas de DNA. Deste modo, o sistema mantém-se específico, eficaz e de simples manuseio, uma vez que consiste apenas de sgRNA e Cas9 (MAEDER; GERSBACH, 2016).

As sequências complementares de DNA alvo para sgRNA são ignorados pelo complexo Cas9/sgRNA, caso PAM não estiver presente a jusante do DNA alvo,

isto porque a formação do heteroduplex DNA-RNA é dependente de PAM (LAUFER; SINGH, 2015).

A especificidade do complexo CRISPR/Cas9 depende de sgRNA que pode ser direcionada a qualquer local do DNA genômico, pelo manuseio da sequência de sgRNA. Ao contrário dos outros sistemas discutidos acima, CRISPR/Cas9 não requer a engenharia de novas proteínas para cada local-alvo de DNA. A facilidade relativa com que os novos locais podem ser segmentados, simplesmente alterando a curta região do sgRNA torna este sistema um método altamente atrativo para a introdução de DSBs em locais específicos. Uma vez que a rica diversidade de sistemas CRISPR naturais tem sido pouco estudada, espera-se o surgimento de muitas novas tecnologias de edição de gene baseada em CRISPR (MAEDER; GERSBACH, 2016).

5 CONSIDERAÇÕES FINAIS

A edição genômica é uma ferramenta da biologia molecular para induzir a mudança definitiva de sequências genômicas de DNA.

Atualmente, alguns sistemas de edição genômica, tais como ZFNs, TALENs e CRISPR/Cas9 têm sido descritos. Os dois primeiros utilizam proteínas com domínios de ligação ao DNA fusionadas a endonucleases sítio-específicos. Deste modo, a especificidade proteína-DNA é determinada pela sequência de aminoácidos dos domínios de ligação da proteína. No sistema CRISPR/Cas9, por outro lado, a especificidade de interação ao DNA é baseada no princípio de Watson-Crick, por uma sequência de RNA guia (sgRNA). O DNA genômico deve ter uma sequência complementar à de sgRNA (proto-espaçador) e uma sequência 'NGG' a jusante (PAM). Uma vez que sgRNA reconhece a sequência do proto-espaçador, a nuclease Cas9 é recrutada ao hetero-complexo e cliva as duas fitas de DNA três pares de bases a montante de PAM.

O sistema CRISPR/Cas9 possui um grande diferencial e maior escalabilidade em vários locais do genoma, uma vez que o sítio de clivagem do DNA pela nuclease é dependente da sequência sgRNA, ou seja, esta pode ser construída conforme o local de interesse.

A edição genômica viabiliza modificações de porções genômicas sem

comprometer o bom funcionamento celular e apresenta um futuro promissor para diversas áreas. Assim, muitos estudos trazem perspectivas quanto à utilização das técnicas, em especial o CRISPR/Cas9, como terapia para doenças monogênicas com a utilização de células tronco. Os tratamentos de doenças por meio desse sistema não se limitam a doenças genéticas, mas também infecciosas causadas por vírus (HIV), vários tipos de câncer, viabilizar transplantes de órgãos entre animais e humanos, e eliminar o fator de rejeição.

Assim CRISPR/Cas9 pode ser aplicado em diversos campos como medicina terapêutica, agricultura e indústria.

REFERÊNCIAS

AGROTIS A., KETTELER R. A new age in functional genomics using CRISPR/Cas9 in arrayed library screening. *Front. Genet* V6. N300, 2015.

BARRANGOU R., MARRAFFINI L.A. CRISPR-Cas Systems: Prokaryotes Upgrade to Adaptive Immunity. *Molecular Cell* V54, 2014.

BEERLI R. R., DREIER B., BARBAS C.F. Positive and negative regulation of endogenous genes by designed transcription factors. *PNAS* V97. N4, 2000.

CHANDRASEGARAN S. E CARROLL D. Origins of Programmable Nucleases for Genome Engineering. *J. Mol Biol* V428. P963–989, 2016.

CERMAK T., DOYLE E. L., CHRISTIAN M. et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research* V39. N12, 2011.

DENG D., YAN C., PAN X., MAHFOUZ M. et al. Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors. *Science* V335. N6069. P720–723, 2012

DING Y., LI H., CHEN L-L., E XIE K. Recent Advances in Genome Editing Using CRISPR/Cas9 V7. N703, 2016.

FINERAN P.C. E CHARPENTIER E. Memory of viral infections by CRISPR - Cas adaptive immune systems: Acquisition of new information. *Virology* V434. P202–209, 2012.

GAJ T., GERSBACH C. A. E BARBAS C. F. ZFN, TALEN and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* V31. N7. P397–405, 2013.

GILLES A. F. E AVEROF M. Functional genetics for all: engineered nucleases, CRISPR and the gene editing revolution. *Evodevojournal* P5-43, 2014.

HEINTZE J., LUFT C. E KETTELER R. A CRISPR Cas e for high-throughput silencing. *Front Genet* V7. N4. P193, 2013.

HUMPHREY S. E E. KASINSKI A. L. RNA-guided CRISPR-Cas technologies for genome-scale investigation of disease processes. *Journal of Hematology & Oncology* V8. N31, 2015.

LAUREANO J. F. M.; *Novas Abordagens Na Terapia Genética Em Doenças Mitocondriais*. Instituto Superior De Ciências Da Saúde Egas Moniz, 2015.

LAUFER B. I. E SINGH S. M. Strategies for precision modulation of gene expression by epigenome editing: an overview. *Epigenetics&Chromatin* P8. N34, 2015.

MAEDER M. L. E GERSBACH C. A. Genome-editing Technologies for Gene and Cell Therapy. *Molecular therapy.Org* V24. N3. P430–446, 2016.

MA D. E LIU F. Genome Editing and Its Applications in Model Organisms. *Genomics. Proteomics Bioinformatics* V13. P336–344, 2015.

MAK A. N-S., BRADLEY P., BOGDANOVA A. J. E STODDARD B. L. TAL effectors: function, structure, engineering and applications. *CurrOpinStructBiol* V23. N1. P93–99, 2012.

MINCZUK M., PAPWORTH M. A., MILLER J. C., MURPHY M. P. E KLUG A. Development of a single-chain, quasi-dimeric zinc-finger nuclease for the selective degradation of mutated human mitochondrial DNA. *Nucleic Acids Research* V36. N12, 2008.

NELSON, D. L; COX, M. M. *Principles of Biochemistry (Lehninger)*. New York: W. H. Freeman and Company, 2008.

NIEWOEHNER O. E JINEK M. Structural basis for the endoribonuclease activity of the type III-A CRISPR-associated protein Csm6. *Rnajournal.Org* V22. P318–329, 2016.

REYON D., KHAYTER C., REGAN M. R., JOUNG J. K. E SANDER J. D. *Current Protocols in Molecular Biology Engineering Designer Transcription Activator-Like Effector Nucleases (TALENs)*. *CurrProtocMolBiol* V12. N15, 2012.

SASAVIC N. E SCHWANK G. *Advances in therapeutic CRISPR/Cas9 genome editing*. Elsevier Inc, 2015.

SCHIML S., FAUSER F. E PUCHTA H. The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. *The Plant Journal* V80. P1139–1150, 2014.

TAKATA M., SASAKI M. S., SONODA E., MORRISON C., MITSUMASA H., HIROSHI U. et al. Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the

maintenance of chromosomal integrity in vertebrate cells. The EMBO Journal V17. N18. P5497– 5508, 1998.