



**INSTITUTO SUPERIOR DE CIÊNCIAS DA SAÚDE
EGAS MONIZ**

MESTRADO INTEGRADO EM CIÊNCIAS FARMACÊUTICAS

**IMUNIDADE CELULAR: CARACTERIZAÇÃO DAS PROTEÍNAS
CELULARES QUE INIBEM A INFECCÃO VIRAL**

Trabalho submetido por
Ana Margarida Martins Gonçalves Lourenço

junho de 2016



**INSTITUTO SUPERIOR DE CIÊNCIAS DA SAÚDE
EGAS MONIZ**

MESTRADO INTEGRADO EM CIÊNCIAS FARMACÊUTICAS

**IMUNIDADE CELULAR: CARACTERIZAÇÃO DAS PROTEÍNAS
CELULARES QUE INIBEM A INFEÇÃO VIRAL**

Trabalho submetido por

Ana Margarida Martins Gonçalves Lourenço

Trabalho orientado por

Prof. Doutora Isabel Barahona

junho de 2016

«O que hoje não sabemos amanhã saberemos.»

Garcia de Orta

AGRADECIMENTOS

À minha orientadora Prof. Doutora Isabel Barahona pela disponibilidade demonstrada.

À Associação de Estudantes do Instituto Superior de Ciências da Saúde Egas Moniz, não só por todos estes anos de trabalho mas também por ser a minha segunda casa ao longo da minha vida Académica capaz de provar que a família não implica ligações sanguíneas.

Aos meus pais, Alzira e Rafael, pela paciência devotada, sem eles nada seria possível, por me deixaram continuar a acreditar que tudo é possível e que basta eu querer.

Ao meu querido Tomás Carvalho.

Ao Samuel.

Ao Pedro, um biólogo de alma e coração.

À Sofia, muito obrigado porque sem a sua ajuda, compreensão e companheirismo ao longo deste percurso, tudo seria muito mais difícil.

Aos meus colegas de curso, António Tanganho e Pedro Coelho por se terem mantido na minha vida do início ao fim.

À Daniela e aos nossos momentos de alegria no trabalho.

À minha avó Maria Teresa que apesar da sua ausência sei que está sempre presente nos momentos das minhas conquistas.

Ao meu padrinho e ao João, que me apoiaram sempre.

Ao meu querido amigo André que nunca me deixou desistir e sempre me acompanhou.

Ao Eduardo Guerreiro pela sua ajuda e amizade.

Aos meus amigos Marta, Ivan, Matão e Joana, que sempre acreditaram em mim e que em todos os momentos fizeram parte deste processo;

Aos meus colegas de trabalho Dr. Filipe, Dra. Marta, Dra. Rita, Dr. Ricardo, Sali e Raul que sempre me apoiaram e pelo conhecimento transmitido tanto a nível pessoal como profissional.

RESUMO

Os sistemas CRISPR-Cas (do inglês *Clustered Regularly Interspaced Short Palindromic Repeats*) emergiram na última década como um sistema imune bacteriano, capaz, de através de nucleases guiadas por um RNA, induzir danos em *loci* específicos do DNA, de forma simples e eficaz.

Esta forma de editar e manipular genes permite não só a sua protecção de elementos móveis mas também funciona como factor de virulência, de reparação de DNA e de evolução genética para além de permitir que a sua transmissão horizontalmente.

O Vírus do Papiloma Humano é responsável por um elevado número de infecções em humanos que, embora possam ser assintomáticas, atingem um elevado número da população mundial e podem evoluir para carcinogénese.

As proteínas E6 e E7 do HPV desempenham um papel crucial no desenvolvimento de tumores.

As aplicações do sistema CRISPR-Cas continuam a desenvolver-se e são consideradas surpreendentes, especialmente no que diz respeito ao sistema CRISPR-Cas9 e à interrupção da expressão dos genes E6 e E7 que permite um controle nunca antes atingido. Estudos mostram que é possível a sua utilização em vírus humanos, incluindo o HPV.

Mediante este contexto, o objectivo desta monografia centra-se na importância do sistema CRISPR e no seu mecanismo de funcionamento com o intuito de demonstrar que a sua utilização a nível humano é possível e ainda tem muito para nos revelar.

Palavras – Chave: *Clustered Regularly Interspaced Short Palindromic Repeats*, Infecção Viral, Proteínas e Vírus do Papiloma Humano

ABSTRACT

The CRISPR-Cas systems (English clustered regularly interspaced Short Palindromic Repeats) emerged in the past decade as a bacterial immune system capable of using nuclease guided by an RNA to induce damage in specific loci of DNA in a simple and effective manner.

This way to edit and manipulate genes allows not only their protection for mobile genetic elements but also virulence factors, DNA repair and genetic evolution in addition to allowing the transmission is horizontal.

The Human Papilloma Virus is responsible for a large number of infections in humans, although they may be asymptomatic, reaching a large number of the world population and may progress to carcinogenesis.

The HPV E6 and E7 proteins play a crucial role in the development of tumors.

The applications of the CRISPR-Cas system continue to be discovered and considered surprising, especially as regards the CRISPR-Cas9 system and the interruption of the expression of E6 and E7 genes allows control never before achieved. Studies allow its use in human viruses, including HPV.

Through this context, the aim of this thesis focuses on the importance of the CRISPR system and its operating mechanism in order to demonstrate that their use on the human level is possible and still has much to tell us.

Key Words: *Clustered Regularly Interspaced Short Palindromic Repeats*, Viral Infections, Proteins e Human Papillomavirus.

ÍNDICE GERAL

1. INTRODUÇÃO.....	12
2. CRISPR.....	14
2.1. COMPONENTES DO SISTEMA CRISPR.....	16
2.1.1. MATRIZ.....	16
2.1.2. SEQUÊNCIA LÍDER.....	16
2.1.3. REPETIÇÕES.....	17
2.1.4. ESPAÇADORES.....	19
2.1.5. PROTEÍNAS Cas.....	19
3. MOTIVOS ADJACENTES AOS PROTOESPAÇADORES.....	23
3.1. TIPOS DE SISTEMA CRISPR.....	23
3.1.1. CRISPR TIPO I.....	23
3.1.2. CRISPR TIPO II.....	26
3.1.3. CRISPR TIPO III.....	29
3.2. MECANISMO DE ACÇÃO DO SISTEMA CRISPR.....	31
4. APLICAÇÕES DO SISTEMA CRISPR-Cas.....	33
4.1. ENGENHARIA GENÉTICA.....	33
4.2. O SISTEMA CRISPR-CAS-9.....	35
4.2.1. TRATAMENTO DE DOENÇAS INFECCIOSAS.....	36
4.2.2. DOENÇAS GENÉTICAS.....	36
5. VÍRUS DO PAPILOMA HUMANO.....	37
5.1. ESTRUTURA DO GENOMA E ORGANIZAÇÃO.....	39
5.2. EPIEMIOLOGIA MUNDIAL DO HPV.....	40
5.3. EPIDEMIOLOGIA NACIONAL DO HPV.....	42
5.4. INFECÇÃO VIRAL POR HPV.....	43
5.5. TIPOS DE HPV.....	44

5.6. PROTEÍNAS DO HPV	45
6. PATOGÉNESE	47
7. CUSTOS ECONÓMICOS E IMPACTO NA QUALIDADE DE VIDA	49
8. APLICÇÕES DO SISTEMA CRISPR-Cas EM VÍRUS HUMANOS	49
9. O SISTEMA CRISPR-CAS 9 E HPV16 E HPV18	51
9.1. INTERRUPTÃO DO GENE E6 DO HPV-16 ATRAVÉS DO SISTEMA CRISPR-CAS	52
10. CONCLUSÃO	57
11. BIBLIOGRAFIA	58

INDÍCE DE FIGURAS

Figura 1 - Versão típica e simplificada dos loci do sistema CRISPR. Adaptado de (R. Edze, Angus, & Fineran, 2014)	15
Figura 2 – Gráfico das Distribuições do tamanho das sequências de repetições do sistema CRISPR. Adaptado de (F. J M & Garrett, 2013).....	18
Figura 3- Estrutura cristalina da proteína Cas1 de <i>P. aeruginosa</i> (A) e Cas2 de <i>S. solfataricus</i> (B). Adaptado de (Van Der Oost, Westra, Jackson, & Wiedenheft, 2014)	20
Figura 4 - Arquitectura dos <i>loci</i> dos subtipos do sistema CRISPR-Cas – Para cada representação genómica apresentam-se os respectivos <i>locus</i> de cada subunidade. A homologia genética encontra-se identificada por um código de cores e pelo nome da família. Adaptado de (Makarova et al., 2015).....	22
Figura 5- Interferência de DNA no sistema tipo I. O sistema Cascada procura no DNA uma sequência protoespaçadora e PAM. Uma vez identificada a PAM correcta e uma pequena sequência de hibridação (1), os pares de bases de crRNA hibridam com a cadeia de DNA complementar e formam uma ansa R (2). A cadeia de DNA deslocada, da ansa, serve como local de ligação da proteína Cas3 (3). Na ausência de ATP, o domínio da nuclease Cas3 (HD) cliva uma cadeia não marcada dentro de um protoespaçador (4) e produz um corte de DNA (5). Na presença de ATP, a Cas3 remodela o complexo, Cascada e DNA, tornando-o um alvo disponível para clivagem pela Cas3 na sequência protoespaçadora seja ele marcado ou não (6). A Cas3 movimenta-se na direcção 3'-5' através do domínio da helicase (Hel) enquanto o domínio HD degrada o DNA (6,7) de forma unidireccional. Adaptado de (Gasiunas et al., 2014).....	25
Figura 6 - Interferência de DNA nos sistemas CRISPR-Cas do tipoII. O complexo ternário Cas9-crRNA-tracrRNA procura no DNA a sequência protoespaçadora complementar e PAM: Uma vez que a correcta PAM e a curta primária sequência de hibridização são identificadas (1), os pares de bases do crRNA formam uma ansa R com a cadeia complementar (2). Uma vez que a ansa R se encontra formada, a proteína Cas9 cliva ambas as cadeias de DNA, marcada e não marcada utilizando o RuvC e os locais activos de HNC (3). A clivagem ocorre 3 nt antes de PAM, produzindo fragmentos DNA (4). Adaptado de (Gasiunas et al., 2014)	27
Figura 7- Interferência pelos sistemas CRISPR-Cas tipo III. No Caso do sistema III-B, o complexo Cmr procura pares de bases de RNA e crRNA com uma sequência	

protoespaçadora correspondente (1). Dois mecanismos de clivagem de RNA são propostos. O complexo Cmr de *P. furiosus* induz um único corte no RNA alvo a partir da extremidade 3' do crRNA (2) para produzir dois fragmentos (3) O complexo Cmr de *S. sulfolobus* guiado por crRNA corta o RNA alvo em vários locais, especificamente nos dinucleótidos UA (4) em várias posições (5). Os componentes do complexo Cmr envolvidos na clivagem ainda não estão identificados. Adaptado de (Gasiunas et al., 2014)..... 31

Figura 8 - Representação das três fases do sistema CRISPR-Cas : adaptação, expressão e interferência. No tipo I e tipo II dos sistemas CRISPR-Cas, mas não no tipo III, a selecção de protoespaçadores do ácido nucleico invasor depende de um protoespaçador adjacente ao motivo (PAM), mas esta etapa de reconhecimento ainda não é clara. Após o reconhecimento inicial, as proteínas Cas1 e Cas2 integram os protoespaçadores nos locus CRISPR. Adaptado de (Makarova et al., 2012) 32

Figura 9 - Marcos na evolução de sistemas de manipulação genética até ao aparecimento do sistema CRISPR-Cas. Adaptado de (GenScript, n.d.) 34

Figura 10- Representação 3D do vírus do Papiloma Humano (HPV) Adaptado de <http://3dciencia.com/blog/wp-content/uploads/2011/03/papiloma-virus-HPV-papillomavirus-what-you-should-know-about-viru.jpg> 37

Figura 11 - Relação evolutiva entre os HPV-. Árvore filogenética representativa da semelhança entre os genes E1, E2, L2 e L1 de 132 HPVs. Divisão entre os cinco grupos evolutivos com diferentes tropismos epiteliais e ligação a patologias. Adaptado de (Doorbar et al., 2012) 38

Figura 12 - O genoma viral e dividido em três regiões: uma região não codificante denominada LCR; e duas codificantes, designadas região precoce e tardia. Os ORFs da região precoce são apresentados através de setas ou barras, enquanto os da região tardia através de setas ou barras, enquanto os da região tardia através de setas cinza ou preta e barras listradas. Podem também verificar-se os quatro promotores (P1, P2, P3 e P4) e os dois locais de poli-adenilação Adaptado (Lima et al., 2013) 40

Figura 13 - Prevalência de HPV em mulheres cuja citologia se apresenta normal. Adaptado de (Forman et al., 2012) 41

Figura 14 - Prevalência do HPV no mundo, adaptado de (Crow, 2012)..... 42

Figura 15 - Processo de infecção por HPV adaptado de (Lazarczyk, Cassonnet, Pons, Jacob, & Favre, 2009) 44

Figura 16- Representação da mutagénese do DNA viral pela CRISPR-Cas9. Adaptado de (White et al., 2015)	50
Figura 17 - Organização genética do HPV-16. Adaptado de (Lazarczyk et al., 2009)..	52
Figura 18 - Representação da edição do gene E6 através do sistema CRISPR-Cas. Em A verifica-se a quebra do gene E6 (DSBs) e em B os gRNAs marcam o gene E6. Adaptado de (Yu et al., 2014).....	53
Figura 19 - Clivagem do DNA do gene E6 do HPV 16 pelo complexo CRISPR gRNA/Cas9.....	54
Figura 20 - Apoptose induzida pelo sistema CRISPR-Cas. SiHa (A), CaSki (B), C33A(C) e HEK293 (D). Adaptado de (Yu et al., 2014)	55
Figura 21 - Análise de Western Blot e análise por ImageJ. Adaptado de (Yu et al., 2014)	55

ÍNDICE DE TABELAS

Tabela 1 - Mecanismos de imunidade adquirida mediada pelo sistema CRISPR tipo I relativamente a seis organismos. Adaptado de (Gasiunas et al., 2014).....	24
Tabela 2 - Mecanismos de imunidade adquirida mediada pelo sistema CRISPR tipo II relativamente a três organismos. Adaptado de (Gasiunas et al., 2014).....	28
Tabela 3 - Mecanismos de imunidade adquirida mediada pelo sistema CRISPR tipo III relativamente a três organismos. Adaptado de (Gasiunas et al., 2014).....	30
Tabela 4- Principais diferenças entre TALENs, ZFNs e CRISPR-Cas. Adaptado de (GenScript, n.d.)	35
Tabela 5 - Funções das proteínas do HPV, adaptado de Boulet et al, 2007	47
Tabela 6 - Relação entre o tipo de HPV e a doença associada. Adaptado de (Souto et al., 2005).....	48
Tabela 7 - Aplicação do Sistema CRISPR-Cas9 a HPV16 e HPV18. Adaptado de (White et al., 2015).....	51

ÍNDICE DE ABREVIATURAS

Cas – Proteínas associadas ao CRISPR (do inglês *CRISPR-associated proteins*)

Cascada – Complexo associado para defesa antiviral (do inglês *associated complex for antiviral defense*)

CRISPR - Repetições palindrômicas curtas agrupadas e regularmente intespaçadas (do inglês *Clustered Regularly Interspaced Short Palindromic Repeats*)

CRISPR-Cas - Repetições palindrômicas curtas agrupadas e regularmente intespaçadas associadas a proteínas Cas (do inglês *Clustered Regularly Interspaced Short Palindromic Repeats-associated Cas proteins*)

crRNA – RNA pequeno processado do locus CRISPR (do inglês *small encoded RNA*)

DNA – Ácido desoxirribonucleico (do inglês *Deoxyribonucleic Acid*)

DSBs – Quebras de cadeia dupla (do inglês *double-strand breaks*)

dsDNA - Ácido desoxirribonucleico de cadeia dupla (do inglês *Deoxyribonucleic Acid double-stranded*)

E – precoce (do inglês *early*)

E. coli – *Escherichia coli*

EMV – Elementos genéticos móveis

gRNA – Ácido ribonucleico guia (do inglês *guide Ribonucleic acid*)

H. ochraceum- *Haliangium ochraceum*

HIV – Vírus da Imunodeficiência adquirida (do inglês *Human Immunodeficiency virus*)

HPV – Vírus do papiloma humano (do inglês *Human Papilloma Virus*)

L - tardia (do inglês *late*)

LCR – região longa de controlo (do inglês *long control region*)

M. Jannaschii – *Methanacaldococcus jannaschii*

mRNA – Ácido ribonucleico mensageiro (do inglês *Messenger Ribonucleic acid*)

NHEJ- Recombinação terminal não homóloga (do inglês *nonhomologous end joining*)

nt – Nucleótido (do inglês nucleotide)

OMS - Organização Mundial de Saúde

ORF – quadro de leitura aberta (do inglês open Reading frame)

P. aeruginosa - Pseudomonas aeruginosa

P. furiosus - Pyrococcus furiosus

PAM – Motivos adjacentes aos protoespaçadores (do inglês Protospacer-adjacent motif)

Pb- pares de bases

Poli-A - poliadenilação

pre-crRNA - precursor crRNA

RAMP – Proteínas receptoras de modificação de actividade (do inglês Receptor activity-modifying proteins)

Rb - Retinoblastoma

RNA – Ácido Ribonucleico (do inglês Ribonucleic acid)

RNAi - Ácido Ribonucleico de interferência (do inglês Ribonucleic acid interference)

RNase - Ribonuclease

RNAss - Ácido Ribonucleico de cadeia única (do inglês single- stranded Ribonucleic acid)

RNP – ribonucleoproteínas (do inglês Ribonucleoproteins)

S. epidermidis - Staphylococcus epidermidis

S. islandicus – Staphylococcus islandicus

S. Pyogenes- StreptococcusPyogenes

S. solfataricus -Sulfolobus solfataricus

S. thermophilus - Streptococcus thermophilus

sgRNA – Ácido ribonucleico guia pequeno (do inglês short guide Ribonucleic acid)

siRNA - Ácido ribonucleico pequeno de interferência (do inglês small interfering ribonucleic acid)

sRNA – Pequeno ácido ribonucleico (do inglês small Ribonucleic acid)

TALEN – (do inglês transcription-activator-like effector nuclease)

TBE – (do inglês Tris-Borate-EDTA)

Tfu or T. fusca - Thermobifida fusca

tracrRNA – (do inglês Trans-activating CRISPR RNA)

ZFN – (do inglês zinc finger nuclease)

1. INTRODUÇÃO

O sistema imunitário é formado por uma complexa rede de órgãos, células e moléculas que concomitantemente se organizam para defender o organismo de diversas agressões.

De forma a compreender os mecanismos de actuação do sistema imunitário é também necessário estudar os potenciais agentes infecciosos como vírus, bactérias, fungos e parasitas.

O estudo dos múltiplos mecanismos de defesa contra os agentes virais é muito importante dado que as doenças virais são muito comuns, e representam uma preocupação no que diz respeito à evolução para doenças cancerígenas e até mesmo morte.

Nos últimos vinte anos, os sistemas de manipulação genética evoluíram de forma extraordinária. Há cerca de uma década surge uma técnica revolucionária, baseada na descoberta do sistema de sobrevivência de eubactérias e de arqueobactérias que constantemente são ameaçadas por vírus, designada de CRISPR, que permitiu vários avanços na área da biologia e engenharia genética. Este sistema imunitário bacteriano permite não só excluir, como substituir ou mesmo impedir a edição de DNA, utilizando uma proteína bacteriana modificada e um RNA que a orienta para uma sequência de DNA específica, proporcionando um controlo nunca anteriormente alcançado, fácil, eficiente e de rápida utilização que ainda hoje surpreende a comunidade científica.

A possibilidade de manipular o DNA permite uma melhor compreensão do funcionamento do genoma, o que é extremamente relevante para o avanço de várias áreas da biologia que permitem a identificação de novos alvos para intervenção médica.

O Vírus do Papiloma Humano é um dos vírus com maior prevalência no que diz respeito a doenças sexualmente transmissíveis e segundo a organização Mundial de Saúde afecta cerca de 440 milhões de pessoas em todo o mundo. Uma vez que o seu genoma tem uma predisposição para a perda de estabilidade, este encontra-se associado a vários tipos de cancro sendo considerado um problema de saúde pública.

Introdução

Com esta revisão bibliográfica pretende-se dar a conhecer o sistema CRISPR-Cas, o seu mecanismo de funcionamento, algumas das suas aplicações, especialmente no que diz respeito ao HPV.

2. CRISPR

Os bacteriófagos são os vírus mais abundantes na atmosfera e infectam bactérias com o objectivo de se reproduzirem e como consequência podem inclusive provocar a morte da célula hospedeira quando a sua replicação se encontra concluída, existindo no entanto fagos que embora possuam capacidade lisogénica não provocam a morte de bactérias (Gasiunas, Sinkunas, & Siksnyš, 2014).

As bactérias (eubactérias e arqueobactérias) desenvolveram por isso diversos mecanismos de defesa que atingem vários passos do ciclo de vida dos fagos, bloqueando a adsorção, prevenindo a injeção de ácido desoxirribonucleico (DNA) viral, restringindo a entrada do mesmo e sistemas de infecção abortivos (Barrangou et al., 2007).

Em 1987, surge a descoberta de um sistema de repetições palindrómicas curtas agrupadas e regularmente interespaçadas, (CRISPR), no genoma de *Escherichia coli*, aquando da sequenciação de um fragmento da mesma (Ishino, Shinagawa, Makino, Amemura, & Nakata, 1987).

Foram no entanto necessários vários anos para que estas matrizes de repetições se passassem a designar por CRISPR (Jansen, Van Embden, Gaastra, & Schouls, 2002) e apenas em 2000 se demonstra que desempenham um papel fundamental tanto no domínio *archae* como eubacteriológico (Mojica, Díez-Villaseñor, Soria, & Juez, 2000) confirmando-se posteriormente que existe um elevado nível de homologia entre espécies filogeneticamente distintas (Mojica, Díez-Villaseñor, García-Martínez, & Soria, 2005)

Estas repetições podem encontrar-se em cerca de 84% de *archae* e em 45% de eubactérias (Rath, Amlinger, Rath, & Lundgren, 2015).

Este sistema confere protecção contra elementos genéticos móveis (EMG), tais como plasmídeos e fagos, cujo DNA se pode integrar no DNA da célula hospedeira no *locus* CRISPR.

As CRISPR são formadas pela aquisição de ácidos nucleicos exógenos, designados por espaçadores, que são incorporados em sequências de DNA endógeno – repetições; podendo os primeiros ser obtidos por diversas fontes tais como fagos ou plasmídeos (Barrangou et al., 2007).

Os *loci* CRISPR sofrem alterações de forma a providenciar imunidade a uma determinada geração permitindo que esta possa ser transmitida para a geração seguinte. Esta herança de características adquiridas é o que mais se assemelha à herança de Lamarck, em que um organismo é capaz de reagir a alterações ambientais e gerar uma resposta adaptativa no genoma (Koonin & Makarova, 2013) (Figura 1)

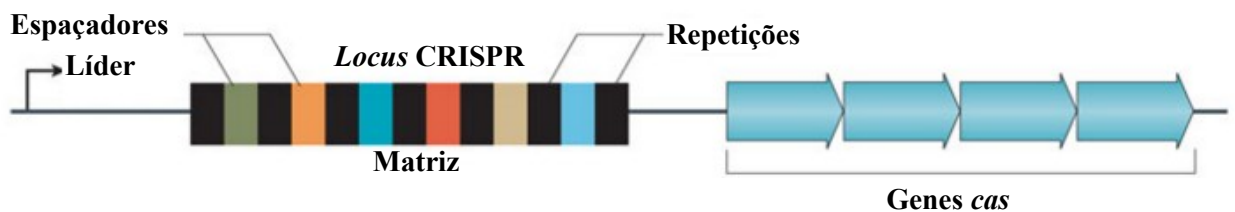


Figura 1 - Versão típica e simplificada dos loci do sistema CRISPR. Adaptado de (R. Edze, Angus, & Fineran, 2014)

Para que desempenhe e mantenha a sua actividade o sistema necessita da presença de genes *cas*, associados a CRISPR (CRISPR associated sequences). Os genes *CAS* geralmente encontram-se adjacentes a CRISPR e codificam proteínas essenciais à resposta imunitária, identificando o DNA dos fagos e incorporando-o no *locus* do sistema (Luiciano A. & Sontheimer, 2010)

A partir do momento em que o genoma é modificado pelo processo de aquisição de espaçadores (DNA exógeno) ocorre a protecção a uma nova infecção pelo mesmo vírus. Novos espaçadores podem ser incluídos na mesma extremidade de CRISPR, fazendo com que haja uma gravação cronológica dos vírus que infectaram as células num determinado momento e também dos vírus antecedentes que também infectaram a mesma célula (Rath et al., 2015). Após a transcrição desta zona do DNA, formam-se ácidos ribonucleicos pequenos (sRNA) processados a partir da matriz de CRISPR que contêm parte da sequência dos espaçadores, que vão formar um complexo com as proteínas CAS expressas. Aquando de uma nova infecção pelo mesmo vírus, formam-se estes complexos ribonucleicos que vão por sua vez ligar-se por complementaridade de bases ao ácido nucleico do invasor, clivando dentro desta sequência de nucleótidos do invasor (interferência), resultando na inactivação do vírus (F. J M & Garrett, 2013).

Desta forma, o sistema foca-se na marcação de DNA ou RNA como meio de proteger o organismo contra vírus e outros EMG (Shah, Erdmann, Mojica, & Garrett, 2013).

2.1. COMPONENTES DO SISTEMA CRISPR

Para que o sistema CRISPR funcione existem elementos fundamentais tais como a Matriz com uma sequência líder a montante repetições, espaçadores e os genes Cas.

2.1.1. MATRIZ

Na matriz (Figura 1) as repetições alternam com as sequências espaçadoras (Ishino et al., 1987). A maior parte das matrizes do sistema CRISPR envolvem de duas a cem repetições, embora existam matrizes ainda mais longas. A mais longa matriz encontrada pertence à bactéria halofílica, *Haliangium ochraceum* DSM 14358 e contém 588 repetições (F. J M & Garrett, 2013).

Existem no entanto organismos do grupo *archae* que possuem diversos *loci* de CRISPR, como o *Methanocaldococcus* sp. FS406-22, com 23 matrizes genômicas. Embora a maioria dos CRISPR se encontrem em regiões do DNA bacteriano com elementos provirais, as repetições também são encontradas em plasmídeos, que funcionam como veículos de transferência horizontal para algumas bactérias (Godde & Bickerton, 2006). Foram também encontradas matrizes em alguns genomas virais livres e em pequenos plasmídeos conjugados de *Sulfolobus* (Mojica, Díez-Villaseñor, García-Martínez, & Almendros, 2009).

Através da utilização das proteínas CAS do hospedeiro, existe uma base para uma interação competitiva entre vírus e plasmídeos no interior da célula (F. J M & Garrett, 2013).

2.1.2. SEQUÊNCIA LÍDER

As sequências líder encontram-se geralmente na matriz do sistema CRISPR e orientam-se especificamente para a sequência de repetição adjacente ao CRISPR. Estas são consideradas sequências que não codificam proteínas e têm baixa complexidade (F. J M & Garrett, 2013).

O seu grau de conservação de sequência total é relativamente baixo e o seu tamanho real pode ser difícil de estimar. Alguns organismos como *M. jannaschii* contêm diversos *loci* de CRISPR com sequências líder idênticas cuja identificação pode ser realizada por comparação da sequência de nucleótidos (Mojica et al., 2009)

Embora o grau de conservação das sequências líder geralmente siga a repetição da matriz CRISPR associada, existem algumas derivações desta tendência que sugerem que trocas entre diferentes líderes e matrizes com repetições podem ocorrer. (Mojica et al., 2009)

As sequências líderes assumem duas funções distintas, uma das funções consiste em conter o principal promotor para transcrição do *locus* contíguo do CRISPR (Westra et al., 2010) e caso o líder esteja ausente observam-se baixos níveis de transcrição, utilizando como promotor motivos retirados aleatoriamente dos espaçadores ou mesmo do interior de uma repetição (Deng et al., 2012). A outra função consiste em ser o local de aquisição onde novos espaçadores são primariamente incorporados, perto da primeira repetição (Held, Herrera, Quiroz, & Whitaker, 2010).

Existem estudos filogenéticos em sistemas CRISPR *crenarchael* que evidenciam a co-evolução de líderes, repetições e sequências de proteínas Cas 1, enfatizando o papel dos líderes no processo de aquisição, facilitando a inserção de novas unidades proto-espaçadoras para um local de montagem perto das proteínas Cas (Luiciano A. & Sontheimer, 2010).

Alguns *loci* do sistema CRISPR que se mantêm em algumas células parecem inactivos, mas se forem transferidos para perto do promotor da sequência líder ou de um promotor dentro da matriz do CRISPR, são capazes de manter a sua actividade de interferência (Lillestøl et al., 2009). Em 2012, Yosef et al., demonstraram que os elementos de DNA na sequência líder são necessários para o processo de aquisição (Yosef, Goren, & Qimron, 2012).

2.1.3. REPETIÇÕES

As repetições associadas ao CRISPR sequenciadas em diferentes genomas variam entre 23 e 55 pb, sendo visível um pico nas 24-25, 29-30 e 36-37 pb, onde cada um sofre separação de meia volta helicoidal de dsDNA (ver Figura 2).

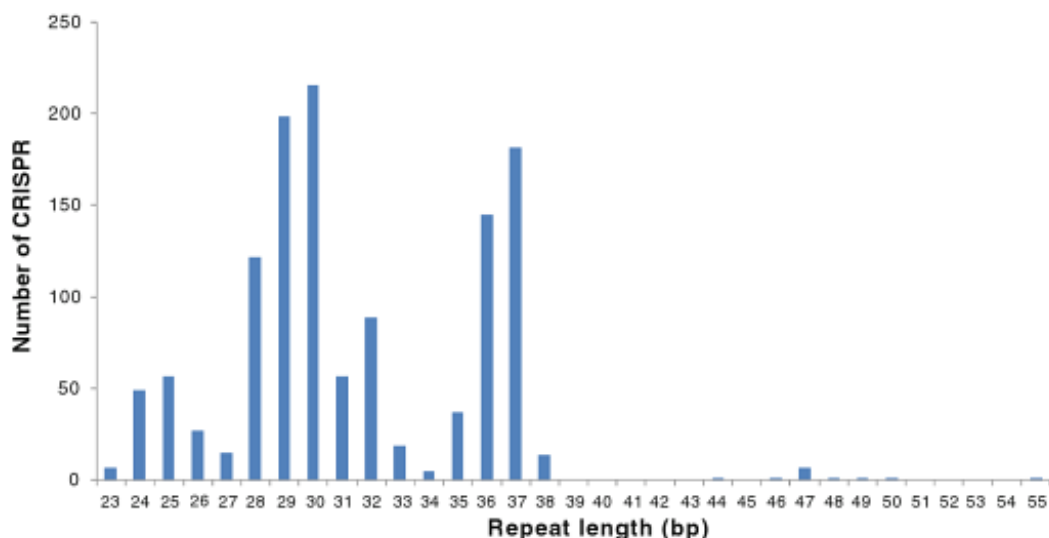


Figura 2 – Gráfico das Distribuições do tamanho das sequências de repetições do sistema CRISPR. Adaptado de (F. J M & Garrett, 2013)

Muito poucas repetições têm tamanho superior a 38 pb e 55 pb, sendo *Desulfobacca acetoxidans* DSM 11109 uma das poucas que excedem os 50 pb (Gráfico 1) (F. J M & Garrett, 2013).

Embora quatro cópias destas repetições sejam agrupadas e regularmente espaçadas, ao longo do genoma não se encontram ligadas aos genes Cas, o que sugere que não sejam funcionais com o sistema CRISPR. O menor tamanho de 23pb é encontrado nas repetições do *archaeon Ferroglobus placidus* DSM 10642. Enquanto a maior parte das matrizes do sistema CRISPR possuem cerca de 100 repetições, estas podem variar de 2 a 588 repetições no *H. ochraceum* DSM 14365, sendo este organismo capaz de possuir 815 repetições (F. J M & Garrett, 2013).

As repetições do sistema classificam-se em 12 grupos *major* e 21 *minor* com base na similaridade da sequência (Kunin, Sorek, & Hugenholtz, 2007).

As sequências de repetição são geralmente altamente conservadas numa matriz, embora a repetição que se encontra na zona mais distal da matriz da sequência líder sofra, por vezes, degeneração (F. J M & Garrett, 2013).

As repetições presentes nos *loci* do sistema CRISPR de um dado organismo podem ser altamente divergentes, no entanto, repetições idênticas são encontradas em organismos pouco relacionados. O que leva a concluir que os sistemas CRISPR sofrem uma transferência genética horizontal (Chakraborty et al., 2010).

2.1.4. ESPAÇADORES

Os espaçadores localizados dentro de uma matriz de repetição apresentam um tamanho similar, variando apenas em alguns nucleóticos e possuem sequências únicas. O seu tamanho varia de 21 a 72 pb, mas a sua maioria ou possui 32 ou varia entre 35-27 pb.

As sequências espaçadoras contidas nas matrizes do sistema CRISPR são altamente variáveis e apenas em diferentes estirpes de algumas espécies existem espaçadores idênticos dentro das suas matrizes. Estes espaçadores tendem a estar concentrados na zona terminal da matriz mantendo a sua ordem. (F. J M & Garrett, 2013)

2.1.5. PROTEÍNAS CAS

Os *loci* do sistema CRISPR contêm grupos de genes codificadores de proteínas designados por *cas*. Estas proteínas possuem diversos domínios característicos de helicases, polimerases e proteínas ligadas a RNA, passíveis de identificação através de análise computacional, que fazem parte do sistema de reparação de DNA (Makarova, Aravind, Grishin, Rogozin, & Koonin, 2002).

Estas desempenham um importante papel na diferenciação entre os tipos e subtipos de CRISPR-Cas. As proteínas Cas1 e Cas2 encontram-se envolvidas na fase de adaptação e consideram-se universais aos sistemas CRISPR-Cas. (Figura 3) Existem no entanto outras proteínas que apenas se associam a alguns tipos específicos de sistemas CRISPR. Tal acontece devido à diversidade das proteínas Cas, à presença de diversos *loci* e à transferência horizontal dos sistemas CRISPR-Cas (Rath et al., 2015)

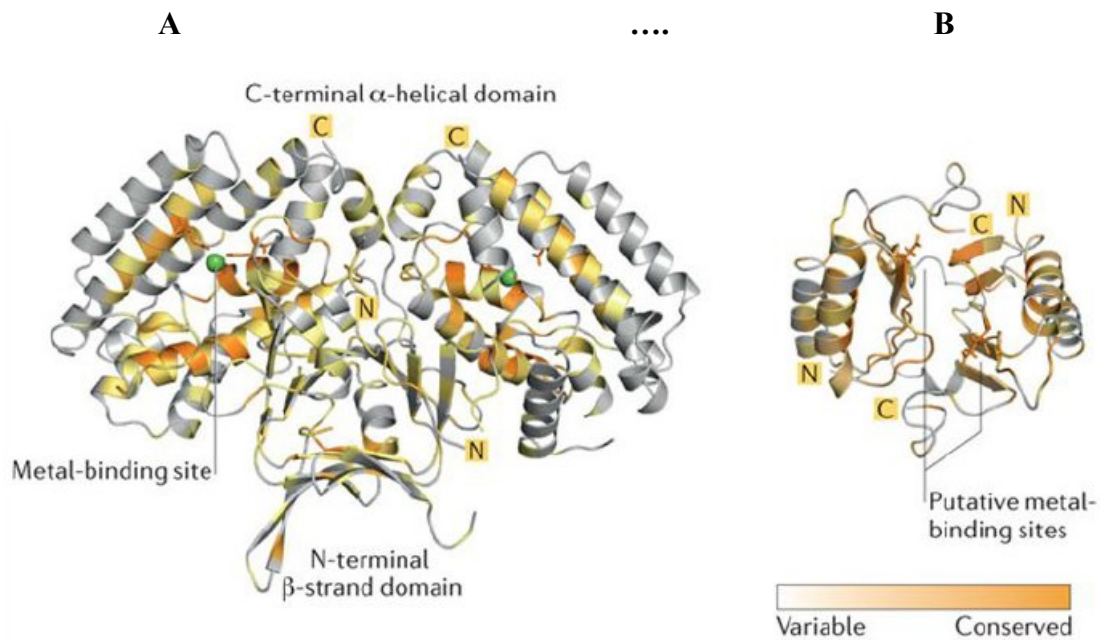


Figura 3- Estrutura cristalina da proteína Cas1 de *P. aeruginosa* (A) e Cas2 de *S. solfataricus* (B). Adaptado de (Van Der Oost, Westra, Jackson, & Wiedenheft, 2014)

Podem assim distinguir-se dois grupos principais de proteínas: o primeiro inclui as proteínas universais Cas1 e Cas2 e as proteínas específicas para um tipo *major*, Cas3, Cas9 e Cas10 que servem de proteínas base para os tipos de sistema I, II e III respectivamente (Makarova, Brouns, Horvath, Sas, & Wolf, 2012).

O segundo grupo inclui proteínas encontradas apenas nos *clusters* de genes de um determinado subtipo, como é o caso da Cse e Csy específicas para os tipos I-E e I-F, respectivamente (Makarova et al., 2012).

Os sistemas podem ser complexos uma vez que algumas bactérias contêm múltiplos subtipos de CRISPR-Cas e cada um deles tem diversas matrizes que funcionam com o *cluster* Cas apropriado (van Belkum, Scherer, van Alphen, & Verbrugh, 1998).

Relativamente à proteína Cas1 da *Pseudomonas aeruginosa*, visível na figura 2, bioquimicamente considera-se uma proteína dependente de metal, cuja sequência não corresponde a uma endonuclease do DNA (Zegans et al., 2009). A sua função prende-se com o reconhecimento e clivagem do DNA invasor e à integração de espaçadores no *locus* do CRISPR (Deveau, Garneau, & Moineau, 2010) bem como à inserção/delecção de espaçadores de CRISPR por recombinação (Mojica et al., 2009).

A proteína Cas2 do *S. solfataricus*, visível na figura 2, exibe um domínio de ferredoxina e apresenta actividade endonucleásica dependente de metal (Beloglazova et al., 2008).

Esta proteína cliva o RNAss (RNA com uma única cadeia, em português) em regiões ricas em uracilo e pressupõe-se que uma das suas funções seja a incorporação de novos espaçadores nos *loci* do CRISPR (Szczepankowska, 2012).

A proteína Cas3, cataliza de forma independente do ATP o emparelhamento RNA -DNA, uma ligação dependente de metal (Howard, Delmas, Ivančić-Baće, & Bolt, 2011).

Relativamente à proteína Cas4, segundo Makarova esta tem como função hidrolisar os DNA invasores (Makarova, Grishin, Shabalina, Wolf, & Koonin, 2006).

Os genes *cas5*, *cas6* e *cas7* codificam para proteínas da superfamília RAMP (Makarova et al., 2006) e encontram-se implicados na clivagem dos transcritos CRISPR RNA (crRNA)(Makarova et al., 2012)

Makarova *et. al*, em 2005 analisam especificamente as proteínas Cas num contexto genómico e dividem-nas em 25 famílias, que se pensa estarem ligadas ao mecanismo de defesa contra fagos e plasmídeos cuja funcionalidade é análoga ao sistema eucariótico de ácido ribonucleico de interferência (RNAi), visível na figura 4 (Makarova et al., 2006).

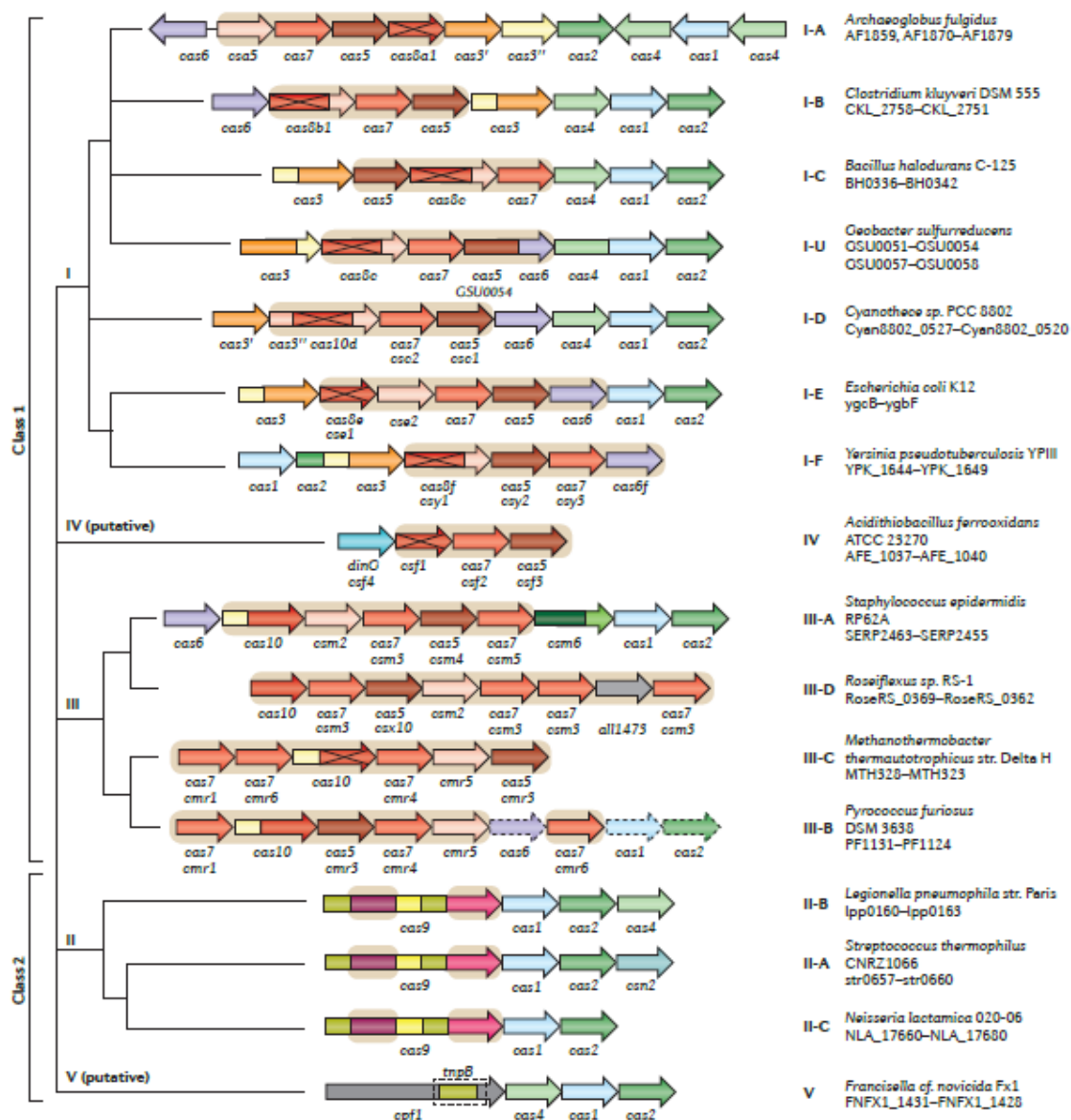


Figura 4 - Arquitetura dos loci dos subtipos do sistema CRISPR-Cas – Para cada representação genómica apresentam-se os respectivos locus de cada subunidade. A homologia genética encontra-se identificada por um código de cores e pelo nome da família. Adaptado de (Makarova et al., 2015).

3. MOTIVOS ADJACENTES AOS PROTOESPAÇADORES

Os motivos adjacentes aos proto-espaçadores (PAMs) consistem numa sequência que ocorre no fim dos proto-espaçadores para cada sistema CRISPR, dependendo do tipo de sistema e do organismo em questão (Gudbergdottir, Chen, Jensen, Jensen, & She, 2011).

Estes são identificados por alinhamentos sequenciais de elementos genéticos que contenham sequências que correspondem espaçadores dentro de matrizes de CRISPR individuais. A identidade da sequência de consenso PAM provavelmente depende de dois processos moleculares distintos: reconhecimento por aquisição do complexo proteico de aquisição e a subsequente selecção de proto-espaçadores determinados (Shah et al., 2013)

A presença deste sistema é essencial para o processo de interferência abaixo descrito, em ambos os sistemas CRISPR tipo I e II (Gudbergdottir et al., 2011) não desempenhando no entanto a mesma importância nos sistemas tipo III (Hale et al., 2012)

3.1. TIPOS DE SISTEMA CRISPR

Os sistemas CRISPR-Cas são classificados em três tipos principais baseados na organização do seu *locus* e no conteúdo do gene *cas*: I, II e III. Mais de um tipo de sistema pode ser encontrado num organismo, sugerindo que os seus sistemas sejam compatíveis e possam partilhar componentes funcionais (Makarova et al., 2012)

3.1.1. CRISPR TIPO I

Os sistemas tipo I encontram-se divididos em seis subtipos que se distinguem entre si pelo número e pelo arranjo dos genes *cas*. Embora existam diferenças entre si, estes sistemas codificam uma proteína Cas3 característica e uma proteína universal Cas1 (Makarova et al., 2012)

No que diz respeito aos mecanismos de imunidade adquirida mediada pelo sistema CRISPR do tipo I, estudaram-se seis organismos. (Tabela 1). As espécies *E. coli* e *Streptococcus thermophilus* pertencem ambas ao subtipo I-E, enquanto as restantes,

Sulfolobus solfataricus, *Haloferax volcanii*, *Bacillus halodurans* e *P. aeruginosa*, pertencem respectivamente ao subtipo A, B, C e F. Embora o comprimento da repetição e sequências no tipo I da matriz do CRISPR variem entre os sistemas do modelo, as sequências nucleotídicas dentro da repetição são palindrômicas, exceptuando no tipo A. (Karvelis et al., 2013)

Organismo	Complexo acrónimo	Subtipo	Comprimento da repetição, nt	Sequência de Repetição	Comprimento de espaçadores, nt	PAM ^b ^c (5'-3')	Comprimento de crRNA	Composição do Complexo
<i>Sulfolobus solfataricus</i>	Ss-Cascade (aCascade)	I-A	24-25	5'-GATAATCTCTT ATAGAATTGAAAAG-3' nd	38-44	Nd	60-70	Cas7, Cas5, Csa5, Cas6
<i>Haloferax volcanii</i> ^e		I-B	30	5'- GCTTCAACCCACAAG GGTTCGTCTGAAAC-3'	34-39	TTC, ACT, TAA, TAT, TAG, CAC	64-69	Nd
<i>Bacillus halodurans</i>	Bh-Cascade	I-C	32	5'- GTCGCACTCTACATGA GTGCGTGGATTGAAAT- 3'	33-36	Nd	65 ^f	Cas8c ₁ :Cas7 ₆ :Cas5d ₂
<i>Escherichia coli</i> K-12	Ec-Cascade	I-E	29 (28) ^h	5'-(G)AGTTCCCGCGCC AGCGGGGATAAACCG- 3' ^h	32 (33) ^h	AWG (AW) _h	61	Cse1 ₁ :Cse2 ₂ :Cas7 ₆ :Cas5 ₁ :Cas6e ₁
<i>Streptococcus thermophilus</i> D GCC7710	St-Cascade	I-E	28	5'- GTTTTTCCCGCACAC GC GGGGGTGATCC-3'	33	W	61	Cse1, Cse2, Cas7, Cas5, Cas6e
<i>Pseudomonas aeruginosa</i> PA14	Pa-Cascade	I-F	28	5'-GTTCACTGCCGTA TAGGCAGCTAAGAAA-3'	32	GG	60	Csy1 ₁ :Csy2 ₁ :Csy3 ₆ :Cas6f ₁

Tabela 1 - Mecanismos de imunidade adquirida mediada pelo sistema CRISPR tipo I relativamente a seis organismos. Adaptado de (Gasiunas et al., 2014)

No tipo I, as matrizes de repetições de espaçadores são transcritas num precursor de crRNA, designado pré-crRNA, onde uma sequência palindrômica de repetições forma uma ansa que é reconhecida e processada pela endorribonuclease Cas6 ou Cas5d de forma a gerar um crRNA maduro. Posteriormente crRNA é incorporado num complexo Cascade, com múltiplas subunidades proteicas, que juntamente com a proteína Cas3 silenciam o DNA invasor (Gasiunas et al., 2014) (Figura 5).

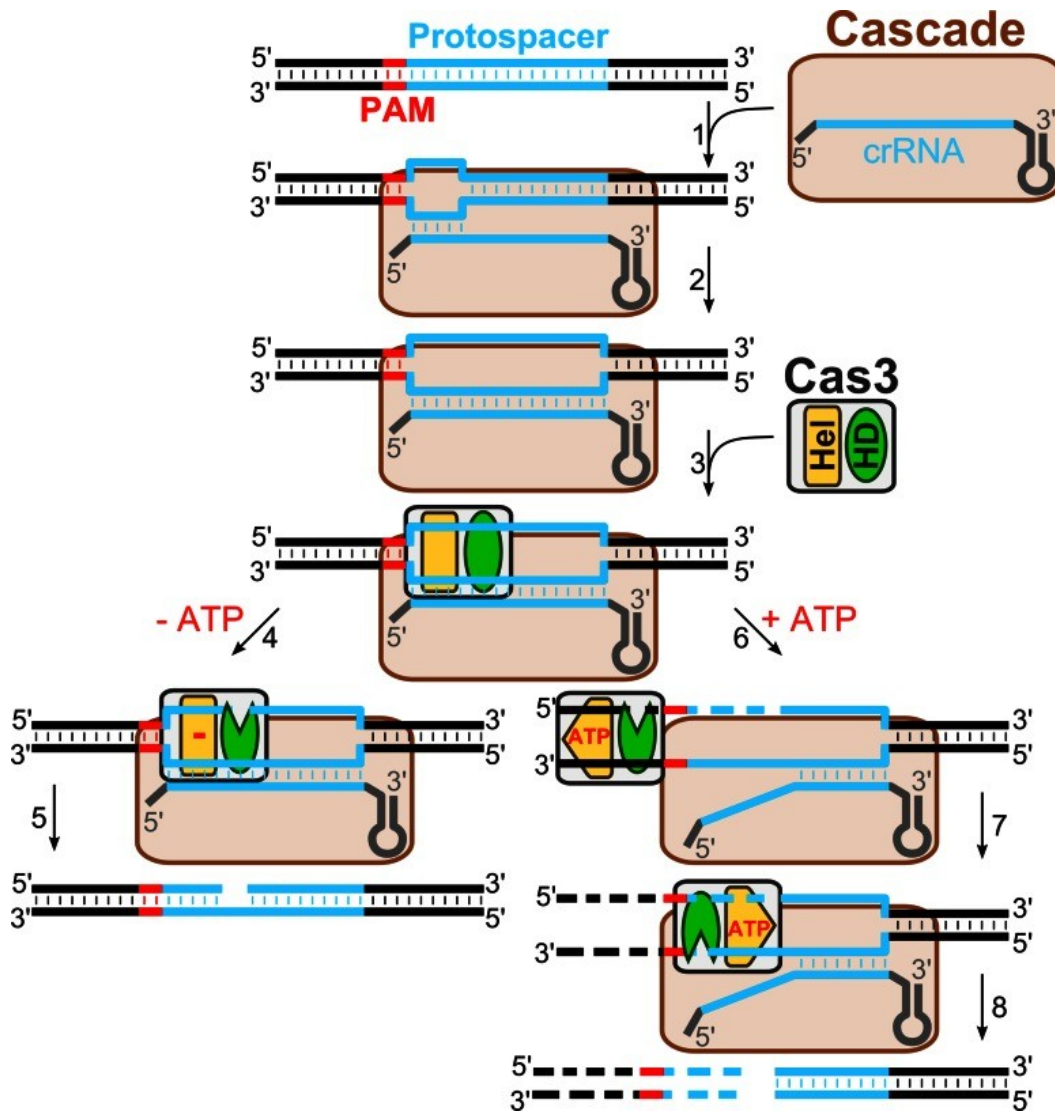


Figura 5- Interferência de DNA no sistema tipo I. O sistema Cascade procura no DNA uma sequência protoespaçadora e PAM. Uma vez identificada a PAM correcta e uma pequena sequência de hibridação (1), os pares de bases de crRNA hibridam com a cadeia de DNA complementar e formam uma ansa R (2). A cadeia de DNA deslocada, da ansa, serve como local de ligação da proteína Cas3 (3). Na ausência de ATP, o domínio da nuclease Cas3 (HD) cliva uma cadeia não marcada dentro de um protoespaçador (4) e produz um corte de DNA (5). Na presença de ATP, a Cas3 remodela o complexo, Cascade e DNA, tornando-o um alvo disponível para clivagem pela Cas3 na sequência protoespaçadora seja ele marcado ou não (6). A Cas3 movimenta-se na direcção 3'-5' através do domínio da helicase (Hel) enquanto o domínio HD degrada o DNA (6,7) de forma unidireccional. Adaptado de (Gasiunas et al., 2014)

3.1.2. CRISPR TIPO II

No sistema CRISPR tipo II, a operação dos genes *cas* codifica apenas três de quatro proteínas Cas (Makarova et al., 2012) incluindo as universais Cas1 e Cas2 e as proteínas Cas4 e Cas2 que se encontram envolvidas no processo de aquisição de espaçadores.

Estes tipos subdividem-se em três subtipos II-A, II-B e II-C, todos eles contendo como assinatura a proteína Cas9 (Makarova et al., 2012).

Neste sistema CRISPR tipo II as endoribonucleases Cas6 e Cas5d que se encontram envolvidas no processamento do pré-crRNA nos sistemas tipo I, encontram-se ausentes. Em *Streptococcus pyogenes*, o sistema CRISPR –Cas II-A consiste num pequeno RNA trans-codificado designado de tracrRNA ou CRISPR RNA trans-activador, com repetições de 24 nt complementares às repetições de pré-crRNA, que contribui para a maturação do crRNA num processo que necessita da presença da RNase III do hospedeiro (Deltcheva et al., 2011) Figura 6.

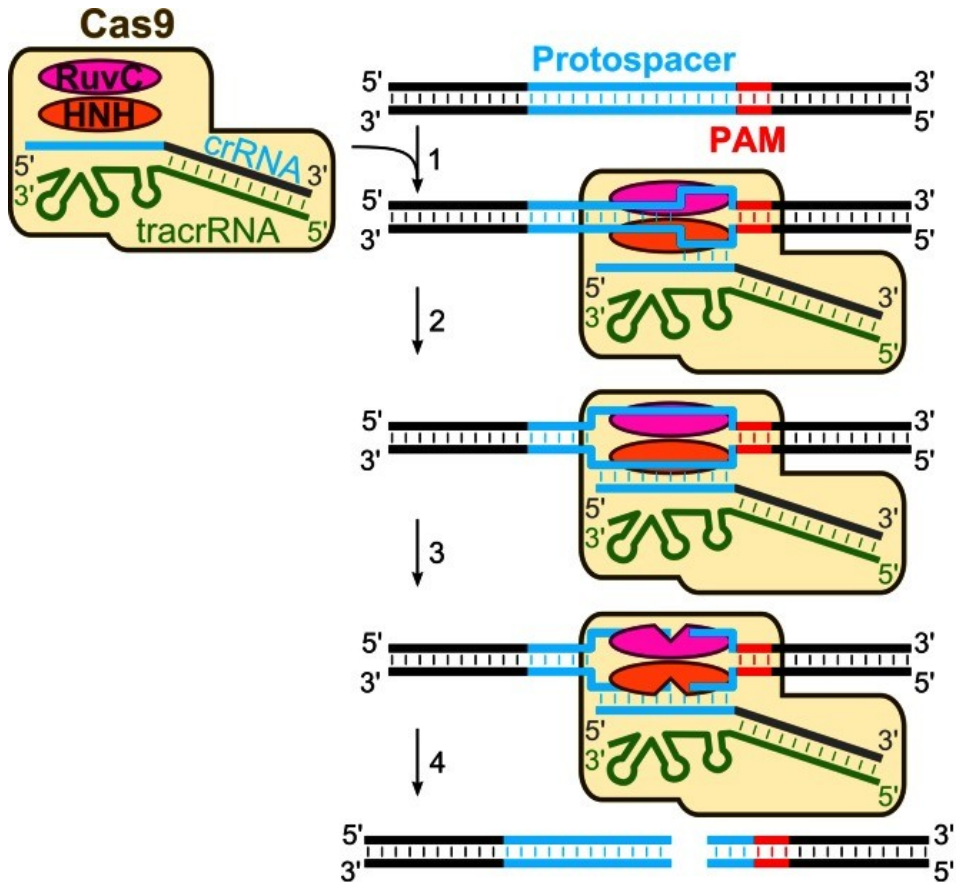


Figura 6 - Interferência de DNA nos sistemas CRISPR-Cas do tipoII. O complexo ternário Cas9-crRNA-tracrRNA procura no DNA a sequência protoespaçadora complementar e PAM: Uma vez que a correcta PAM e a curta primária sequência de hibridização são identificadas (1), os pares de bases do crRNA formam uma ansa R com a cadeia complementar (2). Uma vez que a ansa R se encontra formada, a proteína Cas9 cliva ambas as cadeias de DNA, marcada e não marcada utilizando o RuvC e os locais activos de HNC (3). A clivagem ocorre 3 nt antes de PAM, produzindo fragmentos DNA (4). Adaptado de (Gasiunas et al., 2014)

A importância do tracrRNA no processo de maturação de crRNA é demonstrada directamente nos vários sistemas do tipos II (Karvelis et al., 2013).

Os sistemas CRISPR-Cas associados aos mecanismos de imunidade adquirida foram caracterizados em dois organismos modelo. *S. pyogenes* e *S. thermophilus* DGCC7710.(Tabela 2)

Organismo	Sub-tipo	Compleco acrónimo	Comprimento da repetição	Sequências de Repetição	Comprimento de espaçadores	crRNA, nt	TracrRNA	PAM (5'-3')	Posição de clivagem	Composição do complexo
<i>Streptococcus thermophilus</i> DGCC7710/CRISPR1	II-A ^a		36	5'- GTTTTTGTACT CTCAAGAT TTAAGTAACTG TACAAC-3'	30	~42 ^b	+	NNAAGAW	3 nt upstream PAM	nd
<i>Streptococcus thermophilus</i> DGCC7710/CRISPR3	II-A	St-Cas9t	36	5'- GTTTTAGAGCT GTGTTGT TTCGAATGGTT CCAAAAC-3'	30	42	+	NGGNG	3 nt upstream PAM	Cas9-crRNA-tracrRNA
<i>Streptococcus pyogenes</i> SF 370	II-A	Sp-Cas9t	36	5'- GTTTTAGAGCT ATGCTGT TTTGAATGGTC CCAAAAC-3'	30	42	+	NGG	3 nt upstream PAM	Cas9-crRNA-tracrRNA

Tabela 2 - Mecanismos de imunidade adquirida mediada pelo sistema CRISPR tipo II relativamente a três organismos. Adaptado de (Gasiunas et al., 2014)

Relativamente ao *S. thermophilus* DGCC7710 este possui dois sistemas II-A, designados ST-CRISPR 1 e ST-CRISPR3, sendo ambos activos *in vivo*, isto é, capazes de incorporar novos espaçadores após a infecção por fagos e providenciar resistência para novas possíveis infecções (Horvath et al., 2008)

As sequências repetidas nos sistemas II-A caracterizados, como é possível ver na tabela 3, têm um comprimento de 36 nt e são parcialmente palindrómicas.

O processamento do pré-crRNA na presença de tracrRNA e da RNase III resulta num crRNA maduro de 42 nt, bastante mais pequeno do que os crRNA maduro do sistema tipo I (Tabela 1).

As duas principais diferenças entre os crRNA do sistema II-A e do sistema I são: 1- no tipo II, o crRNA não contém repetições na extremidade 5' e na extremidade 3' contém repetições com apenas 22 nt geradas pela clivagem pela RNase III dentro da região de repetições do duplex pré-crRNA – tracr RNA; 2- o fragmento de DNA espaçador (DNA exógeno) é mais pequeno (20 nt) devido à acção de RNases desconhecidas.

Consequentemente, os 20 nt do crRNA maduro ligam-se a 30 nt da sequência protoespaçadora presente no ácido nucleico invasor. Os nucleótidos protoespaçadores que faltam no crRNA não são relevantes para a imunidade, contudo diminuindo demasiado a sequência protoespaçadora do crRNA (para 19 nt ou menos) revoga a interferência do plasmídeo mediado pelo CRISPR.

3.1.3. CRISPR TIPO III

Os sistemas do tipo III são mais difundidos no grupo *archae* e estão presentes em *loci* de diversos CRISPR-Cas pertencentes a diferentes subtipos, como é o caso do *Pyrococcus furiosus* que contém três módulos de proteínas Cas – I-A, I-B e III-B; e sete matrizes CRISPR, o *archaeon S. solfataricus* que contém cinco módulos de proteínas Cas, duas delas pertencentes ao subtipo III-B e três ao subtipo I-A e seis matrizes CRISPR, respectivamente. (Gasiunas et al., 2014)

O sistema imune CRISPR do tipo III, não é contudo uma característica das espécies *archaea*. A bactéria *S. thermophilus* DGCC7710 tem quatro *loci* diferentes de CRISPR-Cas no cromossoma incluindo o III-B. Outras bactérias incluindo patogénicos importantes como micobactérias e estafilococos também possuem sistemas CRISPR tipo III.

Neste sistema actua a proteína Cas10 que inclui motivos específicos de polimerase. Múltiplas proteínas da família RAMP também se podem encontrar neste sistema tipo III, que estão envolvidas na ligação ao RNA dando origem a uma estrutura em forma de garfo

Nos sistemas tipo III-B de *S. solfataricus* e *P. furiosus* o complexo, formado pelo módulo Cas-RAMP (Cmr) e crRNA, reconhece e cliva RNA sintético *in vitro*. (Tabela 3) enquanto o sistema tipo III-A de *S. epidermidis* cliva DNA *in vivo* (Zhang et al., 2012). Por outro lado, o sistema III-B de *S. islandicus* interfere com DNA plasmídico por um mecanismo dependente da transcrição directa do protoespaçador em RNA (Deng et al 2013).

Organismo	Sub-tipo	Compleco acrónim o	Comprimento da repetição	Sequências de Repetição	Comprimento de espaçadores	crRNA, nt	Alvo	Posição de clivagem	Composição do complexo
<i>Staphylococcus epidermidis</i> P62A ^a	III-A		36	5'-GTTCTCGTCCCCTTTTCTT CGGGGTGGGTATCGATC C-3' ^b	34-35	37 or 43	dsDNA ^c	Nd	Nd
<i>Pyrococcus furiosus</i>	III-B	Pf-Cmr	30	5'-GTTCCAATAAGACTA CAAAGAATTGAAAG-3'	36-42	39 or 45	ssRNA	14 nt from 3' crRNA end	Cmr1, Cas10, Cmr3-6
<i>Sulfolobus solfataricus</i> P2	III-B	Ss-Cmr	24-25	5'-GATAATCTCTTATAG AATTGAAAG-3'	36-44	45-50	ssRNA (including crRNA)	At UA dinucleotide	Cmr1, Cas10, Cmr3-7

Tabela 3 - Mecanismos de imunidade adquirida mediada pelo sistema CRISPR tipo III relativamente a três organismos. Adaptado de (Gasiunas et al., 2014)

Contrariamente ao que sucede nos sistemas de tipo I, as repetições dos subtipos III-B não são palindrômicas. Contudo, o processamento do pré-crRNA nos sistemas tipo III é realizado pela endoribonuclease Cas6 presente no garfo, também utilizada nos sistemas do tipo I, embora o domínio que interage com a sequência não palindrômica seja diferente. O processamento do *P. furiosus* ocorre em dois passos: inicialmente a sua Cas6 corta o pré-RNA na região de repetição para se ligar um intermediário, posteriormente ocorre a degradação da extremidade 3' do intermediário por endonucleases desconhecidas produzindo-se duas populações de crRNA, uma de 45 e outra de 35 nt de comprimento em que ambas contêm uma sequência conservada originada da repetição clivada na extremidade 3'. Os crRNA maduros contêm 37 e 31 nt pertencentes à região espaçadora (Gasiunas et al., 2014)(Figura 7)

A sequência de repetição no subtipo III-A é palindrômica. O processamento de pre-crRNA *in vivo* necessita da Cas6, Csm4 e ou proteínas Cas10 e a clivagem ocorre na base do pente. O padrão de clivagem de pre-crRNA segue um mecanismo similar ao da endoribonuclease Cas6 do tipo I-E e I-F e a Se-Cas6 também está envolvida no processamento do pr-crRNA de *S. epidermidis*. A posterior maturação do intermediário

dá origem ao crRNA com 43 e 37 nt que necessita das proteínas Csm2, Csm3 e Csm5 para actuar (Wang, Preamplume, Terns, & Li, 2011).

O mecanismo de degradação de invasão do ácido nucleico foi estudado *in vitro* para os tipos III-B de *P. furiosus* e *S. solfataricus* e *in vivo* para o sistema III-A de *S. epidermidis*. Complexos efectores de *P. furiosus* e *S. solfataricus* (Pf-Cmr e Ss-Cmr) foram purificados bioquimicamente e caracterizados. Contrariamente a outros sistemas CRISPR-Cas que clivam o DNA, ambos os complexos efectores de tipo III-B clivam RNA.

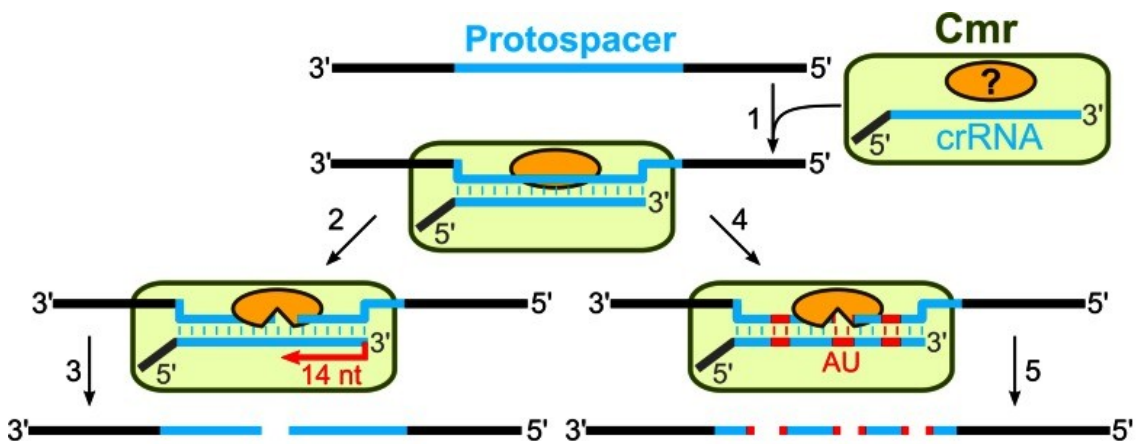


Figura 7- Interferência pelos sistemas CRISPR-Cas tipo III. No Caso do sistema III-B, o complexo Cmr procura pares de bases de RNA e crRNA com uma sequência protoespacçadora correspondente (1). Dois mecanismos de clivagem de RNA são propostos. O complexo Cmr de *P. furiosus* induz um único corte no RNA alvo a partir da extremidade 3' do crRNA (2) para produzir dois fragmentos (3) O complexo Cmr de *S. solfataricus* guiado por crRNA corta o RNA alvo em vários locais, especificamente nos dinucleótidos UA (4) em várias posições (5). Os componentes do complexo Cmr envolvidos na clivagem ainda não estão identificados. Adaptado de (Gasiunas et al., 2014)

3.2.MECANISMO DE ACÇÃO DO SISTEMA CRISPR

O sistema CRISPR-Cas medeia a imunidade aos elementos genéticos invasores através de um processo com três fases: adaptação, expressão e interferência. (Figura 8) (Makarova et al., 2012)

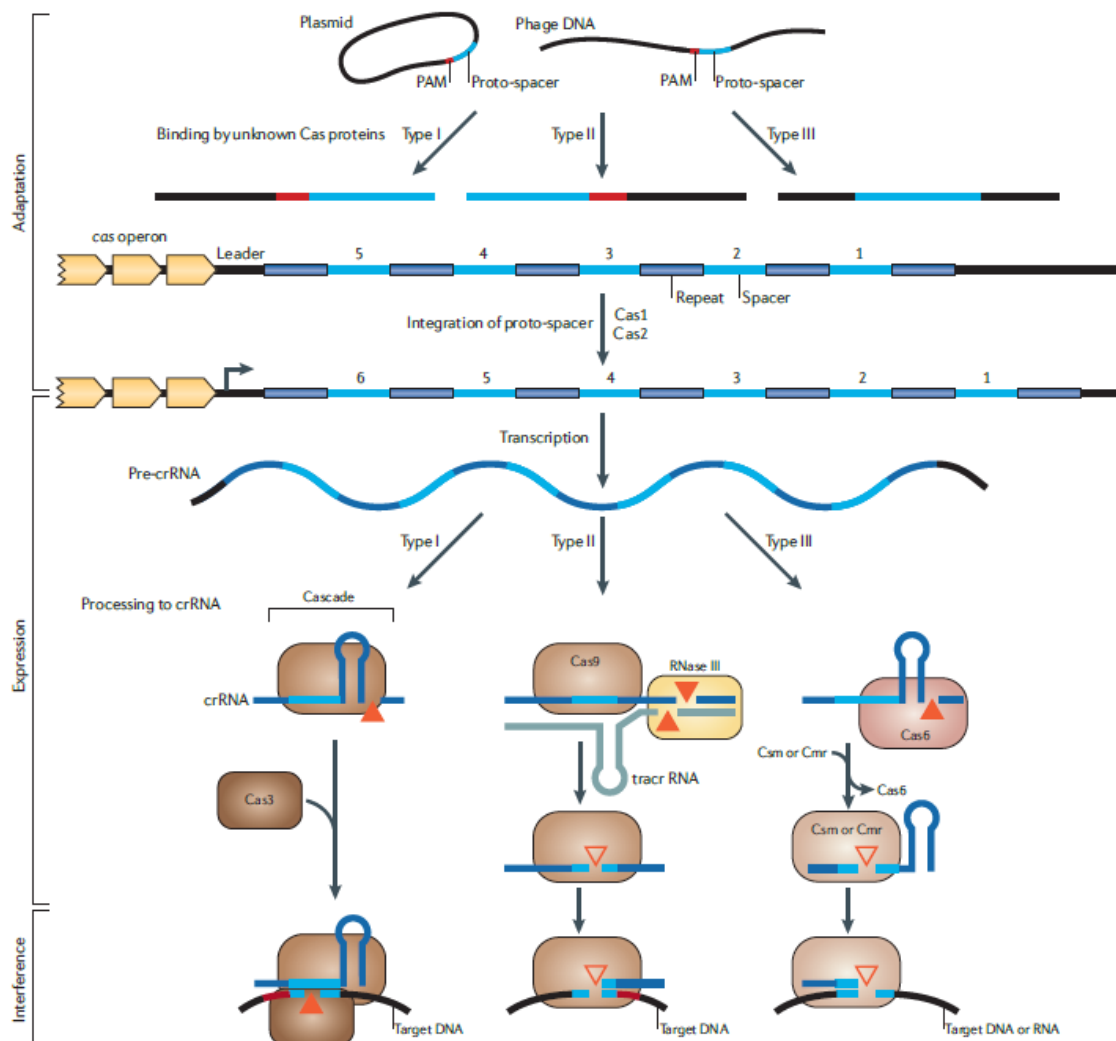


Figura 8 - Representação das três fases do sistema CRISPR-Cas : adaptação, expressão e interferência. No tipo I e tipo II dos sistemas CRISPR-Cas, mas não no tipo III, a seleção de protoespaçadores do ácido nucleico invasor depende de um protoespaçador adjacente ao motivo (PAM), mas esta etapa de reconhecimento ainda não é clara. Após o reconhecimento inicial, as proteínas Cas1 e Cas2 integram os protoespaçadores nos locus CRISPR. Adaptado de (Makarova et al., 2012)

Durante a fase de adaptação, porções de DNA exógeno de vírus ou plasmídeos são integradas nos *loci* CRISPR (Bhaya, Davison, & Barrangou, 2011).

Nesta fase, no local líder do *locus* do CRISPR o DNA exógeno é inserido como espaçador conferindo resistência ao invasor. Aquisições de vários espaçadores do mesmo fago ocorrem de forma menos frequente. Cada momento de integração é acompanhado de duplicação de uma repetição e por sua vez cria uma nova unidade de espaçador-repetição. A seleção de um precursor de espaçadores a partir do DNA invasor é determinada pelo reconhecimento de um motivo adjacente aos protoespaçadores (PAMs)(Makarova et al., 2012)

A segunda fase é a expressão, durante a qual é gerado um transcrito primário longo do *locus* do CRISPR (pré-CrRNA) posteriormente processado em pequenos crRNA. O passo de processamento é catalisado por endoribonucleases que ou operam como subunidades de um complexo maior (como o complexo CRISPR para defesa antiviral (Cascada) em *Escherichia coli*) ou como uma simples enzima (Cas6 no Caso do *archaeon Pyrococcus furiosus*). Relativamente ao complexo Cascada do tipo I do sistema CRISPR-Cas, o crRNA mantém-se associado ao complexo após a clivagem inicial pela endonuclease,

Na terceira e última fase chamada de interferência, o DNA ou RNA invasor é reconhecido como sequência protoespaçadora e clivado (Westra et al., 2010) Os crRNA correspondentes guiam os complexos de proteínas Cas, bem como o complexo Cascada de *Escherichia coli*, para as sequências de vírus ou plasmídeos a destruir.

4. APLICAÇÕES DO SISTEMA CRISPR-CAS

Devido ao polimorfismo dos genes *cas* e à correspondente diversidade das proteínas Cas, o sistema CRISPR-Cas permitiu que diversos investigadores em todo o mundo desenvolvem-se tecnologias com diversas aplicações em biologia, algumas das quais abaixo descritas, especialmente o sistema CRISPR-Cas9.

4.1. ENGENHARIA GENÉTICA

Em 1980 surge o conceito de edição genética, através de recombinação homóloga – utilizada para marcar genes específicos de células estacionárias embrionárias de ratos de forma a criar células *knock-in* (KI) e *knock-out* (KO) (Capecchi, 1989).

Com a necessidade da adaptação dos modelos à escala de patologias de mamíferos surge a tecnologia Cre-lox que permitiu controlar a expressão génica espacial e temporalmente (Utomo, Nikitin, & Lee, 1999).

Posteriormente são reconhecidas as nucleases *zinc finger* (ZFN) que consistem num sistema de fusão de domínios não específicos de proteínas de clivagem que permitem introduzir quebras em cadeias duplas de DNA (DSBs) (Gaj, Gershbach, & Barbas, 2013). Em 2009, desenvolvem-se as *transcription-like effector nucleases* (TALENs) – enzimas derivadas de proteínas efectoras TAL, secretadas por bactérias xanthomonas,

que permitem cortar segmentos específicos de DNA (Wright, Li, Yang, & Spalding, 2014).

No ano de 2000, Mojica *et al* identificam pela primeira vez o sistema de defesa CRISPR em procariotas, mas apenas em 2013 foi desenvolvido o sistema de CRISPR tipo II observado das bactérias *S. Thermophilus* e *S. Pyogenes*. Actualmente, o sistema CRISPR-Cas é utilizado como ferramenta de edição genética e prevê-se o seu desenvolvimento como meio de edição de genomas de mamíferos (Cong et al., 2013).

Esta evolução é visível na figura 9.

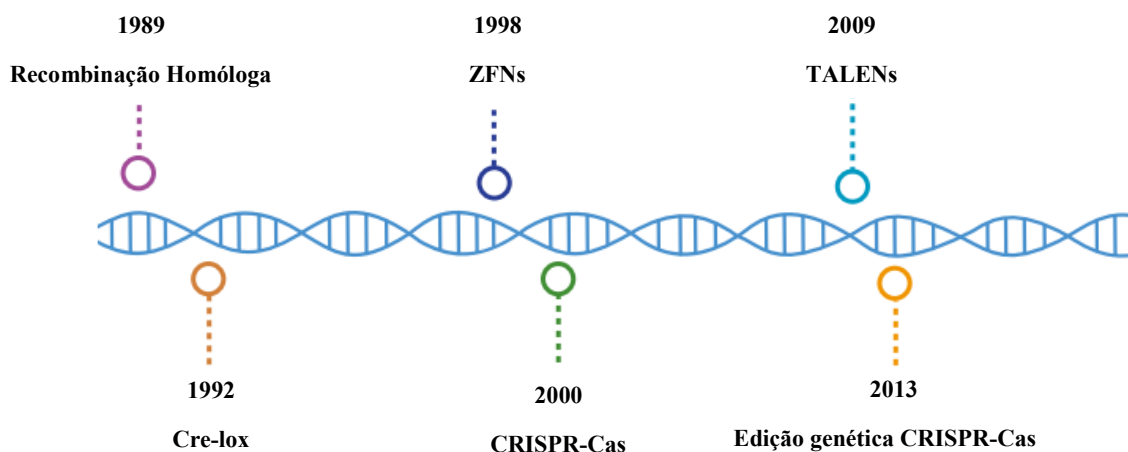


Figura 9 - Marcos na evolução de sistemas de manipulação genética até ao aparecimento do sistema CRISPR-Cas. Adaptado de (GenScript, n.d.)

A adaptação do sistema CRISPR para células de mamíferos revolucionou a edição genética, pois contrariamente às ZF e às nucleases TALE, o sistema CRISPR não necessita de engenharia proteica para marcação dos genes, necessitando apenas de porções de DNA que permitam codificar um gRNA e uma Cas9. Caso o KI seja alcançado, ocorre a recombinação homóloga, podendo também editar vários genes simultaneamente.

Na tabela abaixo podem observar-se as principais diferenças entre os sistemas CRISPR-Cas, TALENs e ZFNs.

	TALENs	ZFNs	CRISPR-Cas
Alvo	Proteína : DNA	Proteína : DNA	gRNA-Cas9 : DNA
Domínio	Proteínas que possuam domínios de ligação de DNA reconhecem sequências específicas de DNA através dos pares de base	Os motivos de ligação de DNA <i>zinc finger</i> em determinada configuração reconhecem segmentos de três pares de bases de DNA	O crRNA de 20 nt é fundido com um tracrRNA e uma endonuclease Cas9 que reconhece sequências específicas de pares de bases
	Desvantagem		Vantagens
	Necessitam de uma proteína para cada sequência de genes		Sistema funciona apenas com um vector gRNA-Cas9
	Menor eficiência de entrega		Permite a edição de vários genes

Tabela 4- Principais diferenças entre TALENs, ZFNs e CRISPR-Cas. Adaptado de (GenScript, n.d.)

4.2. O SISTEMA CRISPR-CAS-9

O sistema CRISPR-Cas 9 surge com a capacidade extraordinária de direccionar especificamente nucleases, o que permite não só a sua utilização a nível de engenharia genética bem como a nível biomédico.

4.2.1. TRATAMENTO DE DOENÇAS INFECCIOSAS

O Sistema CRISPR-Cas9, pode ser utilizado para o tratamento de doenças infecciosas através da erradicação de genomas patogénicos nos indivíduos afectados, como no caso do HIV (vírus da imunodeficiência adquirida), podendo mesmo prevenir uma re-infecção (Ebina, Misawa, Kanemura, & Koyanagi, 2013).

Estudos realizados recentemente, evidenciam a alta especificidade do complexo Cas9/sgRNAs para inactivar a expressão do DNA proviral de HIV e evitar a replicação do vírus em células latentes (Hu et al., 2014).

4.2.2. DOENÇAS GENÉTICAS

Através da injeção de mRNA de Cas9 e sgRNA em zigotos de ratinhos, o sistema CRISPR-Cas9 permite corrigir uma mutação genética responsável por uma distrofia muscular na linha germinativa dos mesmos ao mesmo tempo que previne o desenvolvimento da mesma patologia em ratinhos mutantes (Schwank et al., 2013).

5. VÍRUS DO PAPILOMA HUMANO

Os vírus do papiloma humano (HPV) pertencem à família *Papillomaviridae* e ao género do Papilomavírus infectam tecidos epiteliais de anfíbios, répteis, pássaros e mamíferos (Lima, Silva, & Rabenhorst, 2013) (Figura 10).

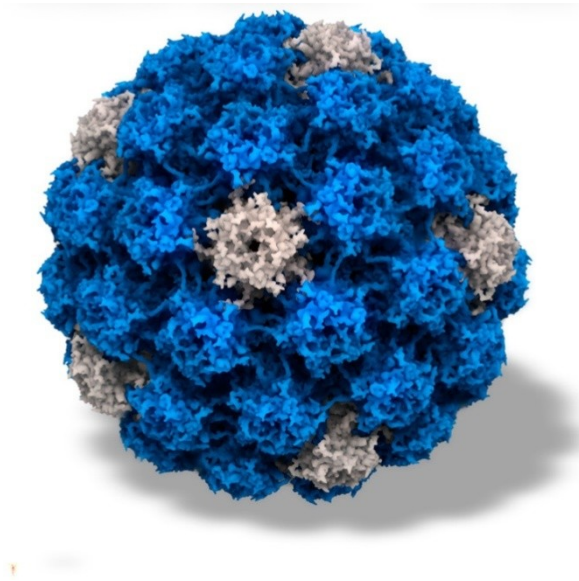


Figura 10- Representação 3D do vírus do Papiloma Humano (HPV) Adaptado de <http://3dciencia.com/blog/wp-content/uploads/2011/03/papiloma-virus-HPV-papillomavirus-what-you-should-know-about-viru.jpg>

Estes vírus não são envolvidos por membrana, têm simetria icosaédrica e um genoma de DNA de cadeia dupla, com cerca de 8000 pares de bases (Ferraz, Beatriz, Santos, & Discacciati, 2012)

Até à data existem mais de 150 tipos de HPV completamente sequenciados (Figura 11) e cerca de 60 papiloma vírus animais (Bernard et al., 2010)

Os Gamapapilomavírus por sua vez abrangem cinco espécies causadoras de lesões cutâneas, HPVs 4, 48, 50, 60, 88, 65, 95.

Os *Mupapilomavírus* englobam os HPVs 1 e 63 sendo o primeiro um dos mais estudados no que diz respeito a verrugas vulgares e palmares.

Os *Nupapilomavírus* englobam apenas uma espécie, o HPV 41 (Pereira, Porro, Francisco, & Tomimori, 2011)(Doorbar, 2005)

5.1. ESTRUTURA DO GENOMA E ORGANIZAÇÃO

Relativamente ao seu genoma (Figura 12), este pode ser dividido em três regiões principais, sendo que uma delas não codifica *long control region* (LCR) e duas delas são regiões codificantes a região precoce (E) e a região tardia (L).

O LCR consiste numa região não codificante com 400 a 900 pb que assume funções regulatórias e contém a origem de replicação, promotores precoces onde se ligam proteínas reguladoras e que se encontra entre as regiões precoce e tardia, representando cerca de 15% do genoma viral

As duas regiões codificantes, são compostas por *open reading frames* (ORFs) – sequências de DNA compreendidas entre um codão de iniciação e um codão stop; e são região precoce, E, que contém seis ou mais ORFs (E1 a E6), que se encontram envolvidos na transcrição génica, replicação, transformação e adaptação celular. A região tardia, L, que contém duas ORFs codificadores de proteínas tardias, L1 e L2 procedidas de sequências poli-adenilação (poli-A)(Mu et al., 2004)

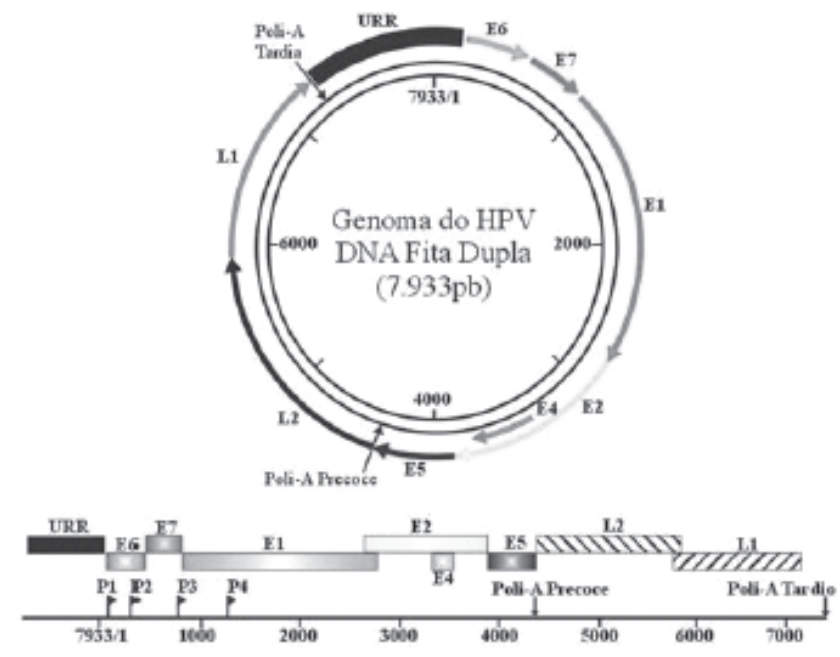


Figura 12 - O genoma viral é dividido em três regiões: uma região não codificante denominada LCR; e duas codificantes, designadas região precoce e tardia. Os ORFs da região precoce são apresentados através de setas ou barras, enquanto os da região tardia através de setas cinza ou preta e barras listradas. Podem também verificar-se os quatro promotores (P1, P2, P3 e P4) e os dois locais de poli-adenilação Adaptado (Lima et al., 2013)

5.2.EPIEMIOLOGIA MUNDIAL DO HPV

Segundo a Organização Mundial de Saúde (OMS) o vírus do papiloma humano (HPV), causa cancro cervical e este é o quarto cancro mais comum em mulheres, causando cerca de 266,000 mortes 528,000 novos casos em 2012, sendo que a maioria das infecções ocorre em países menos desenvolvidos (85%) representando cerca de 12% dos cancros femininos. (OMS, 2015), considerando-se a segunda maior causa de mortes por cancro em indivíduos do sexo feminino (Bosch, Lorincz, Muñoz, Meijer, & Shah, 2002).

Estudos efectuados no ano de 2012, em mulheres cujo exame citológico se apresenta normal indicam que a prevalência é cerca de 11 a 12%, existindo uma variação regional cuja maior prevalência se encontra em África e nas Caraíbas (superando os 30%) e a menor na Ásia (inferior a 2%), surgindo a Europa do Norte como uma excepção ao padrão com cerca de 12%, como é visível na figura 13 (Forman et al., 2012).

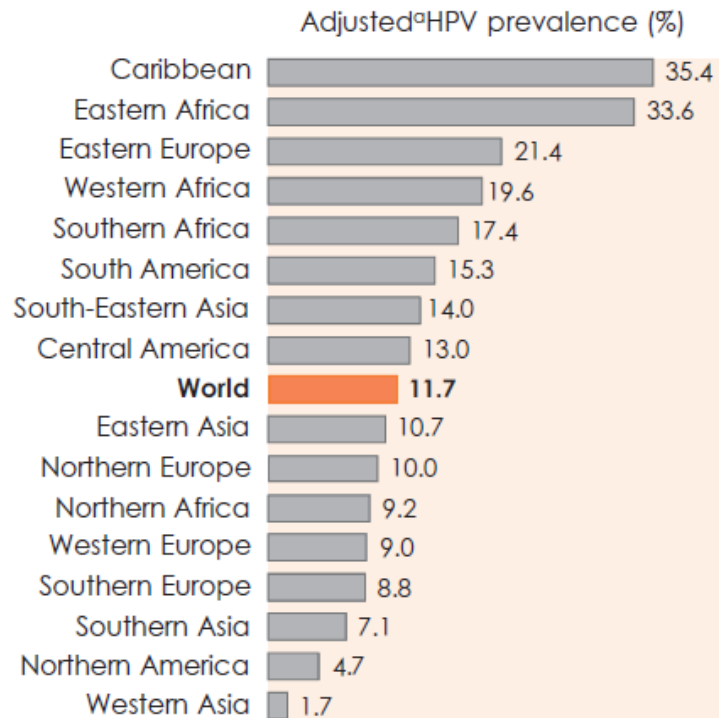


Figura 13 - Prevalência de HPV em mulheres cuja citologia se apresenta normal. Adaptado de (Forman et al., 2012)

Os dados acima referidos são importantes, uma vez que mesmo das mulheres saudáveis na figura abaixo pode verificar-se o número preocupante de mulheres que irão desenvolver cancro cervical, estando no entanto associados a factores de risco como doenças auto-imunes, tabaco, mães de idade jovem e um elevado número de nascimentos(Crow, 2012).

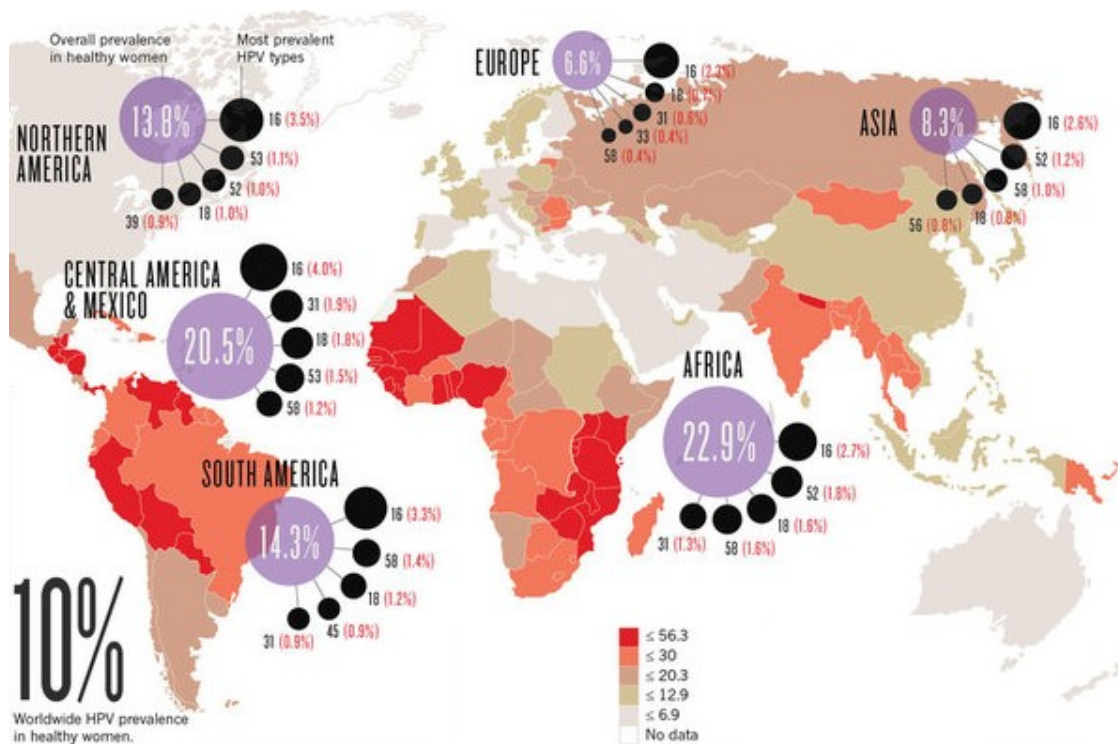


Figura 14 - Prevalência do HPV no mundo, adaptado de (Crow, 2012)

O HPV é por estes números considerado o vírus mais prevalente envolvido em doenças sexualmente transmissíveis no mundo e, constitui assim um importante desafio no que diz respeito à saúde pública, representando o segundo carcinogéneo mais importante a seguir ao tabaco, correspondendo a 5% dos cancros em geral (“WHO,” n.d.) (Figura 14)

5.3.EPIDEMIOLOGIA NACIONAL DO HPV

Em Portugal, não abundam estudos de prevalência em relação aos diferentes tipos de HPV, no entanto recentemente numa amostra de 2326 mulheres, 19,4% eram portadoras de infecção. No grupo etário mais jovem, dos 20 aos 24 anos, a prevalência foi a mais elevada, apresentando 28,8%, diminuindo com o aumento da idade. Na faixa etária de 49 o valor diminui cerca de 10% e dos 60 aos 65 os valores são de 6%.(Pista, Oliveira, Paixão, Real, & Cunha, 2011).

5.4. INFECCÃO VIRAL POR HPV

As partículas de HPV desde a superfície celular até ao citoplasma e núcleo sofrem vários passos até ao local onde irá ocorrer a replicação, sendo a cápside um dos factores mais importantes no que diz respeito à infecção viral (Lowy & Schiller, 2006).

A infecção encontra-se dependente de uma diferenciação celular por parte dos queratinócitos epiteliais. Inicialmente, as células epiteliais da camada basal, por micro-abrasão, são infectadas por viriões de HPV através de interacções entre proteoglicanos (HSPGs) - *heparan sulfate proteoglycans*; e a proteína da cápside L1 (Sapp & Bienkowska-Haba, 2009).

Desta forma, na camada basal, o genoma é mantido como epissomas (plasmídeos com capacidade de se integrarem no cromossoma), facilitando a entrada do vírus cujo DNA depois de duplicado pode atingir cerca de 50 cópias por célula.

Posteriormente, inicia-se replicação prolífica do vírus, devido à diferenciação das células infectadas, levando o genoma viral a multiplicar-se até 1000 cópias e à formação de viriões devido à expressão das proteínas L1 e L2, capazes de provocar nova infecção.

Uma vez que os epissomas dependem da fase S de replicação, os vírus do HPV através das proteínas E6 e E7 desenvolveram a capacidade de poderem introduzir-se novamente na fase S de forma a manterem o estado de diferenciação das células infectadas através da anulação do p53 e do Rb (supressor tumoral retinoblastoma) por E6 e E7.

Ambas as oncoproteínas virais permitem a replicação do DNA viral, permitindo uma nova entrada na fase S do ciclo celular de queratinócitos diferenciados (E7) bem como evitam a paragem do ciclo dependente do p53 e a resposta apoptótica (E6). Este processo é visível na figura 15 (D'Abramo & Archambault, 2011).

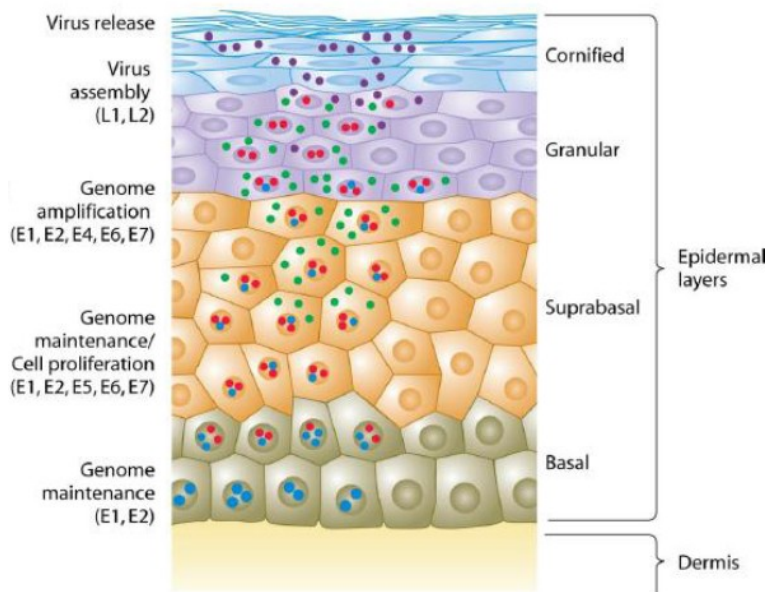


Figura 15 - Processo de infecção por HPV adaptado de (Lazarczyk, Cassonnet, Pons, Jacob, & Favre, 2009)

5.5. TIPOS DE HPV

Até aos dias de hoje, já foram identificados mais de 120 tipos de vírus diferentes de HPV, baseando esta identificação no seu genoma, sendo por sua vez classificados, de acordo com as frequências de diagnóstico viral associados a cancro, como baixo, intermédio e alto risco ou oncogénicos.

Relativamente aos tipos de alto risco, o mais frequente é o HPV-16, que é responsável por cerca de 70% dos casos de carcinomas ligados à infecção viral. Os tipos 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 69, 73 e 82 (Waggoner, 2003) são classificados como oncogénicos. Quanto aos tipos 6, 11, 40, 42, 43, 44, 54, 61, 70, 72, 81 e CP6108 são considerados de baixo risco (Munoz N, Bosch FX, de Sanjose S, Herrero R, Castetellaasague X, 2003).

O HPV-6 e HPV-11 são os mais frequentes no que diz respeito aos tipos de baixo risco e causam, na sua maioria, lesões benignas (Mu et al., 2004).

5.6. PROTEÍNAS DO HPV

As proteínas L1 e L2 são as principais proteínas da estrutura da cápside. A primeira tem como principal função mediar a ligação dos vírus à superfície celular, enquanto a segunda é responsável pela formação da cápside no genoma viral, a organizando o virião através da sua ligação ao ADN viral (Fay, IV, Roden, & Moroianu, 2004).

A proteína E1 encontra-se principalmente envolvida na replicação de DNA, tem funções de helicase de forma a separar as cadeias de DNA viral e forma também um complexo de replicação com a proteína E2 (Conger, Liu, Kuo, Chow, & Wang, 1999).

A proteína E2 encontra-se envolvida na regulação transcricional de DNA, sendo a sua principal função a recrutar de genomas virais, P97 e P105, que se encontram na região longa de controlo (LCR) e codifica os genes precoces e tardios, respectivamente (Longworth & Laimins, 2004) sendo responsável pela inibição da apoptose, e potencial transformação do HPV.

A proteína E4 é uma proteína citoplasmática que perturba a estrutura da queratina e influencia a formação de HPV-1. É responsável pela regulação do ciclo celular e pode aumentar a espessura da camada espinhosa da epiderme. (Doorbar, 2005). Estas são visíveis em momentos de infecção geralmente aquando da iniciação da amplificação do genoma (Gnanamony, Peedicayil, & Abraham, 2007).

A proteína E5, encontra-se envolvida na transformação e participa na replicação de DNA viral, permitindo que a célula infectada seja reconhecida pelo sistema imunitário.

As proteínas E6 e E7 constituem proteínas de alto risco mediadoras da carcinogénese devido ao seu mecanismo de interacção com diversos alvos celulares. O gene E6 é um dos primeiros genes a ser expressado durante a infecção por HPV (Longworth & Laimins, 2004).

As proteínas E6 têm aproximadamente 150 aminoácidos na sua constituição e contêm quatro motivos C-x-x-C que formam dois domínios de ligação ao zinco. Uma das suas principais funções consiste na ligação ao supressor de tumores p53 –que consiste numa proteína de ligação específica ao DNA expresso em resposta a agentes que danificam o ADN ou induzem a marcação do mesmo, resultando na paragem do ciclo celular ou em apoptose (Wu & Levine, 1994).

De forma a degradar o p53, as proteínas E6 ligam-se a uma ligase celular de ubiquitina formando um complexo- E6-AP. A degradação do p53 pela E6 resulta na omissão dos sinais de paragem de crescimento no G1/S e G/M, que leva à aquisição de inúmeras alterações genéticas que contribuem para a manifestação maligna (Fehrmann & Laimins, 2003).

As proteínas E6 têm também como função activar a telomerase – enzima com múltiplas subunidades responsáveis pela replicação de DNA dos telómeros na terminação dos cromossomas; sendo a activação da telomerase um processo crítico no que diz respeito à transformação celular. A maioria das células tumorais mantem o comprimento do telómero e actividade da telomerase (Bryan et al., 1997; Elenbaas et al., 2001).

O mecanismo preciso da activação da telomerase pela E6, continua no entanto desconhecido, alguns estudos porém demonstram que esta é capaz de regular os níveis do estado estacionário do hTERT endógeno nos queratinócitos através da activação do seu promotor.

Outras funções de E6 independentes de p53 podem contribuir para as suas actividades em transformação. Uma delas é a interacção das proteínas E6 com proteínas com vários domínios PDZ – consistem em longas unidades de interacção proteína-proteína de com cerca de 90 aminoácidos e são mais frequentemente encontrado no interior das proteínas celulares em áreas de contacto como junções sinápticas em neurónios e *tight junctions*, em células epiteliais (Fanning & Anderson, 1999) com proteínas resultando na sua degradação mediada por ubiquitina.

Várias proteínas virais envolvidas na oncogénese com adenovirus E4-ORF1, HTLV-1Tax, bem como proteínas E6 de HPV de alto risco possuem motivos de ligação de domínio PDZ na sua terminação carboxílica (Lee et al., 1997). A mutação deste domínio de ligação PDZ em proteínas E6 de alto risco tem sido utilizada para anular as suas funções de transformação (Howie, Katzenellenbogen, & Galloway, 2009).

Tanto a proteína E7 como a E6 resultam na maior transformação de HPV(Fehrmann & Laimins, 2003).

A proteína E7 das espécies de HPV de alto risco, sozinha é capaz de imortalizar queratinócitos e é composta por cerca de 100 aminoácidos cuja função primária consiste na associação com a família de repressores de tumores de retinoblastomas (Rb) de forma a facilitar a progressão para a fase S do ciclo de replicação. Através da ligação desta a E7 vai promover a progressão do ciclo celular em

células epiteliais diferenciadas favorecendo a produção de genes de HPV(Fehrmann & Laimins, 2003)

Já a proteína E7 das espécies de HPV de baixo risco do tipo HPV-1 é capaz de se ligar à Rb, falhando no entanto na sua marcação para degradação(Munger et al., 2001).

Na tabela 5 é possível ver as funções das proteínas virais de forma resumida.

Proteína	Função
E1	Helicase. Essencial para a replicação viral e controle da transcrição genética
E2	Replicação e expressão da mesma; regula a transcrição e replicação viral; controla a região de expressão precoce
E4	Interação com proteínas do citoesqueleto o que implica envolvimento na liberação de virões nas células do hospedeiro
E5	Estimulação de crescimento por interação com factores de crescimento; Transformação celular
E6	Transformação celular. Liga-se e degrada a p53; induz instabilidade genómica e activa a telomerase
E7	Transformação celular. Liga-se e inactiva a pRb, liberando o factor de transcrição E2F. Induz instabilidade genómica
L1	Compõem a cápside viral; papel na infecção e organização do ADN viral
L2	

Tabela 5 - Funções das proteínas do HPV, adaptado de Boulet et al, 2007

6. PATOGÉNESE

O HPV é transmitido pelo contacto directo ou indirecto com a lesão do individuo afectado, podendo esta lesão ser epidérmica, da mucosa, por contacto sexual ou durante o parto, tendo como principal alvo o epitélio escamoso da pele –verrugas; e mucosas – genital, oral e conjuntivas. Na tabela 6 abaixo, pode verificar-se a relação existente entre o tipo de HPV e a doença associada (Souto, Falhari, & Cruz, 2005)

Tropismo	Doença	Tipo de HPV
Baixo Risco	Verruga plantar	1, 2, 4, 63
	Verruga comum	1, 2, 3, 4, 7, 10, 26, 27, 28, 29, 41, 57, 65 e 77
	Verruga vulgar	3, 10, 26, 27, 28, 38, 41, 49, 75 e 76
	Lesões cutâneas	6, 11, 16, 30, 33, 3, 37, 38, 41, 48, 60, 72 e 73
	Epidermodisplasia verruciformis	2, 3, 5, 8, 9, 10, 12, 14, 15, 17, 19, 20, 21, 22, 23, 24, 25, 36, 37, 38, 47 e 50
	Papilomatose respiratória recorrente	6 e 11
	Papiloma/Carcinoma conjuntivo	6, 11 e 16
	Verrugas genitais	6, 11, 30, 42, 43, 45, 51, 54, 55 e 70
	Neoplasias Intraepitelial cervical (NIC)	
Mucosotrópico (Alto Risco)	→ Não específico	30, 34, 39, 40, 53, 57, 59, 61, 62, 64, 66, 67, 68 e 69
	→ Baixo Risco	6, 11, 16, 18, 31, 33, 35, 42, 43, 44, 45, 51, 52 e 74
	→ Alto Risco	6, 11, 16, 18, 31, 33, 34, 35, 39, 42, 44, 45, 51, 52, 56, 58 e 66
	Carcinoma cervical	16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 66, 68 e 70

Tabela 6 - Relação entre o tipo de HPV e a doença associada. Adaptado de (Souto et al., 2005)

7. CUSTOS ECONÓMICOS E IMPACTO NA QUALIDADE DE VIDA

A infecção por HPV, sendo uma das infecções mais comuns no tracto genital feminino, tem repercussões a nível da qualidade de vida da doente mas também influencia os custos com a saúde e consequentemente os recursos utilizados, especialmente por afectar mulheres jovens e em idade reprodutiva.(Navarro, Martín, Carlos, Ferreira, & Neto, 2013)

8. APLICÇÕES DO SISTEMA CRISPR-CAS EM VÍRUS HUMANOS

Como se sabe os vírus representam um enorme desafio devido ao seu ciclo de vida e à forma como utilizam proteínas celulares, dificultando a identificação de alvos terapêuticos quando comparados com bactérias que possuem um metabolismo celular específico (White, Hu, & Khalili, 2015)

O Sistema CRISPR associado à proteína Cas9 (CRISPR-Cas9) é simples e de fácil utilização, podendo ser utilizado para diferentes alvos de forma flexível (Ran et al., 2013)

Este sistema de imunidade adaptativa, tem sido desenvolvido como uma ferramenta de engenharia genética através da utilização de um pequeno RNA guia (gRNA) para a degradação directa de ácidos nucleicos específicos. Podem distinguir-se dois componentes distintos no sistema CRISPR-Cas9: um gRNA e uma endonuclease –Cas9. Quando ambos os componentes são expressos nas células a sequência de DNA alvo pode ser modificada ou interrompida como se verifica na figura 16 (White et al., 2015)

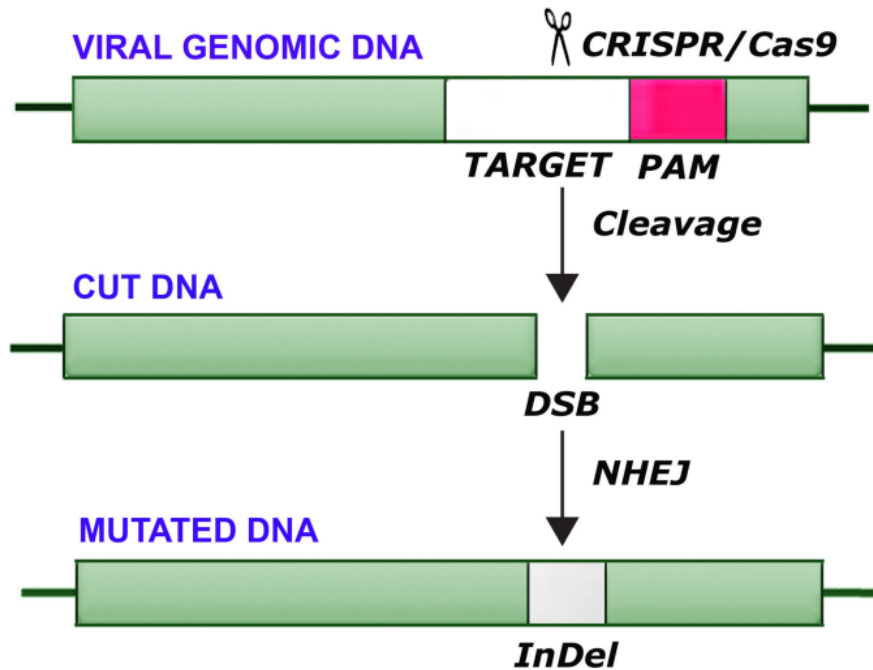


Figura 16- Representação da mutagênese do DNA viral pela CRISPR-Cas9. Adaptado de (White et al., 2015)

O gRNA com a sua sequência de cerca de 20pb recruta o complexo gRNA/Cas9 para o alvo através da complementaridade de bases de Watson-Crick. Para que a ligação de Cas9 ao alvo seja alcançada é também necessária uma sequência PAM após a sequência alvo.

Desta forma, a endonuclease Cas9 corta ambas as cadeias de DNA causando uma DSB que pode ser reparada através de uma via de reparação de DNA, NHEJ. Assim, pode resultar na criação de inserções ou deleções no DSB levando a alterações de frames ou à criação de códons stop prematuros resultando na quebra dos ORFs do gene alvo.

Uma vez que a Cas9 desempenha funções de uma endonuclease, necessita apenas de um gRNA que pode ser sintetizado de forma química, copiado *in vitro* ou expresso para que se torne uma ferramenta específica de marcação genética (White et al., 2015).

9. O SISTEMA CRISPR-CAS 9 E HPV16 E HPV18

Os HPVs 16 e 18 são responsáveis pela maioria dos câncros associados ao Vírus do Papiloma Humano (White, Pagano, & Khalili, 2014).

As propriedades oncogénicas os tipos de HPV de alto risco incluem as proteínas virais E6 e E7, sendo que a maior proteína supressora de tumores, a p53 é um alvo da E6 enquanto a proteína viral E7 complexa com proteína da família do retinoblastoma cuja função é regular o ciclo celular (pRb, p107 e p130), resultando na fosforilação e libertação dos factores de transcrição que promovem o ciclo celular.

Desta forma, os genes E6 e E7 são alvos para a intervenção no que diz respeito a doenças malignas associadas ao HPV (White et al., 2015).

Na tabela 7, estão resumidas alguns dos estudos em que foi utilizado o sistema CRISPR-Cas e seus efeitos.

Vírus	Alvo	Células	Resultados
HPV18	E6 ou E7	HeLA	Indução do p53 ou paragem do ciclo celular do Rb
HPV16	E6 ou E7	SiHa	
HPV16	E7	SiHa e CaSki	Apoptose e inibição de crescimento
HPV16	E6, E7 ou promotor E6/E7	SiHA	Indução do p53 e p21, inibição da proliferação celular e carcinogénese

Tabela 7 - Aplicação do Sistema CRISPR-Cas9 a HPV16 e HPV18. Adaptado de (White et al., 2015)

9.1. INTERRUPTÃO DO GENE E6 DO HPV-16 ATRAVÉS DO SISTEMA CRISPR-CAS

O virão do HPV-16 possui aproximadamente 55 nm de diâmetro e consiste numa cápside icosaédrica composta por duas proteínas estruturais, a L1 e L2, como referido anteriormente. (Doorbar et al., 2012) (Figura 17).

Durante o processo de carcinogénese os oncogenes E6 e E7 promovem e mantêm o fenótipo maligno das células cancerígenas (Munoz N, Bosch FX, de Sanjose S, Herrero R, Castetellaasague X, 2003).

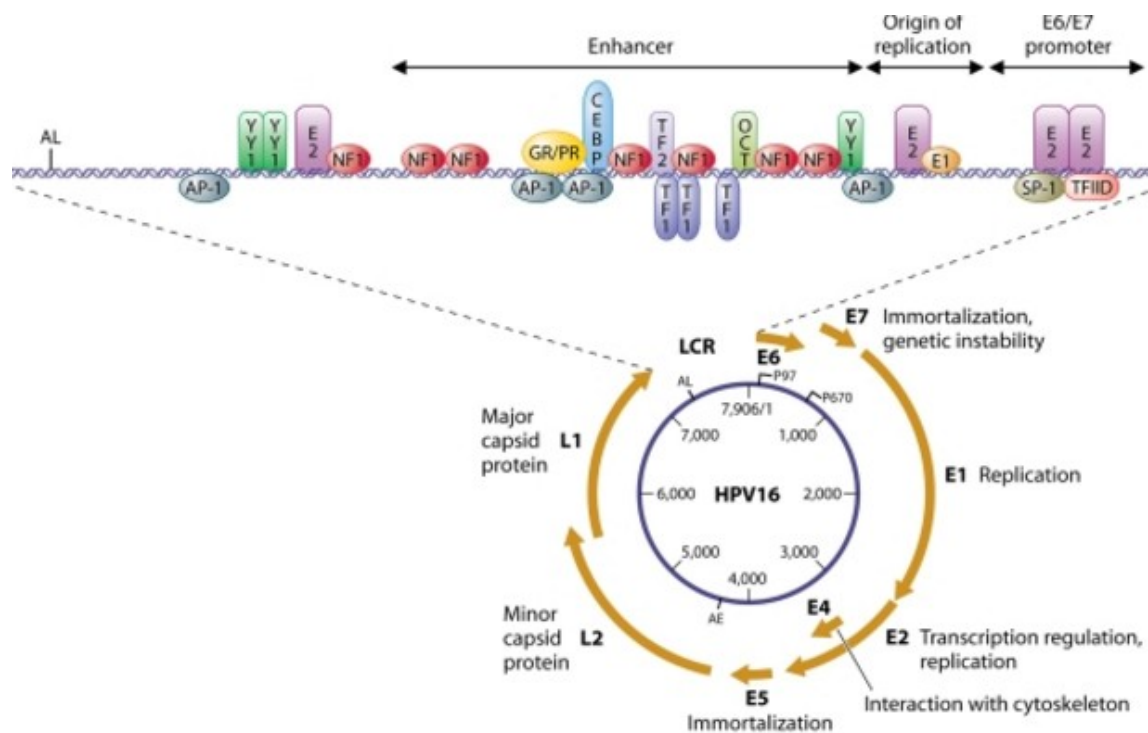


Figura 17 - Organização genética do HPV-16. Adaptado de (Lazarczyk et al., 2009)

Através da enzima Cas9 do sistema CRISPR-Cas e um RNA orientador (RNAg) específico para uma sequência de DNA é possível induzir quebras na cadeia dupla de uma determinada sequência de DNA complementar à sequência de gRNA. (Figura 18).

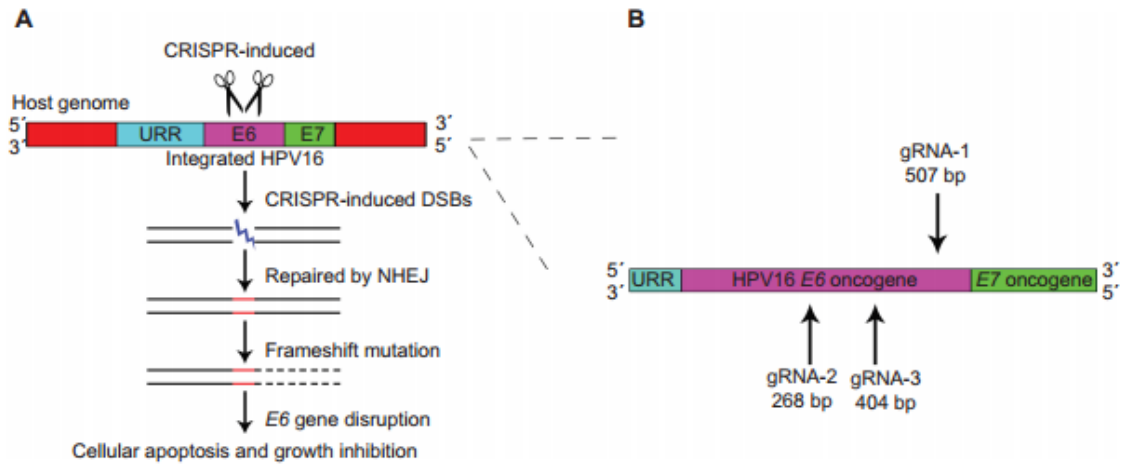


Figura 18 - Representação da edição do gene E6 através do sistema CRISPR-Cas. Em A verifica-se a quebra do gene E6 (DSBs) e em B os gRNAs marcam o gene E6. Adaptado de (Yu et al., 2014)

Estas quebras induzidas, são reparadas por uma via mutagénica terminal não homóloga (NHEJ) que leva à interrupção do gene marcado (E6). Esta quebra pode levar a apoptose e à inibição do crescimento de células infectadas com HPV-16. No estudo de Yu *et al*, sintetizaram-se três gRNA específicos para o gene E6 do HPV16, de forma a garantir que se obteriam DSB em locais específicos com a utilização de células humanas embrionárias de rins (HEK293) infectadas com HPV. Ao ocorrer a quebra a recombinação homóloga conduz à formação de um gene *luciferase*.

Na figura 19 é possível verificar em A que o plasmídeo utilizado como controlo de citotoxicidade induzida pelo CRISPR quando comparado com o controlo negativo apenas induz infecção com o plasmídeo Cas9, indicando também que os complexos gRNA-2/Cas9 e gRNA-3/Cas9 demonstraram maior actividade relativamente ao grupo gRNA-1/Cas9. Relativamente à citotoxicidade da *Renilla luciferase* utilizada em B verifica-se que esta não obteve alteração de sinal.(Yu et al., 2014)

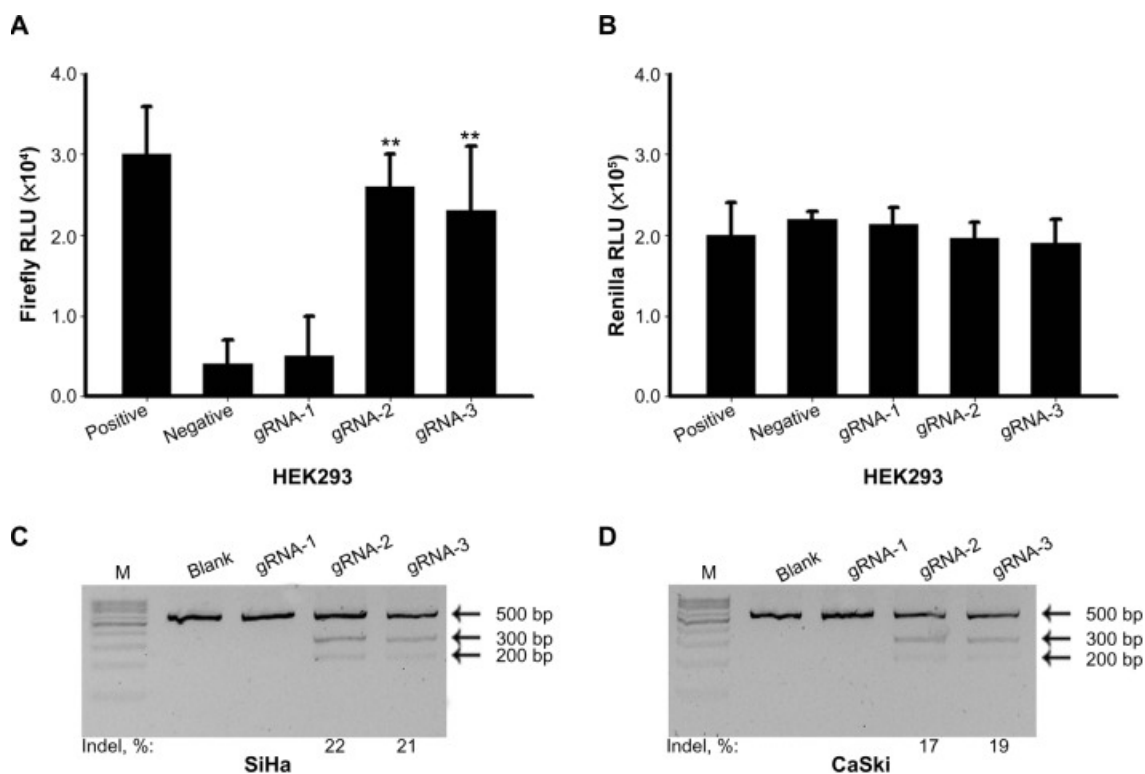


Figura 19 - Clivagem do DNA do gene E6 do HPV 16 pelo complexo CRISPR gRNA/Cas9

Na análise efectuada para confirmação da actividade dos três gRNAs, cada um destes foi introduzido em células SiHA e CaSki. O DNA das células infectadas de gRNA/Cas9 foi previamente amplificado por PCR e estas tratadas com T7E1 (T7 endonuclease I). Na figura 19 C e D verifica-se que o sistema CRISPR-Cas induziu mutações no HPV 16, permitindo concluir que o DSB seria possível no oncogene E6 (Yu et al., 2014).

Para determinar os efeitos do gRNA E6 no que diz respeito a apoptose, na figura 20 pode verificar-se que tanto em A como em B as células de SiHA e CaSki infectadas com HPV 16 quando comparadas com os controlos demonstraram níveis elevados de apoptose quando tratadas com gRNA-2/Cas9 e gRNA-3/Cas9 enquanto as células C33A HPV-negativas e HEK293, recebendo o mesmo gRNA/Cas9 não apresentaram sinais de apoptose (Yu et al., 2014).

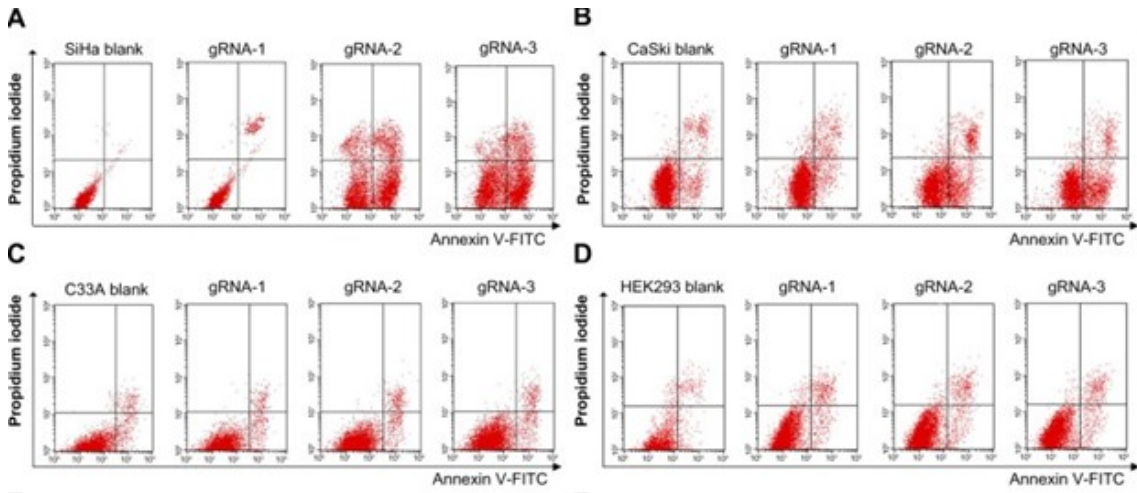


Figura 20 - Apoptose induzida pelo sistema CRISPR-Cas. SiHa (A), CaSki (B), C33A(C) e HEK293 (D). Adaptado de (Yu et al., 2014)

De forma a compreender os efeitos nos níveis proteicos de E6 e p53, utilizou-se uma análise de Western Blot onde a oncoproteína E6 leva a degradação proteossomal do supressor de tumor p53 permitindo estimar que a função da p53 pode ser recuperada, levando a apoptose e inibição de crescimento em células de HPV16, podendo também a expressão da p53 ser aumentada e contribuir para a reversão do fenótipo maligno em células cervicais cancerígenas (Figura 21)

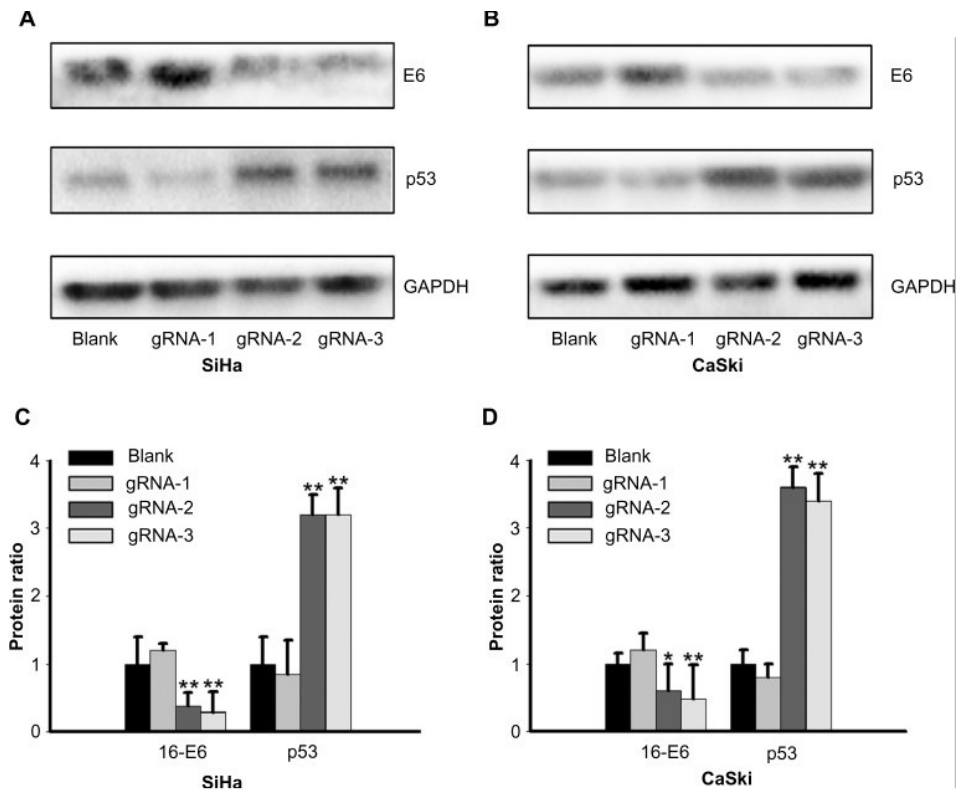


Figura 21 - Análise de Western Blot e análise por ImageJ. Adaptado de (Yu et al., 2014)

Comparando o sistema CRISPR-Cas com o siRNA, este último apenas tem efeito durante alguns dias, enquanto a mutação induzida pelo sistema CRISPR-Cas é permanente e como anteriormente descrito uma das capacidades e vantagens do sistema é que pode ser passado de geração em geração. Para além disso o sistema CRISPR é de mais fácil utilização e neste caso específico permite afirmar que esta técnica pode ser utilizada como forma de combater o cancro cervical causado pelo HPV.

No entanto, uma das preocupações relativas à utilização técnica é a sua especificidade para a sequência de DNA do HPV, isto pois embora induza apoptose nas células SiHa e CaSki o mesmo não se verifica nas células C33A e HEK293, como se constata(Yu et al., 2014).

10. CONCLUSÃO

O sistema CRISPR-Cas tem vindo a ganhar destaque na comunidade científica

Desde a sua descoberta e do seu modo de funcionamento simples, que compreende apenas três elementos: uma enzima – a Cas9; capaz de clivar as duas cadeias de DNA de forma específica, uma molécula de RNA, que funciona como guia, ligando-se ao local específico onde tem de ocorrer o corte pela enzima e uma porção de DNA, iniciando-se assim todo um processo de elevado nível de complexidade na sua simplicidade.

Ocorre assim uma recombinação e posterior integração da sequência. Uma vez que o gene considerado saudável não possui a sequência mutada, o gRNA não o reconhece permitindo que ocorra uma selecção apenas das células saudáveis, conferindo assim ao sistema a capacidade imunitária.

O Vírus do Papiloma Humano, tem ganho terreno nas últimas décadas, embora já existam meios de prevenção do mesmo, os números e estatísticas, revelam que continua a ser um problema de saúde pública e a sua estreita ligação ao desenvolvimento de carcinomas leva a uma certa preocupação.

Estudos efectuados, revelam já a sua importância no que diz respeito a alterações genéticas, devido à capacidade das alterações efectuadas poderem ser transmitidas de geração em geração.

O sistema CRISPR-Cas, e as descobertas que lhe estão associadas juntamente com o alucinante ritmo de desenvolvimento nas áreas da biologia e medicina permitem afirmar que é de forma indiscutível uma nova forma quase que terapêutica relativamente a diversas patologias e acima de tudo cria uma esperança na resolução das mesmas, conduzindo de certeza com a sua evolução à cura de diversas patologias e trará inúmeras surpresas

11.BIBLIOGRAFIA

- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Patrick Boyaval, Moineau, S., ... Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, 315(2007), 1709–1712. doi:10.1126/science.1138140
- Beloglazova, N., Brown, G., Zimmerman, M. D., Proudfoot, M., Makarova, K. S., Kudritska, M., ... Yakunin, A. F. (2008). A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats. *Journal of Biological Chemistry*, 283(29), 20361–20371. doi:10.1074/jbc.M803225200
- Bernard, H. U., Burk, R. D., Chen, Z., van Doorslaer, K., Hausen, H. zur, & de Villiers, E. M. (2010). Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology*, 401(1), 70–79. doi:10.1016/j.virol.2010.02.002
- Bhaya, D., Davison, M., & Barrangou, R. (2011). CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annual Review of Genetics*, 45, 273–97. doi:10.1146/annurev-genet-110410-132430
- Bosch, F. X., Lorincz, A., Muñoz, N., Meijer, C. J. L. M., & Shah, K. V. (2002). The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol*. 2002 Apr; 55(4): 244–265., (January), 244–265.
- Bryan, J., Jansen, K., Lowe, R., Fife, H., McClowry, T., Glass, D., & Brown, D. (1997). Human papillomavirus type 11 neutralization in the athymic mouse xenograft system: correlation with virus-like particle IgG concentration. *Med Virol.*, 53, 185–188.
- Capecchi, M. (1989). Altering the genome by homologous recombination. *Science*, 244(4910), 1288–1292.
- Chakraborty, S., Snijders, A. P., Chakravorty, R., Ahmed, M., Tarek, A. M., & Hossain, M. A. (2010). Comparative network clustering of direct repeats (DRs) and cas genes confirms the possibility of the horizontal transfer of CRISPR locus among bacteria. *Molecular Phylogenetics and Evolution*, 56(3), 878–887. doi:10.1016/j.ympev.2010.05.020
- Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., ... Zhang, F. (2013).

- Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, 339(6121).
- Conger, K. L., Liu, J.-S., Kuo, S.-R., Chow, L. T., & Wang, T. S.-F. (1999). Human Papillomavirus DNA Replication. *Journal of Biological Chemistry*, 274(5), 2696–2705. doi:10.1074/jbc.274.5.2696
- Crow, J. M. (2012). Hpv□: the Global Burden. *Nature*, 488.
- D’Abramo, C. M., & Archambault, J. (2011). Small molecule inhibitors of human papillomavirus protein - protein interactions. *Open Virol J*, 5(514), 80–95. doi:10.2174/1874357901105010080
- Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., ... Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, (471), 602–607.
- Deng, D., Yin, P., Yan, C., Pan, X., Gong, X., Qi, S., ... Shi, Y. (2012). Recognition of methylated DNA by TAL effectors. *Cell Research*, 22(10), 1502–1504. doi:10.1038/cr.2012.127
- Deveau, H., Garneau, J. E., & Moineau, S. (2010). CRISPR/Cas system and its role in phage-bacteria interactions. *Annual Review of Microbiology*, 64, 475–493. doi:10.1146/annurev.micro.112408.134123
- Doorbar, J. (2005). The papillomavirus life cycle. *Journal of Clinical Virology*, 32(SUPPL.), 7–15. doi:10.1016/j.jcv.2004.12.006
- Doorbar, J., Quint, W., Banks, L., Bravo, I. G., Stoler, M., Broker, T. R., & Stanley, M. a. (2012). The biology and life-cycle of human papillomaviruses. *Vaccine*, 30 Suppl 5, F55–70. doi:10.1016/j.vaccine.2012.06.083
- Ebina, H., Misawa, N., Kanemura, Y., & Koyanagi, Y. (2013). Harnessing the CRISPR-Cas9 systems disrupt latent HIV-1 provirus. *Scientific Reports*.
- Elenbaas, B., Spirio, L., Koerner, F., Fleming, M. D., Zimonjic, D. B., Donaher, J. L., ... Weinberg, R. a. (2001). Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. *Genes & Development*, 15(1), 50–65. doi:10.1101/gad.828901.monly
- F. J M, A. R. M., & Garrett. (2013). CRISPR-Cas Systems. In *CRISPR-Cas Systems RNA-Mediated Adaptive Immunity in Bacteria and Archae* (pp. 1–26).
- Fanning, A. S., & Anderson, J. M. (1999). Protein modules as organizers of membrane

- structure. *Current Opinion in Cell Biology*, 11(4), 432–439. doi:10.1016/S0955-0674(99)80062-3
- Fay, A., IV, W. H. Y., Roden, R. B. S., & Moroianu, J. (2004). The Positively Charged Termini of L2 Minor Capsid Protein Required for Bovine Papillomavirus Infection Function Separately in Nuclear Import and DNA Binding. *Virology*, 78, 13447–13454.
- Fehrmann, F., & Laimins, L. a. (2003). Human papillomaviruses: targeting differentiating epithelial cells for malignant transformation. *Oncogene*, 22(33), 5201–5207. doi:10.1038/sj.onc.1206554
- Ferraz, L. D. C., Beatriz, A., Santos, R., & Discacciati, M. G. (2012). Ciclo celular , HPV e evolução da neoplasia intraepitelial cervical□: seleção de marcadores biológicos. *Journal of the Health Sciences Institute*, 30(2), 107–111.
- Forman, D., de Martel, C., Lacey, C. J., Soerjomataram, I., Lortet-Tieulent, J., Bruni, L., ... Franceschi, S. (2012). Global burden of human papillomavirus and related diseases. *Vaccine*, 30 Suppl 5, F12–23. doi:10.1016/j.vaccine.2012.07.055
- Gaj, T., Gershbach, C. A., & Barbas, C. F. (2013). ZFN, TALEN and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnology*, 31, 397–405.
- Gasiunas, G., Sinkunas, T., & Siksnys, V. (2014). Molecular mechanisms of CRISPR-mediated microbial immunity. *Cellular and Molecular Life Sciences*, 71(3), 449–465. doi:10.1007/s00018-013-1438-6
- GenScript. (n.d.). *CRISPR Handbook*.
- Gnanamony, M., Peedicayil, A., & Abraham, P. (2007). An overview of human papillomaviruses and current vaccine strategies. *Indian Journal of Med. Micro.*, 25(1), 10–17.
- Godde, J. S., & Bickerton, A. (2006). The repetitive DNA elements called CRISPRs and their associated genes: Evidence of horizontal transfer among prokaryotes. *Journal of Molecular Evolution*, 62(6), 718–729. doi:10.1007/s00239-005-0223-z
- Gudbergsdottir, S. D. L., Chen, Z., Jensen, J., Jensen, L., & She, Q. (2011). Dynamic properties of the Sulfolobus CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers. *Molecular Microbiology*, 79, 35–49.

- Hale, C. R., Majumdar, S., Elmore, J., Pfister, N., Compton, M., Olson, S., ... Terns, M. P. (2012). Essential Features and Rational Design of CRISPR RNAs that Function with the Cas RAMP Module Complex to Cleave RNAs. *Molecular Cell*, 45(3), 292–302. doi:10.1016/j.molcel.2011.10.023
- Held, N. L., Herrera, A., Quiroz, H. C., & Whitaker, R. J. (2010). CRISPR associated diversity within a population of *Sulfolobus islandicus*. *PLoS ONE*, 5(9). doi:10.1371/journal.pone.0012988
- Horvath, P., Romero, D. A., Coûté-Monvoisin, A. C., Richards, M., Deveau, H., Moineau, S., ... Barrangou, R. (2008). Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *Journal of Bacteriology*, 190(4), 1401–1412. doi:10.1128/JB.01415-07
- Howard, J. a L., Delmas, S., Ivančić-Baće, I., & Bolt, E. L. (2011). Helicase dissociation and annealing of RNA-DNA hybrids by *Escherichia coli* Cas3 protein. *The Biochemical Journal*, 439(1), 85–95. doi:10.1042/BJ20110901
- Howie, H. L., Katzenellenbogen, R. A., & Galloway, D. A. (2009). Papillomavirus E6 proteins. *Virology*, 384(2), 324–334. doi:10.1016/j.virol.2008.11.017
- Hu, W., Kaminski, R., Yang, F., Zhang, Y., Cosentino, L., Li, F., ... Khalili, K. (2014). RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. *Proceedings of the National Academy of Sciences of the United States of America*, 111(31), 11461–6. doi:10.1073/pnas.1405186111
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology*, 169(12), 5429–33. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=213968&tool=pmcentrez&rendertype=abstract>
- Jansen, R., Van Embden, J. D. A., Gaastra, W., & Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, 43(6), 1565–1575. doi:10.1046/j.1365-2958.2002.02839.x
- Karvelis, T., Gasiunas, G., Miksys, A., Barrangou, R., Horvath, P., & Siksnys, V. (2013). crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA Biology*, 10(5), 841–51. doi:10.4161/rna.24203

- Koonin, E. V, & Makarova, K. S. (2013). CRISPR-Cas: evolution of an RNA-based adaptive immunity system in prokaryotes. *RNA Biology*, 10(5), 679–86. doi:10.4161/rna.24022
- Kunin, V., Sorek, R., & Hugenholtz, P. (2007). Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biology*, 8(4), R61. doi:10.1186/gb-2007-8-4-r61
- Lazarczyk, M., Cassonnet, P., Pons, C., Jacob, Y., & Favre, M. (2009). The EVER proteins as a natural barrier against papillomaviruses: a new insight into the pathogenesis of human papillomavirus infections. *Microbiology and Molecular Biology Reviews*: *MMBR*, 73(2), 348–70. doi:10.1128/MMBR.00033-08
- Lillestøl, R. K., Shah, S. A., Brügger, K., Redder, P., Phan, H., Christiansen, J., & Garrett, R. A. (2009). CRISPR families of the crenarchaeal genus *Sulfolobus*: Bidirectional transcription and dynamic properties. *Molecular Microbiology*, 72(1), 259–272. doi:10.1111/j.1365-2958.2009.06641.x
- Lima, M. A. P., Silva, C. G. L., & Rabenhorst, S. H. B. (2013). Papel das Proteínas Precoces do Papilomavírus Humano na Carcinogênese Role of the Human Papillomavirus Early Proteins in the Carcinogenesis, 59(4), 565–573.
- Longworth, M. S., & Laimins, L. a. (2004). Pathogenesis of Human Papillomaviruses in Differentiating Epithelia. *Microbiology*, 68, 362–372.
- Lowy, D. R., & Schiller, J. T. (2006). Review series Prophylactic human papillomavirus vaccines. *The Journal of Clinical Investigation*, 116(5), 1167–1173. doi:10.1172/JCI28607.ond
- Luiciano A., E. J. M., & Sontheimer. (2010). CRISPR Interference: RNA-directed adaptive immunity in bacteria and archae. *Nature Reviews. Genetics*, 11, 181–190.
- Makarova, K. S., Aravind, L., Grishin, N. V, Rogozin, I. B., & Koonin, E. V. (2002). A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Research*, 30(2), 482–496. doi:10.1093/nar/30.2.482
- Makarova, K. S., Brouns, S. J. J., Horvath, P., Sas, D. F., & Wolf, Y. I. (2012). Evolution and Classification of the CRISPR-Cas systems. *Nature Reviews ...*, 9(6), 467–477. doi:10.1038/nrmicro2577.Evolution

- Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., & Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology Direct*, *1*(1), 7. doi:10.1186/1745-6150-1-7
- Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., ... Koonin, E. V. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nature Reviews. Microbiology*, *13*(11), 722–736. doi:10.1038/nrmicro3569
- Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, *155*(3), 733–740. doi:10.1099/mic.0.023960-0
- Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*, *60*(2), 174–182. doi:10.1007/s00239-004-0046-3
- Mojica, F. J. M., Díez-Villaseñor, C., Soria, E., & Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Molecular Microbiology*, *36*(1), 244–246. doi:10.1046/j.1365-2958.2000.01838.x
- Mu, K., Baldwin, A., Edwards, K. M., Hayakawa, H., Nguyen, C. L., Owens, M., & Grace, M. (2004). Mechanisms of Human Papillomavirus-Induced Oncogenesis, *78*(21), 11451–11460. doi:10.1128/JVI.78.21.11451
- Munger, K., Basile, J. R., Duensing, S., Eichten, A., Gonzalez, S. L., Grace, M., & Zaczny, V. L. (2001). Biological activities and molecular targets of the human papillomavirus E7 oncoprotein. *Oncogene*, *20*(54), 7888–7898. doi:10.1038/sj.onc.1204860
- Munoz N, Bosch FX, de Sanjose S, Herrero R, Castetellaasague X, S. K. et al. (2003). Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med*, *348*, 518–527.
- Navarro, C., Martin, R., Carlos, L., Ferreira, D. L., & Neto, G. B. (2013). Custoe-efetividade das estratégias de prevenção primária e secundária do câncer de colo de

- útero para o Brasil, *9*, 11–17.
- Pereira, G., Porro, A. M., Francisco, G., & Tomimori, J. (2011). Infecção pelo papilomavírus humano: etiopatogenia, biologia molecular e manifestações clínicas, 306–317.
- Pista, A., Oliveira, C., Paixão, M., Real, O., & Cunha, M. (2011). Epidemiological study of prevalence of HPV infection of women in Portugal. *Int. J Gynecol Cancer*, 1150–1158.
- R. Edze, P. C. W., Angus, B., & Fineran. (2014). CRISPR-Cas systems: beyond adaptive immunity. *Nature*, *12*.
- Ran, F. A., Hsu, P. D. P., Wright, J., Agarwala, V., Scott, D. a, & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, *8*(11), 2281–2308. doi:10.1038/nprot.2013.143.Genome
- Rath, D., Amlinger, L., Rath, A., & Lundgren, M. (2015). The CRISPR-Cas immune system: Biology, mechanisms and applications. *Biochimie*, *117*, 119–128. doi:10.1016/j.biochi.2015.03.025
- Sapp, M., & Bienkowska-Haba, M. (2009). Viral entry mechanisms: human papillomavirus and a long journey from extracellular matrix to the nucleus. *FEBS J*, 7206–7216.
- Schwank, G., Koo, B. K., Sasselli, V., Dekkers, J. F., Heo, I., Demircan, T., ... Clevers, H. (2013). Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell*, *13*(6), 653–658. doi:10.1016/j.stem.2013.11.002
- Shah, S. A., Erdmann, S., Mojica, F. J. M., & Garrett, R. A. (2013). Protospacer recognition motifs: mixed identities and functional diversity. *RNA Biology*, *10*(5), 891–9. doi:10.4161/rna.23764
- Souto, R., Falhari, J., & Cruz, A. (2005). O Papilomavírus Humano: um fator relacionado com a formação de neoplasias. *Revista Brasileira de Canceroogia*, *51*(2), 155–160.
- Szczepankowska, A. (2012). Bacteriophages, Part A. In *Bacteriophages, Part A* (pp. 289–338).
- Utomo, A., Nikitin, A., & Lee, W. (1999). Temporal, spatial, and cell type-specific

- control of Cre-mediated DNA recombination in transgenic mice. *Nature Biotechnology*, 17.
- van Belkum, a, Scherer, S., van Alphen, L., & Verbrugh, H. (1998). Short-sequence DNA repeats in prokaryotic genomes. *Microbiology and Molecular Biology Reviews*: *MMBR*, 62(2), 275–293.
- Van Der Oost, J., Westra, E. R., Jackson, R. N., & Wiedenheft, B. (2014). Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nature Reviews Microbiology*, 12, 479–492.
- Waggoner, S. E. (2003). Cervical cancer. *Lancet*, 361(9376), 2217–2225. doi:10.1016/S0140-6736(03)13778-6
- Wang, R., Preamplume, G., Terns, M. P., & Li, H. (2011). Interaction of the Cas6 ribonuclease with CRISPR RNAs: recognition and cleavage, 19(2), 257–264.
- Westra, E. R., Pul, U., Heidrich, N., Jore, M. M., Lundgren, M., Stratmann, T., ... Brouns, S. J. J. (2010). H-NS-mediated repression of CRISPR-based immunity in Escherichia coli K12 can be relieved by the transcription activator LeuO. *Molecular Microbiology*, 77(6), 1380–1393. doi:10.1111/j.1365-2958.2010.07315.x
- White, M. K., Hu, W., & Khalili, K. (2015). The CRISPR/CAS9 genome editing methodology as a weapon against human virus. *Discovery Medicine*, 73(4), 389–400. doi:10.1530/ERC-14-0411.Persistent
- White, M. K., Pagano, J., & Khalili, K. (2014). Viruses and Human Cancers: a Long Road of Discovery of Molecular Paradigms. *Clinical Microbiology Reviews*, 27(3), 463–481.
- WHO. (n.d.). Retrieved from <http://www.who.int/hpvcentre/en/>
- Wright, D. A., Li, T., Yang, B., & Spalding, M. H. (2014). TALEN-mediated genome editing: prospects and perspectives. *Biochemical Journal*, 462(1), 15–24. doi:10.1042/BJ20140295
- Wu, X., & Levine, A. J. (1994). p-53 and E2F-1 cooperate to mediate apoptosis. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 3602–3606.
- Yosef, I., Goren, M. G., & Qimron, U. (2012). Proteins and DNA elements essential for

- the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Research*, 40(12), 5569–5576. doi:10.1093/nar/gks216
- Yu, L., Wang, X., Zhu, D., Ding, W., Wang, L., Zhang, C., ... Wang, H. (2014). Disruption of human papillomavirus 16 E6 gene by clustered regularly interspaced short palindromic repeat/Cas system in human cervical cancer cells. *Oncotargets and Therapy*, 8, 37–44. doi:10.2147/OTT.S64092
- Zegans, M. E., Wagner, J. C., Cady, K. C., Murphy, D. M., Hammond, J. H., & O'Toole, G. A. (2009). Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *Journal of Bacteriology*, 91(1), 210–219. doi:10.1128/JB.00797-08
- Zhang, J., Rouillon, C., Kerou, M., Reeks, J., Brugger, K., Reimann, J., ... White, M. F. (2012). Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Mol Cell*, 45(3), 303–313. doi:10.1016/j.molcel.2011.12.013.Structure