



# Ferramentas Moleculares para Investigação da Função Gênica

Janaina de Freitas Nascimento, PhD

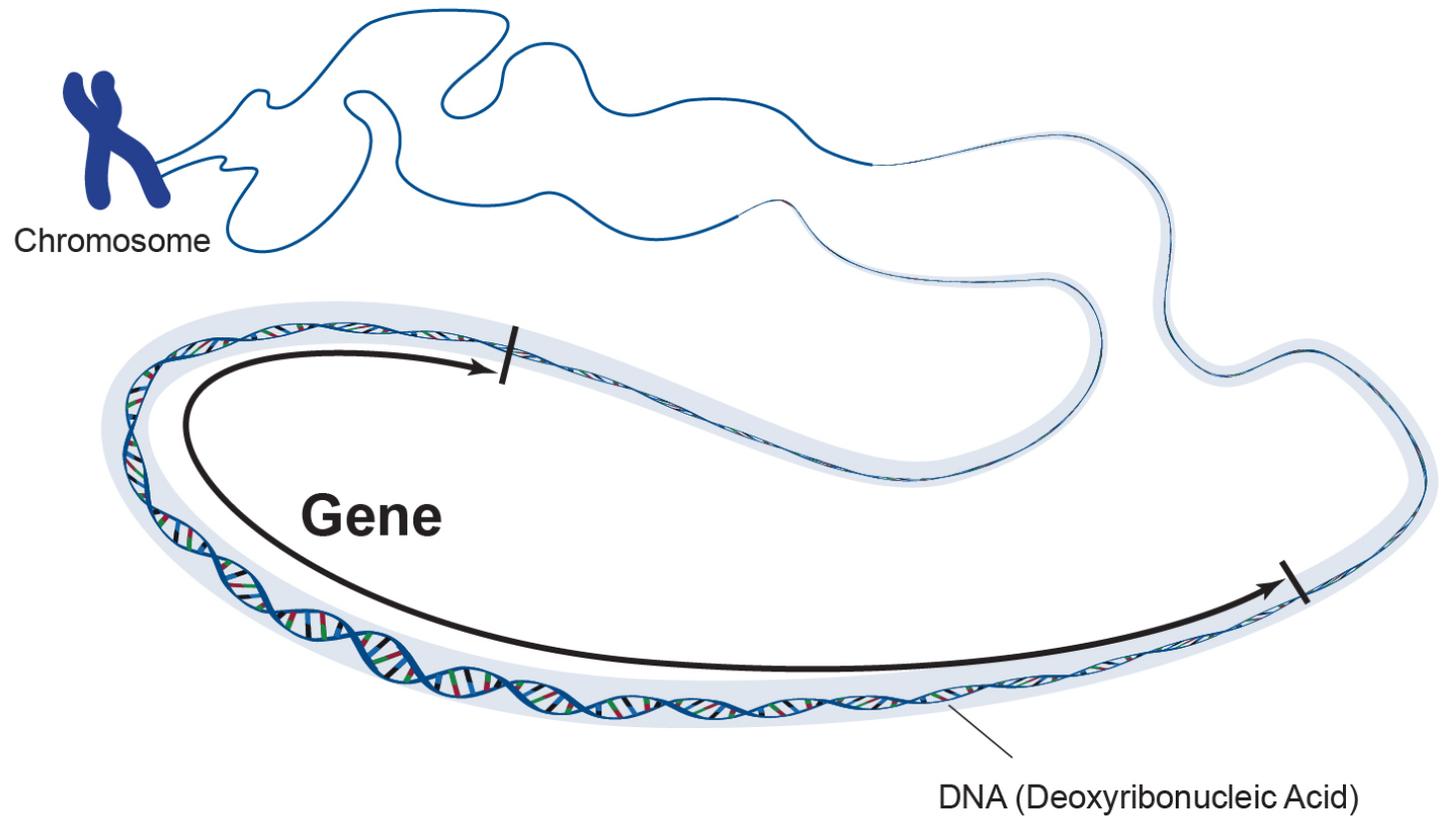
Pós-doutoranda

Laboratório de Bioquímica de Tryps

Departamento de Parasitologia - USP

[janainafn@usp.br](mailto:janainafn@usp.br)

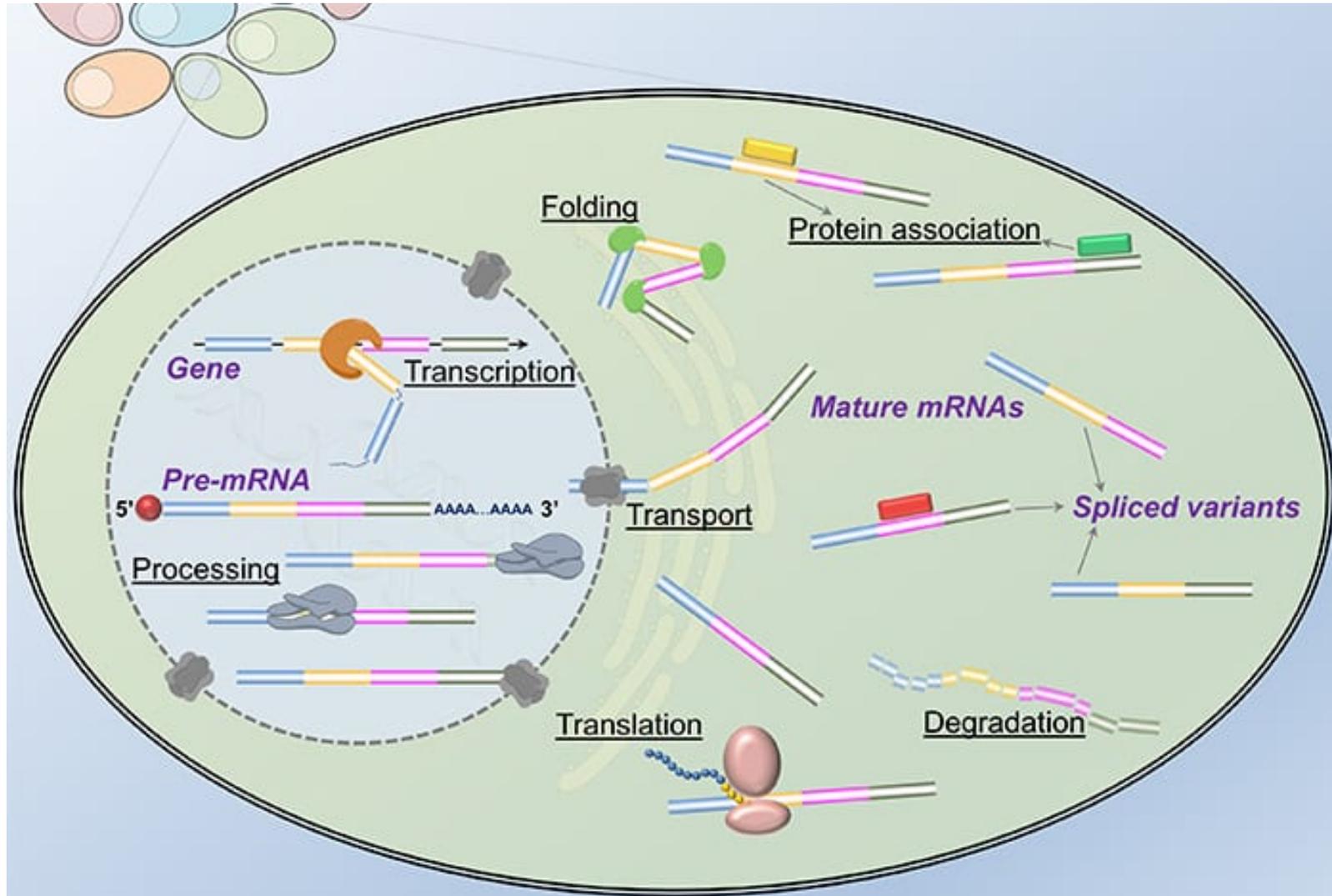
# O que é um gene?



<https://www.genome.gov/sites/default/files/tg/en/illustration/gene.jpg>

*Um gene é uma sequência de DNA (cujos segmentos componentes não precisam necessariamente ser fisicamente contíguos) que codificam um ou mais RNAs ou proteínas. (Portin & Wilkins (2017) Genetics)*

# Expressão gênica



Como podemos investigar a  
função de um gene?

# Tópicos

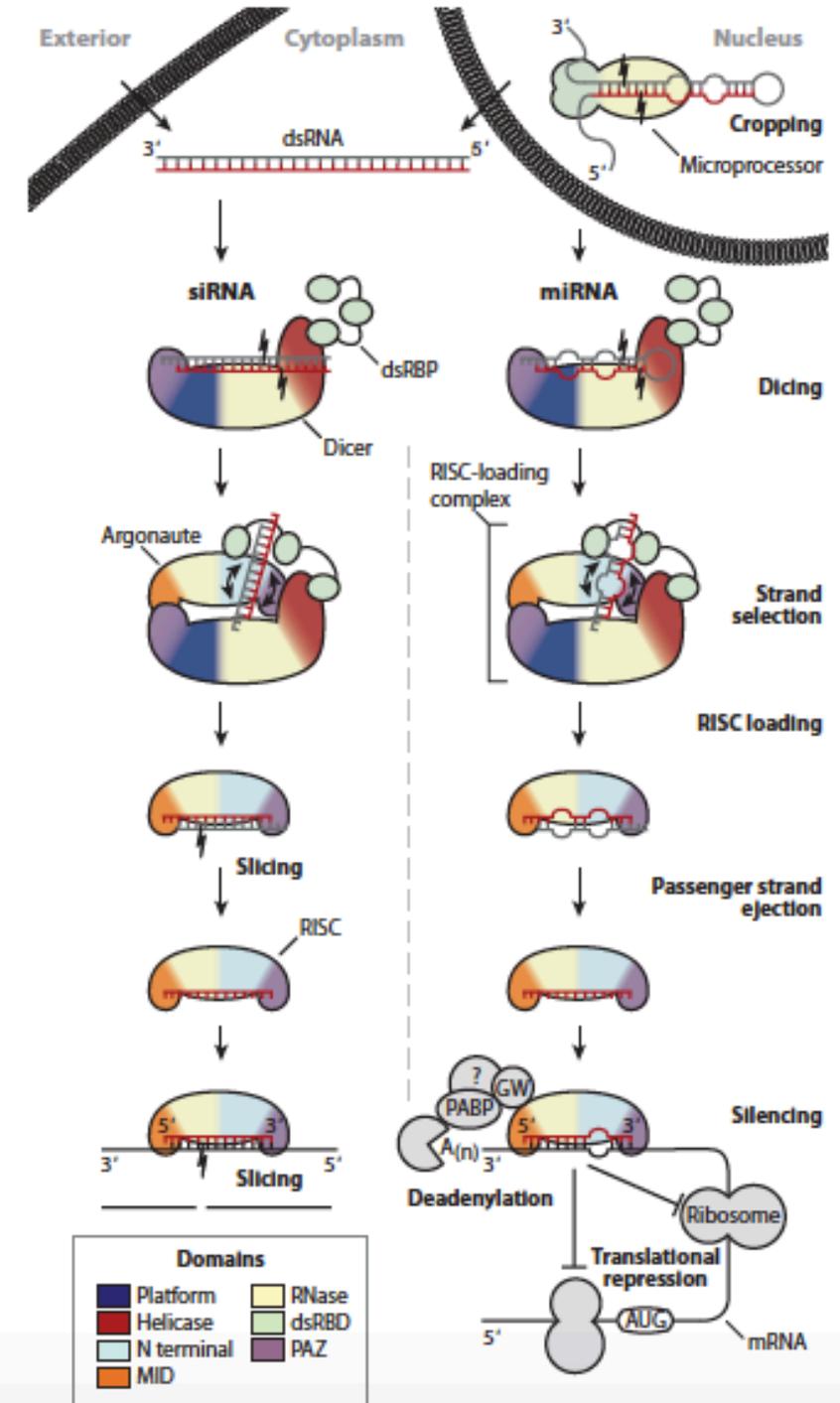
- **Sistemas para regulação dos níveis de mRNA**
  - RNA de interferência – RNAi
  - Ribozimas - *riboswitches*
- **Sistemas para regulação dos níveis proteicos**
  - ddFKBP-Shld1
  - Auxin-inducible degron
- **Sistemas para edição gênica**
  - Cre-loxP
  - CRISPR-Cas

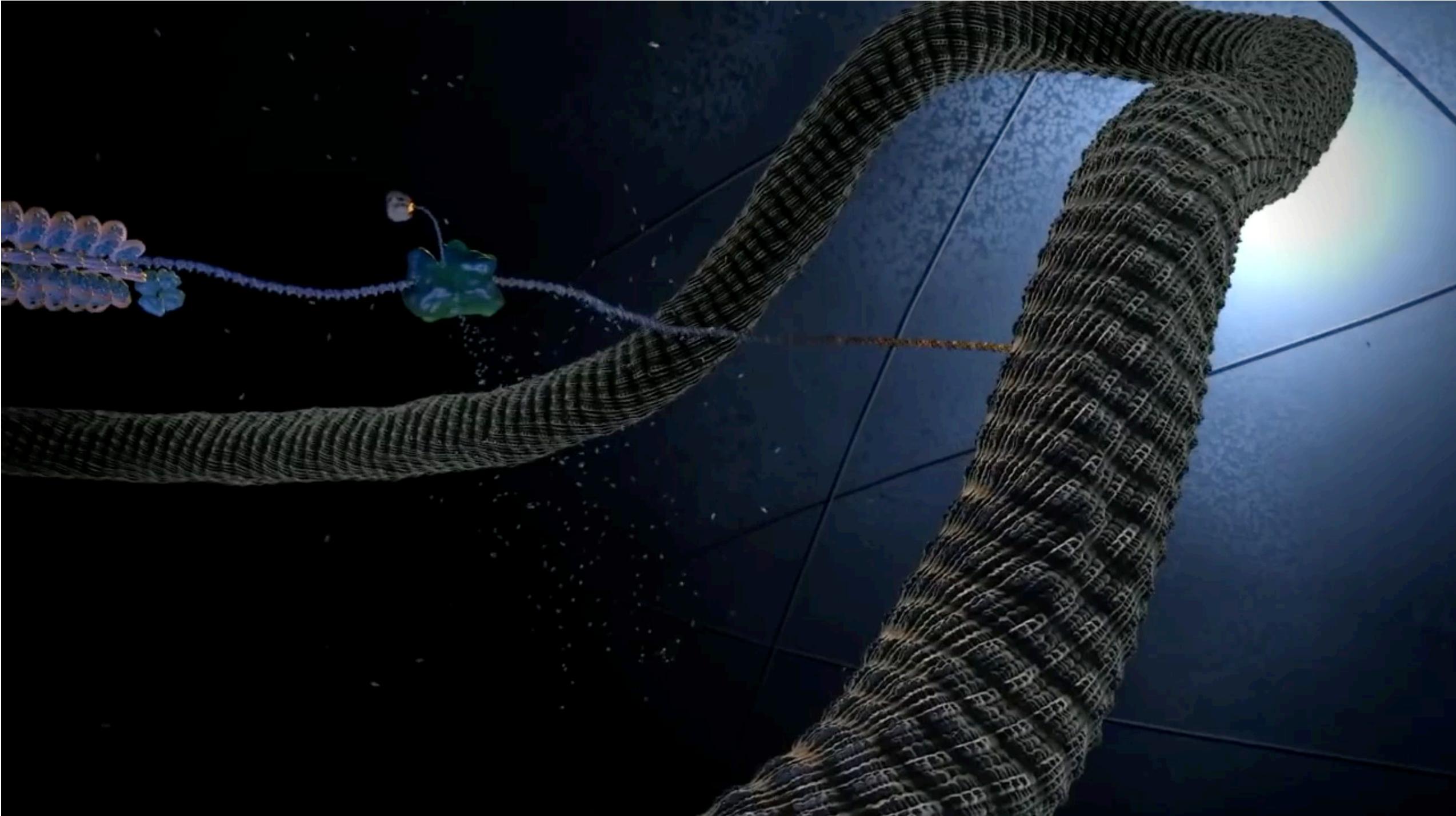
## Sistemas para regulação dos níveis de mRNA

RNA de interferência

# RNA de interferência - RNAi

- Mecanismo de controle de expressão gênica (miRNAs)
- Defesa contra vírus e transposons (siRNA)
- Conservada em eucariotos
- Componentes:
  - siRNA (small interfering RNA)/ miRNA (microRNA)
  - Dicer – endonuclease
  - dsRBP – proteína de ligação a dsRNA
  - Argonaute – proteína efetora





# RNA de interferência - RNAi

## Conservação entre parasitas protozoários

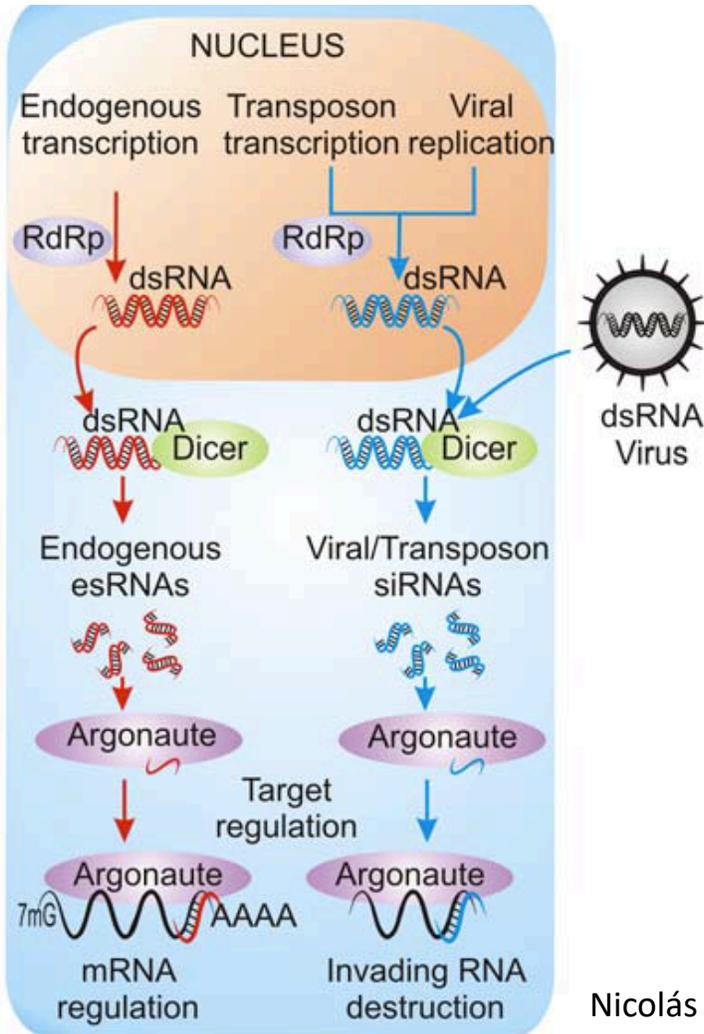


TABLE 1. RNAi genes and classes of small RNAs identified in protozoan parasites

Parasite	siRNAs	miRNAs	Argonaute	Dicer-like	RdRp
<i>T. brucei</i>	+	-	+	+	-
<i>L. braziliensis</i>	+	?	+	+	-
<i>E. histolytica</i>	+	?	+	?	+
<i>G. intestinalis</i>	+	+	+	+	+
<i>T. gondii</i>	+	+	+	+	+

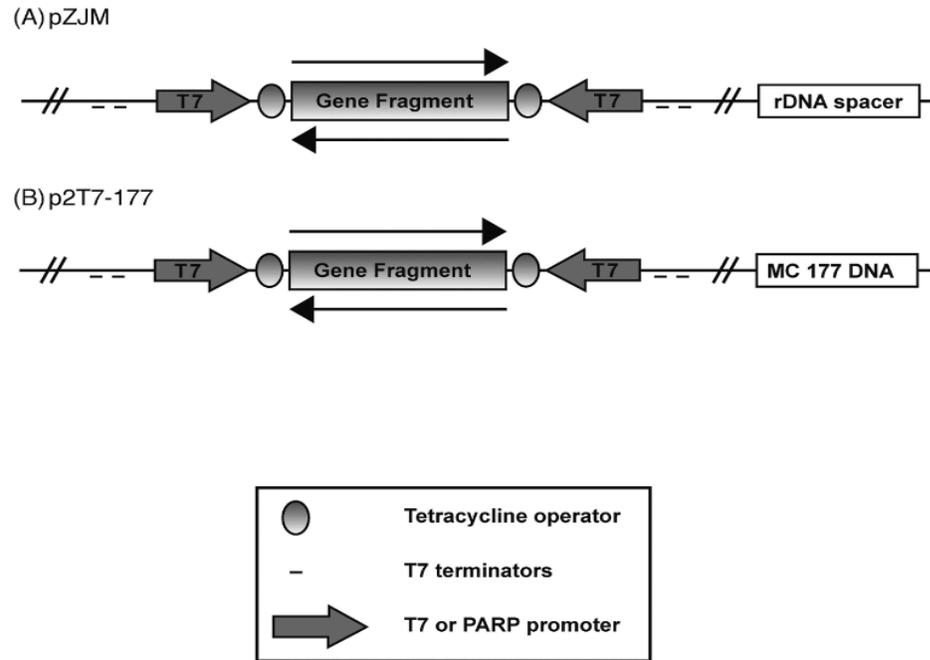
Kolev et al (2011) Eukaryotic Cell

Nicolás et al (2011) Plos Pathogens

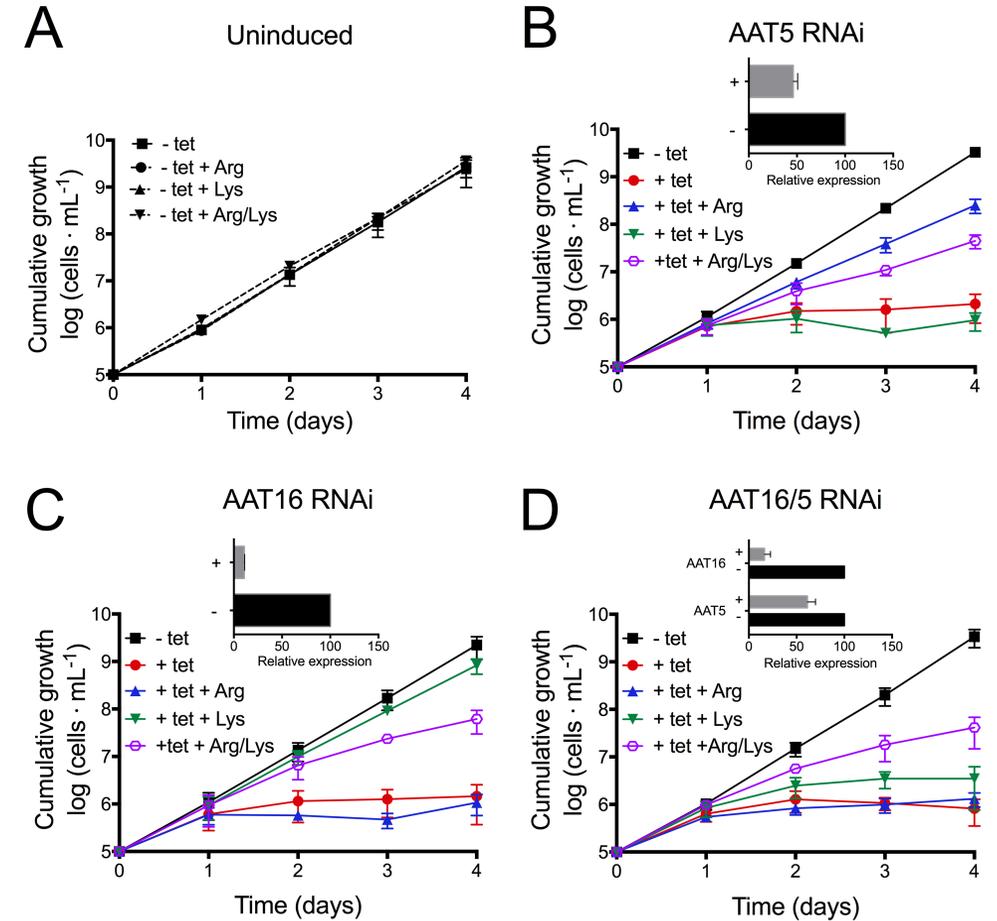
# RNA de interferência - RNAi

## Exemplos de sistemas utilizados em *T. brucei*

Linhagem transgênica de *T. brucei* expressando TetR e RNA T7 Pol



Bellofatto and Palenchar (2008) Methods in Molecular Biology



Mathieu et al (2017) Plos One

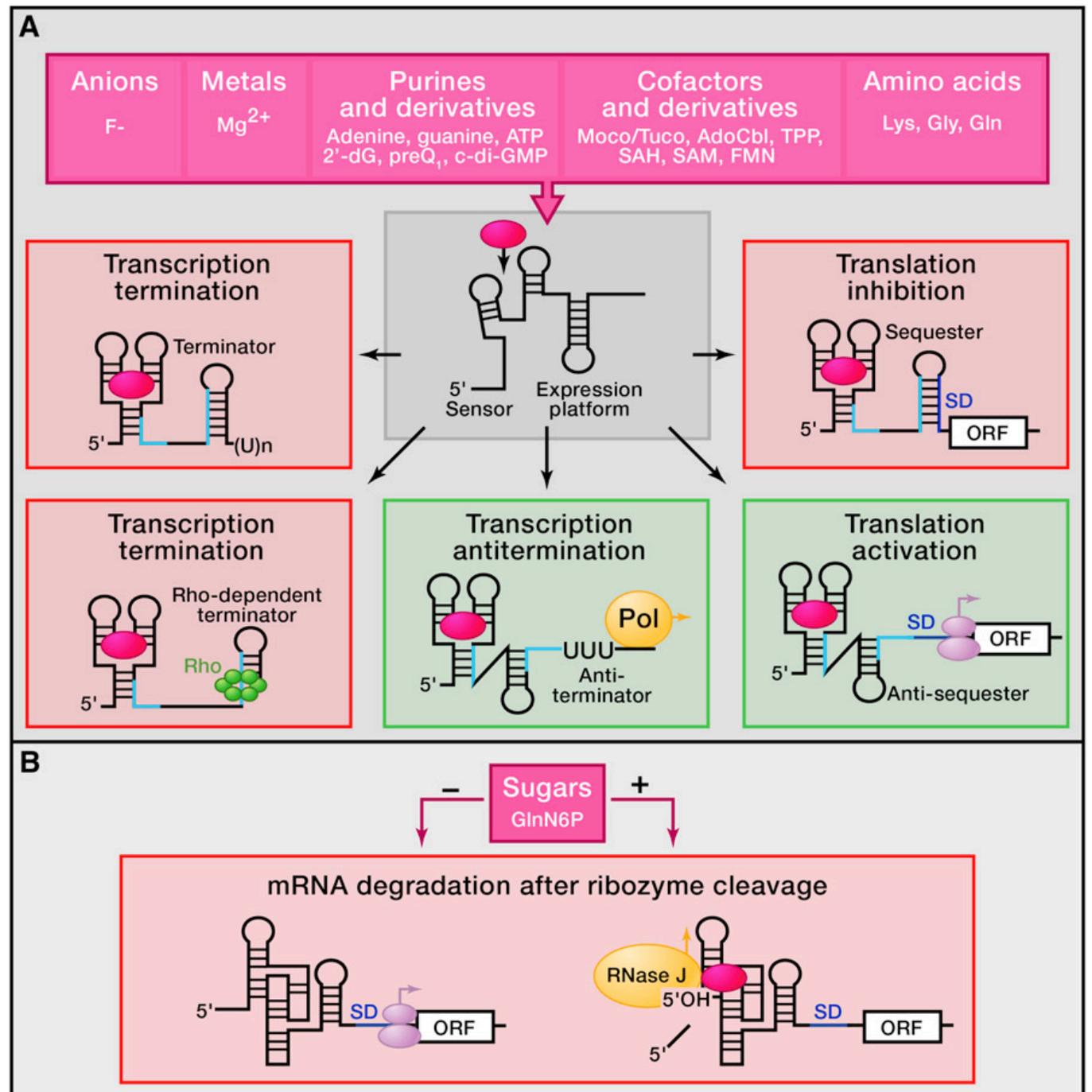
## Sistemas para regulação dos níveis de mRNA

Ribozimas - *Riboswitches*

# Ribozimas e *Riboswitches*

- Ribozimas são RNAs com atividade autocatalítica
- Riboswitches é uma estrutura encontrada nas regiões 5' não-traduzidas (UTRs) de mRNAs que codificam genes metabólicos, onde controlam a expressão gênica adotando estruturas alternativas de RNA na presença ou ausência de pequenos ligantes.

# Riboswitches

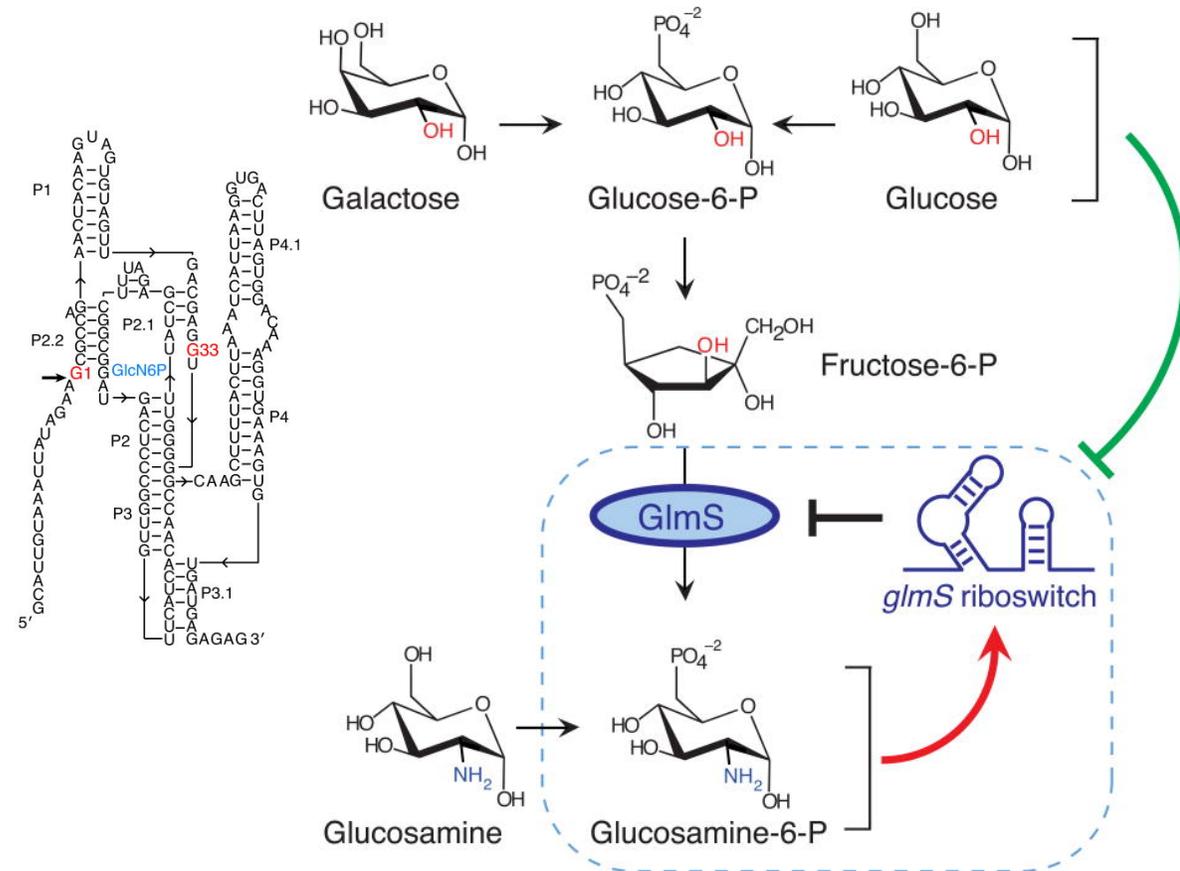


# Ribozimas - *Riboswitches*

## Mecanismo de regulação do riboswitch *glmS*

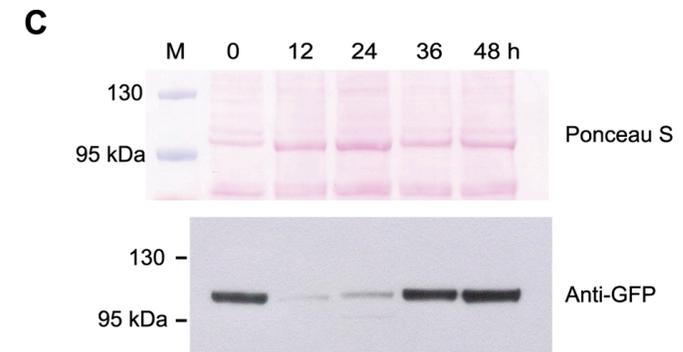
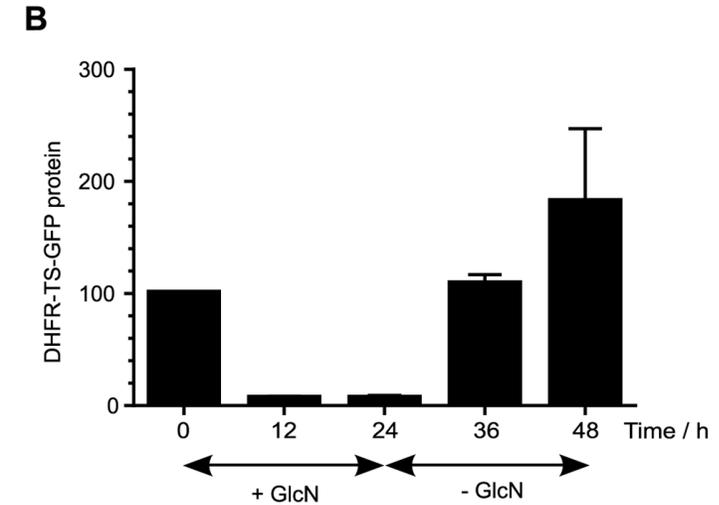
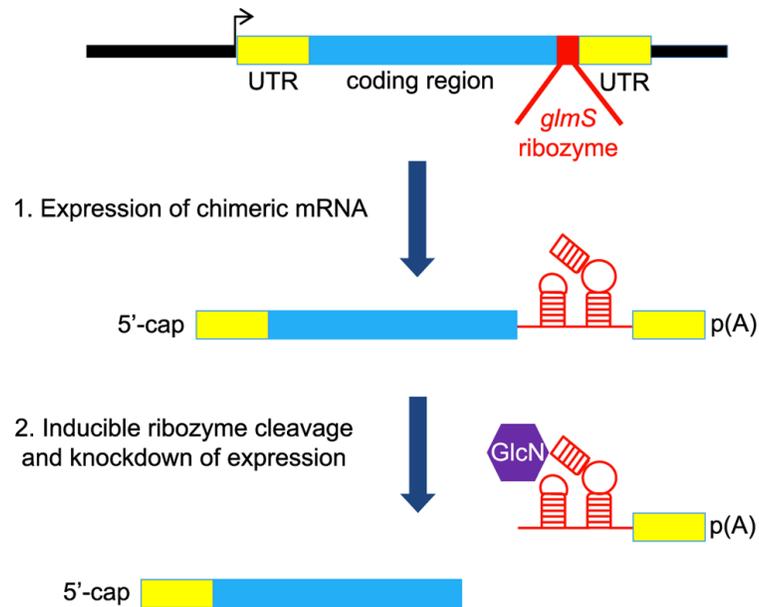
O riboswitch *glmS* (do gene glucosamina-6-fosfato aminotransferase) pertence à família de RNAs reguladores que fornecem regulação de feedback de genes metabólicos e tem atividade de ribozima que se autocliva ao se ligar a glucosamina-6-fosfato (GlcN6P), o produto da enzima codificada por *glmS* em bactérias gram-positivas.

a



# Ribozimas - *Riboswitches*

## Exemplos de sistemas utilizados em *Plasmodium*

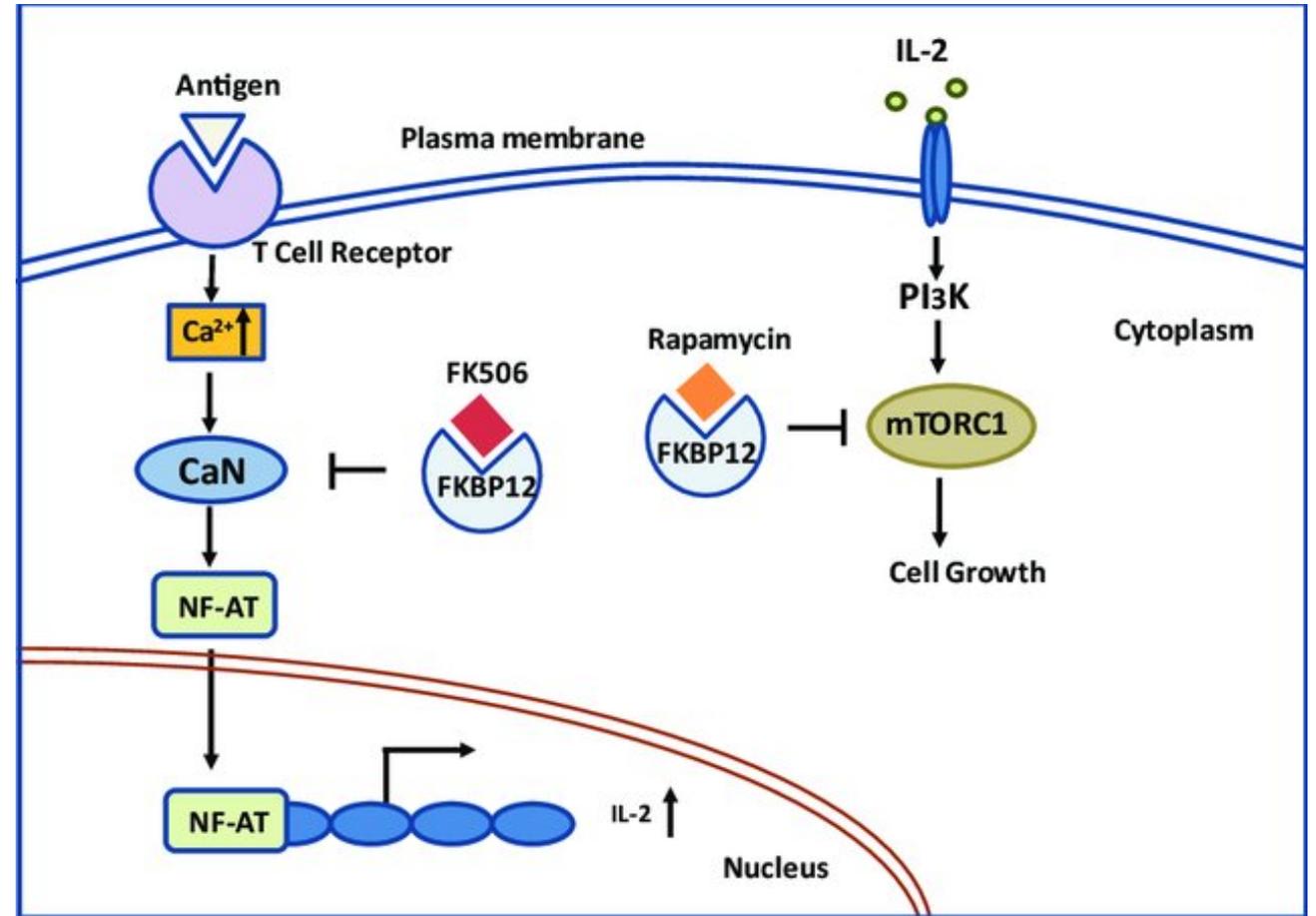


## Sistemas para regulação dos níveis de proteínas

Sistema ddFKBP-ShId1

# Sistema ddFKBP-Shld1

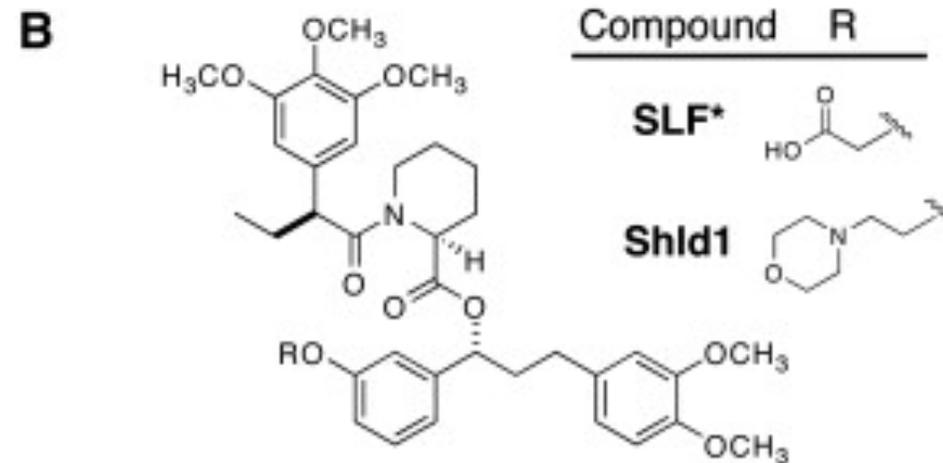
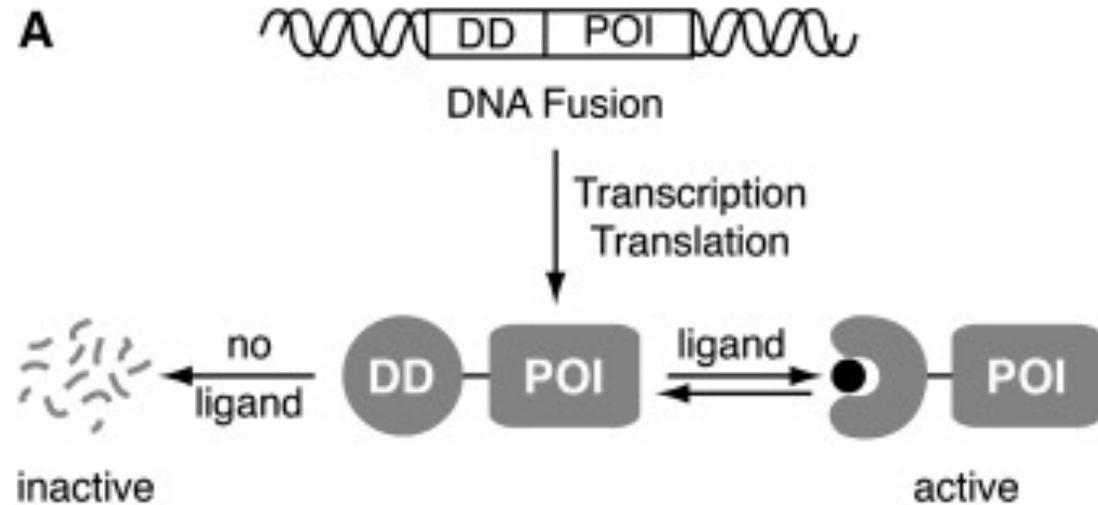
- FK506 binding proteins (FKBPs) são uma família de proteínas (imunofilinas) que tem função regulatória ou de chaperona. FKBP12 se liga ao produto natural imunossupressor FK506 e rapamicina.



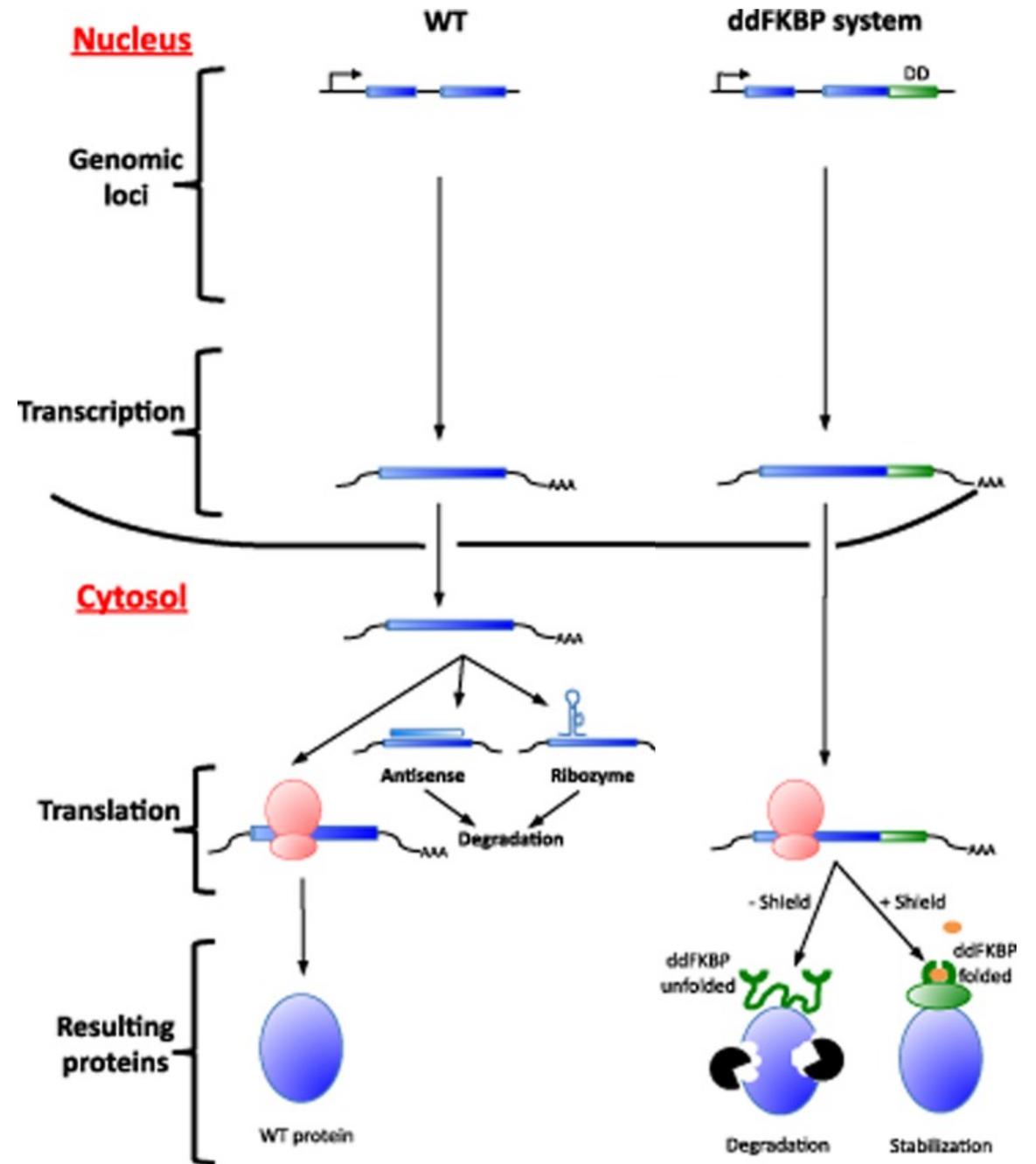
# Sistema ddFKBP-Shld1

## Mecanismo

- Utiliza a fusão de um domínio de desestabilização controlada por ligante (ddFKBP) a uma proteína de interesse.
- A estabilização seletiva é alcançada por ligação reversível de um ligante sintético, denominado Shield-1 (Shld1)

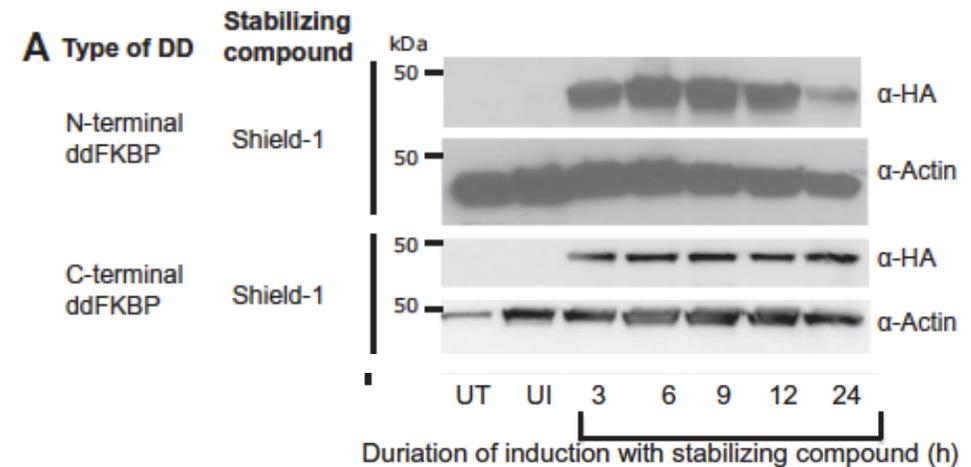
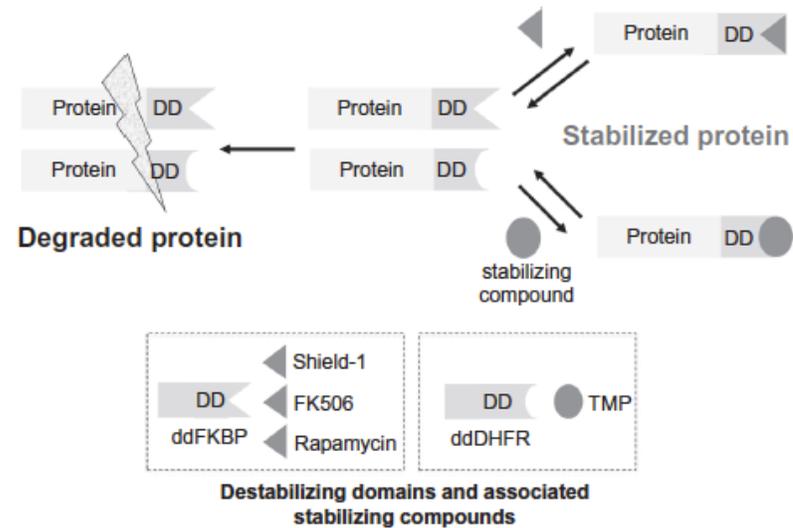


# *Sistema ddFKBP-Shld1*



# Sistema ddFKBP-Shld1

Exemplo de sistema utilizado em *Entamoeba histolytica*

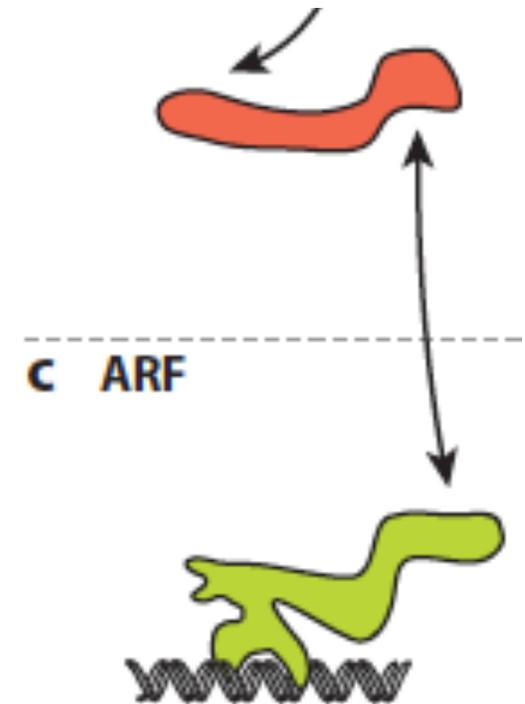


## Sistemas para regulação dos níveis de proteínas

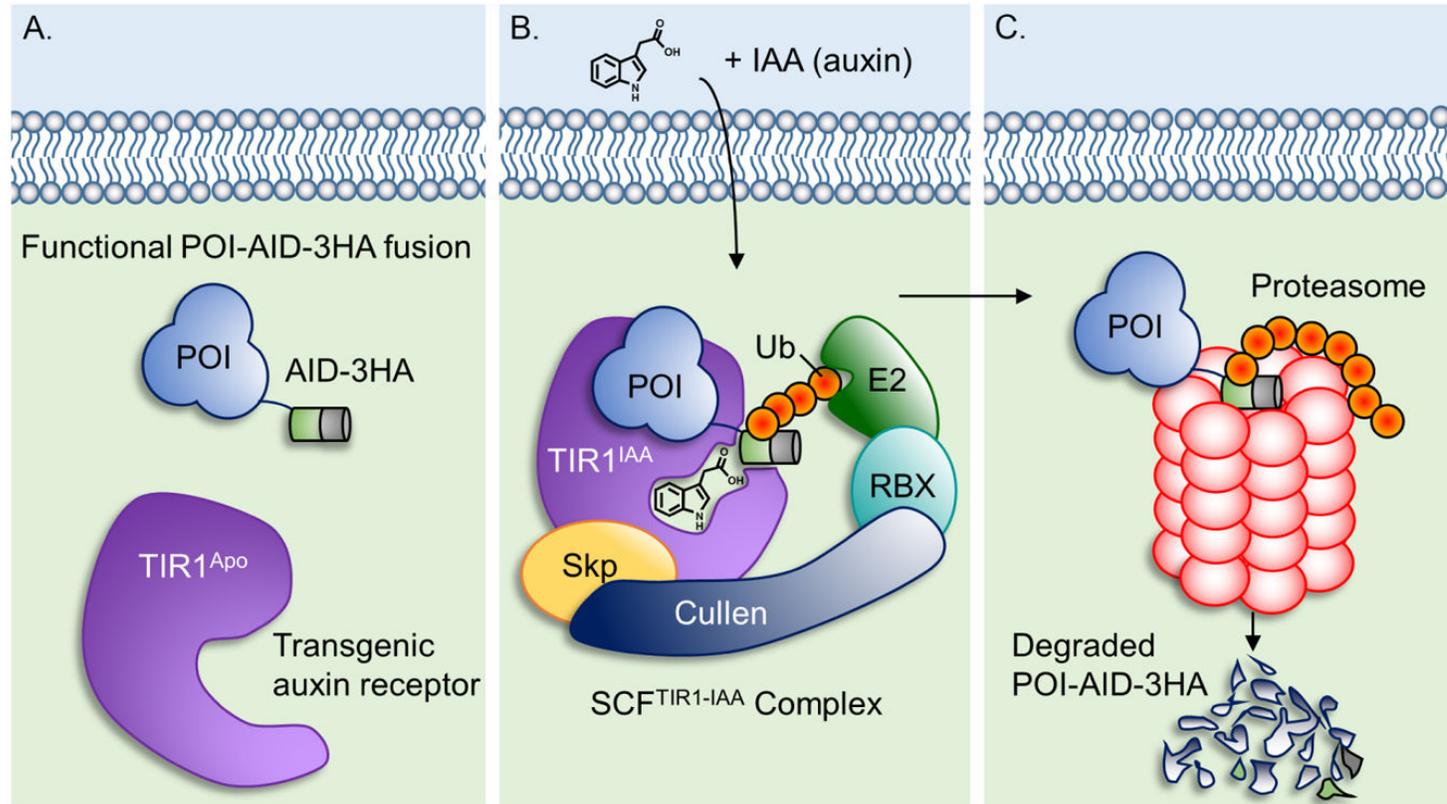
Sistema de degradação induzido por auxina

# Mecanismo de resposta a auxina

- Auxina é uma das principais moléculas sinalizadoras em plantas
- Componentes do sistema:
  - *Transport inhibitor resistant 1/auxin signaling f-box* (TIR1/AFB) – faz parte do complexo de ligação de ubiquitina SCF<sup>TIR1/AFB</sup> que contém SKP1, CUL1 e RBX1
  - *Auxin/indole-3-acetic acid transcriptional coregulators* (Aux/IAA) – co-reguladores transcrpcionais
  - *Auxin reponse factors* (ARFs) – proteínas de ligação ao DNA sequências específicas
  - Auxina promove a interação entre TIR1/AFB e Aux/IAA e desencadeia a degradação das proteínas Aux-IAA via ubiquitinação



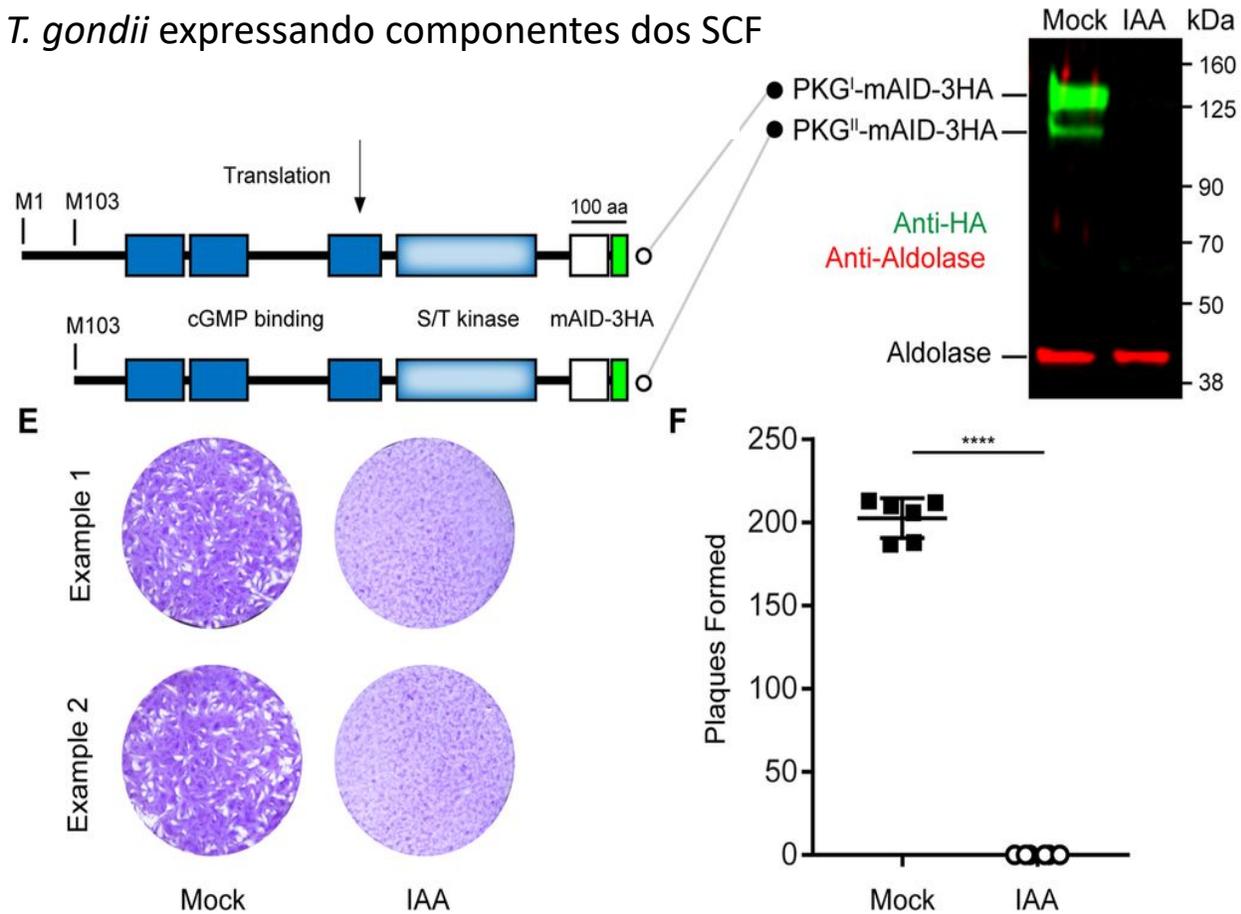
# *Sistema degron induzido por auxina (AID)*



# Sistema AID

## Exemplo de sistema utilizado em *Toxoplasma gondii*

Linhagem transgênica de *T. gondii* expressando componentes dos SCF



## Sistemas para edição gênica

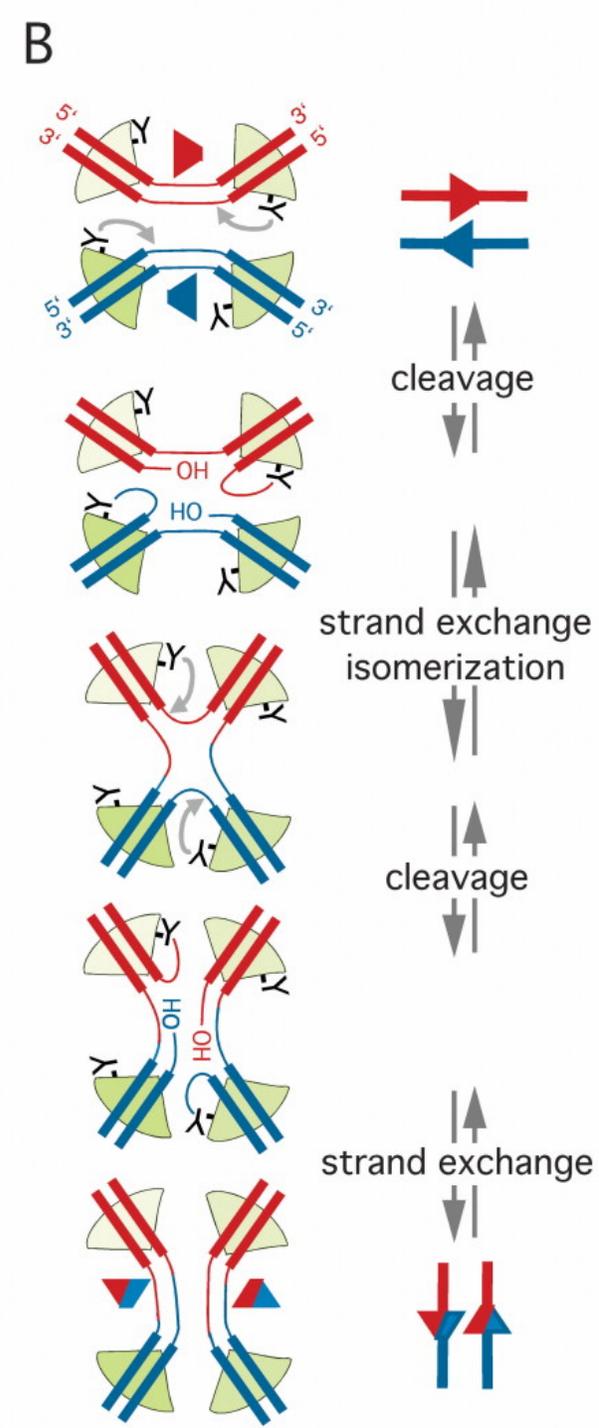
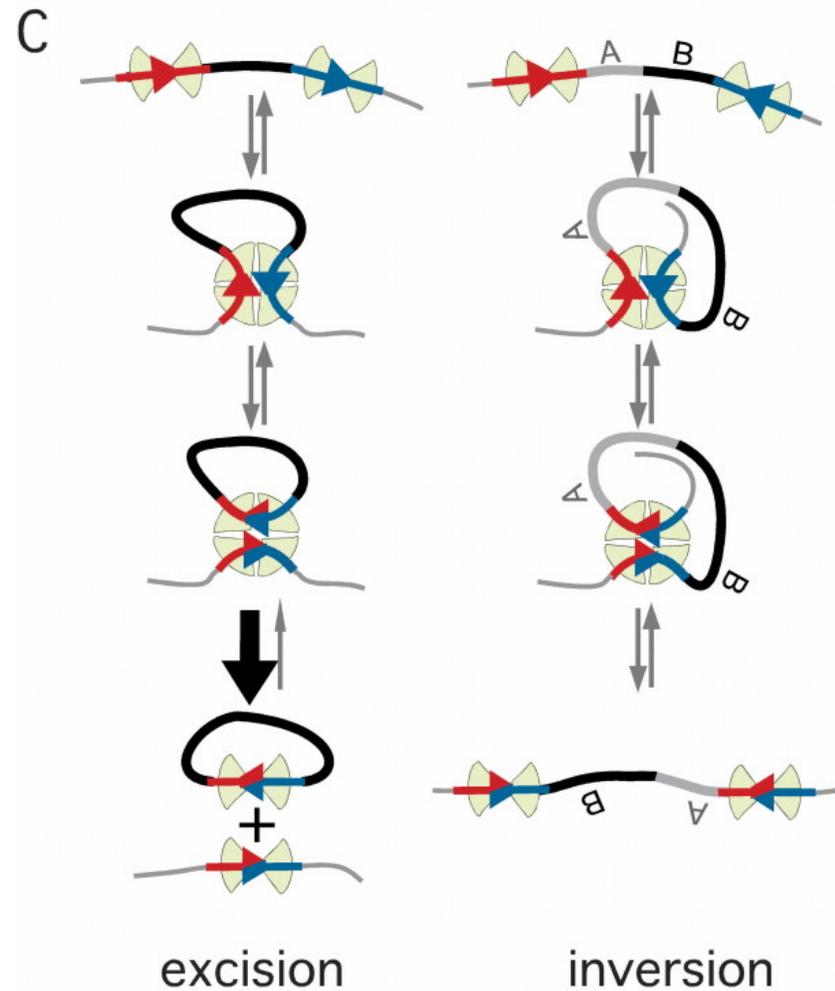
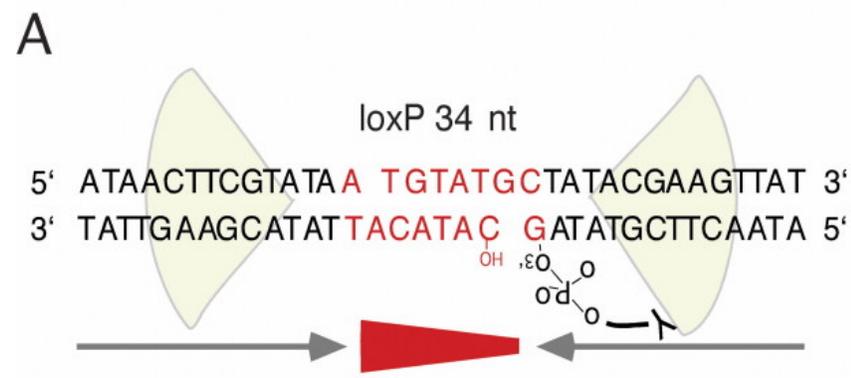
Sistema Cre-loxP

# *Sistema Cre-loxP*

- Utiliza dois componentes derivados do bacteriófago P1: a recombinase Cre e um sítio de reconhecimento loxP.
- Cre (*cause recombination*) recombinase, proteína de 38 kDa responsável pela recombinação intra e inter-molecular nos locais de reconhecimento de loxP. Atua independentemente de quaisquer outras proteínas acessórias ou cofatores.
- Os locais LoxP (*locus of X(cross)-over in P1*) são sequências de reconhecimento de 34 pares de bases que consistem em duas repetições palindrômicas de 13 bp separadas por uma sequência espaçadora assimétrica de 8 bp de comprimento: **ATAACTTCGTATA-GCATACAT-TATACGAAGTTAT**.

# Sistema Cre-loxP

## Mecanismo



# Sistema Cre-loxP

## Aplicações do sistema Cre-loxP

### Inversion

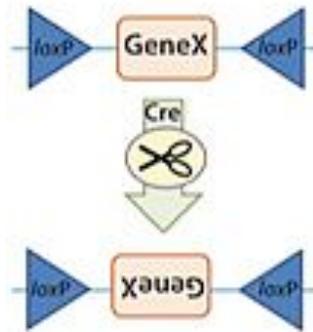
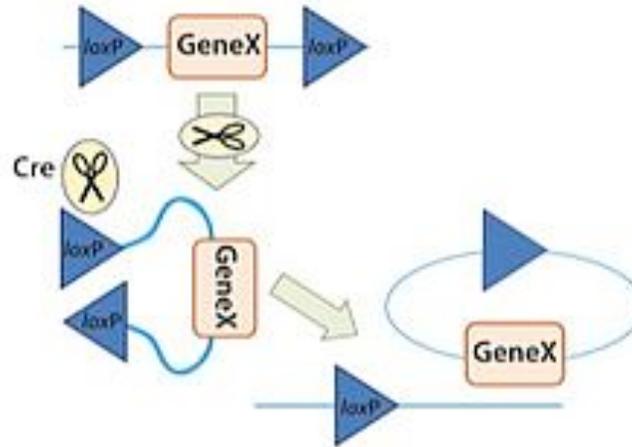


Image by Larissa Halw

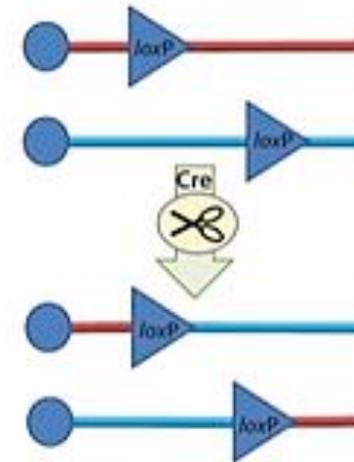
- Ativação ou inativação gênica

### Deletion



- Deleção gênica

### Translocation

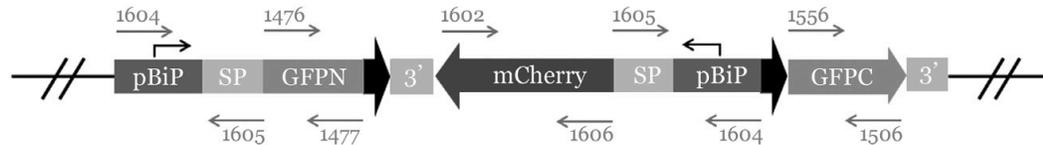


- Inserção gênica

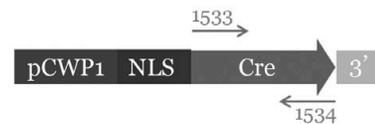
# Sistema Cre-loxP

## Exemplo de sistema utilizado em *Giardia*

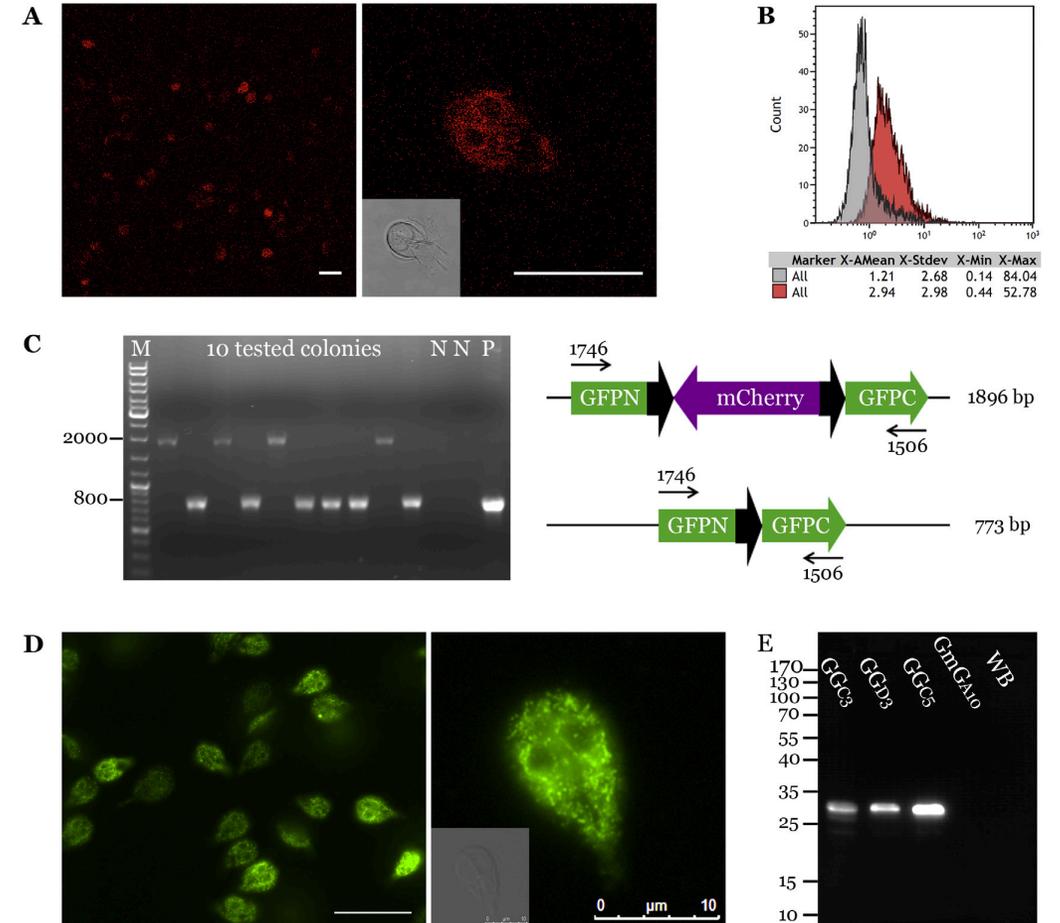
### A Construct GFPN-mCherryfl-GFPC



### B Construct NEO-Cre



### C Reconstruction of the interrupted GFP ORF



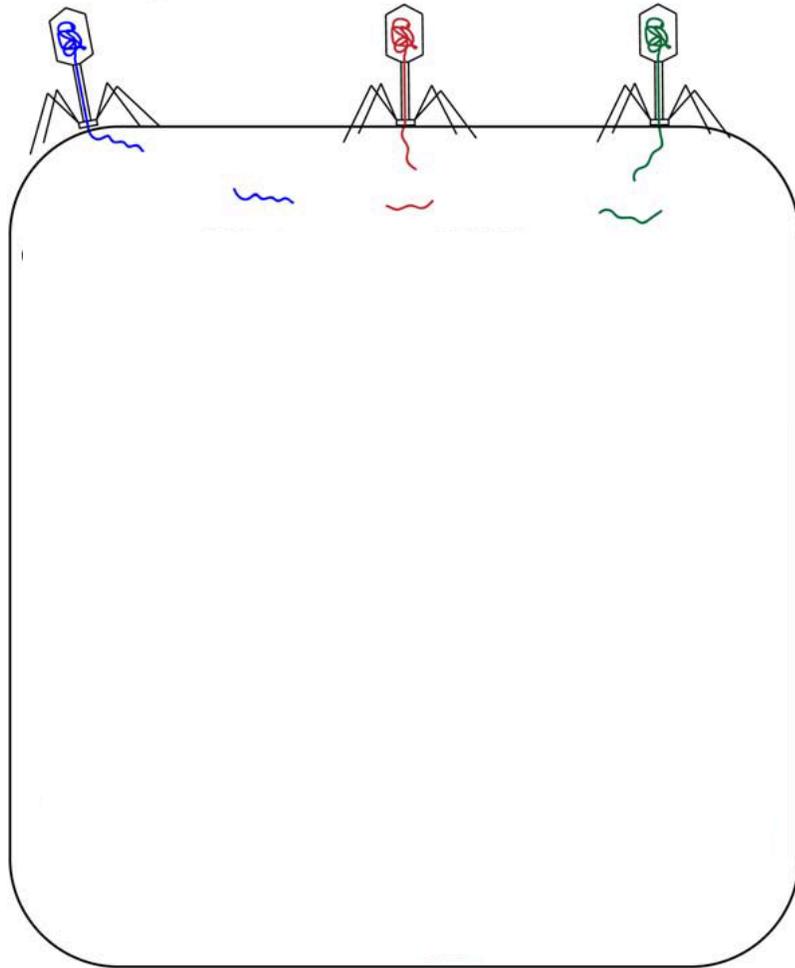
## Sistemas para edição gênica

Sistema CRISPR-Cas

# CRISPR-Cas: Parte do sistema imune bacteriano

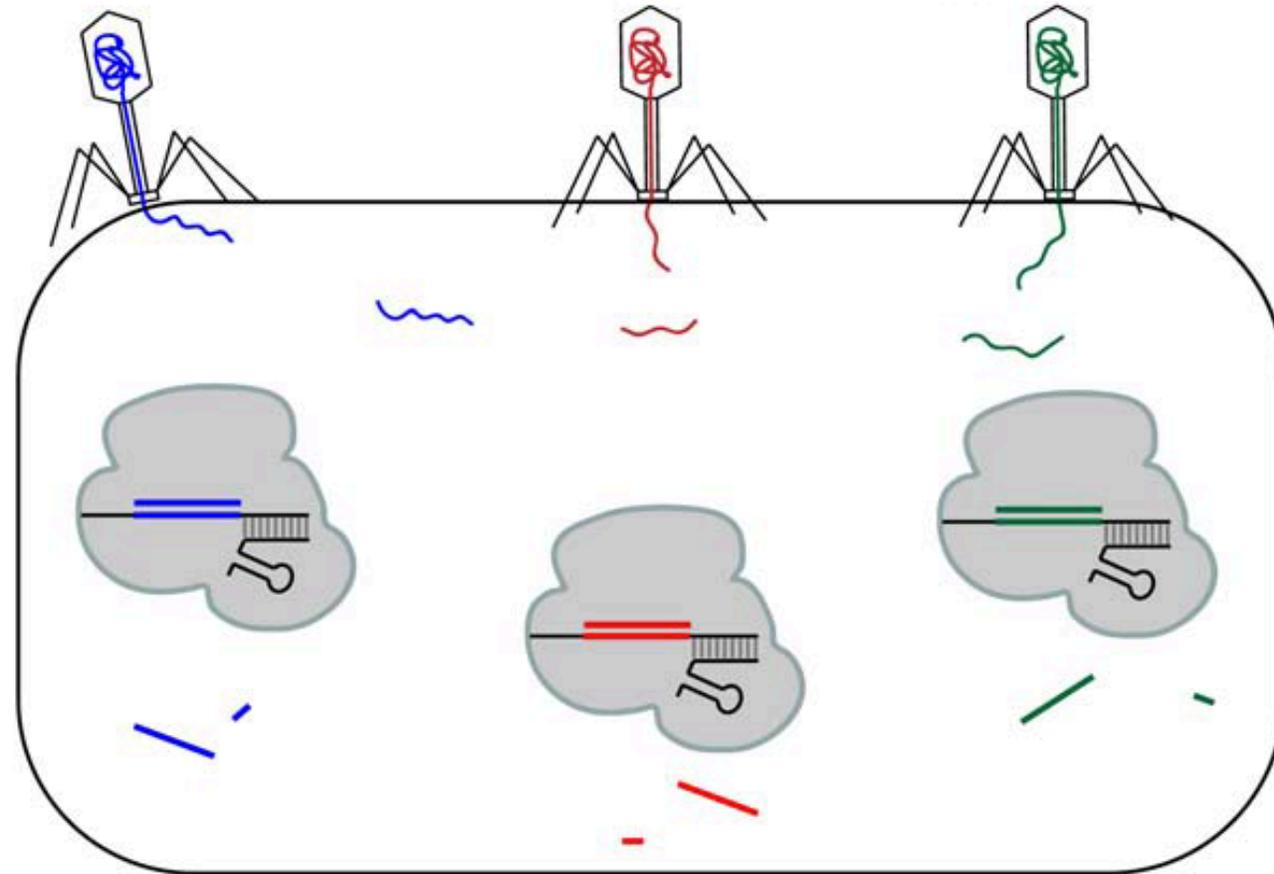
- CRISPR: **C**lustered **R**egularly **I**nterspaced **S**hort **P**alindromic **R**epeats  
(Repetições Palindrômicas Curtas Agrupadas e Regularmente Interespaçadas)
- Cas: **C**RISPR **a**ssociated proteins (Proteínas associadas ao CRISPR)
- Identificados pela primeira vez por análise genômica comparativa de *S. thermophilus*

# Como as bactérias adquirem memória imunológica?



- pre-crRNA – CRISPR RNA precursor
- tracr-RNA – trans-activating CRISPR RNA

# Como as bactérias utilizam a memória imunológica?



# Os diferentes tipos de sistema CRISPR-Cas

Divididos em duas grandes classes baseados na complexidade das proteínas associadas ao CRISPR:

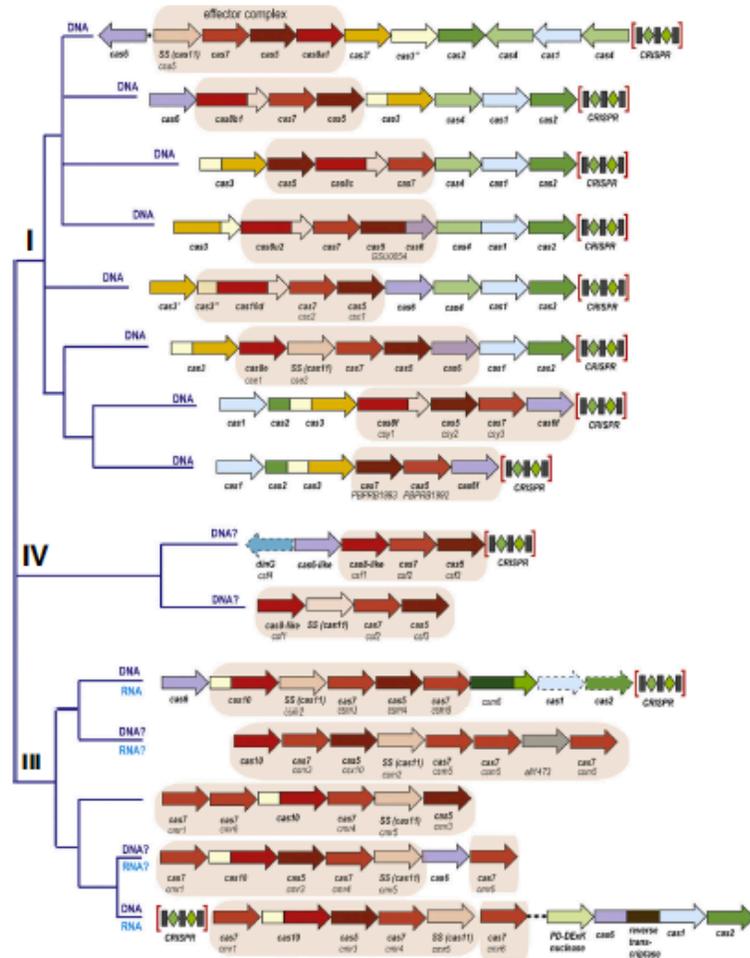
**Classe 1:** Cas é um complexo com múltiplas subunidades. Subdividida em tipos I, III e IV.

**Classe 2:** Cas é uma única proteína. Subdividida em tipos II, V e VI.

Os tipos são subdivididos em subtipos baseados na similaridade de sequência entre múltiplas proteínas Cas, na filogenia da Cas1 (a proteína Cas mais bem conservada), na organização do gene nos loci CRISPR-Cas e na própria estrutura do CRISPR

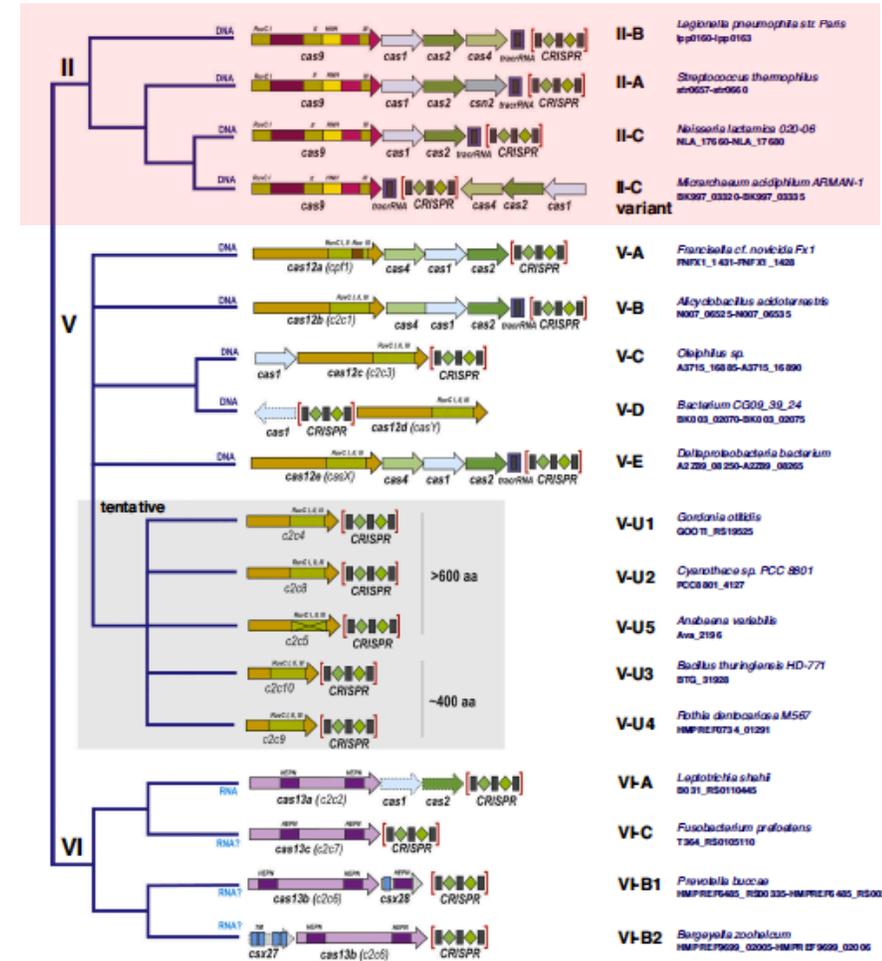
# Os diferentes tipos de sistema CRISPR-Cas

## Classe 1



- I-A *Archaeoglobus fulgidus*  
AF185, AF1870-AF1873
- I-B *Clostridium kluyveri*  
CKL\_2759-CKL\_2751
- I-C *Bacillus halodurans*  
BH0306-BH0342
- I-U *Geobacter sulfurreducens*  
GSU0051-GSU0054  
GSU0057-GSU0058
- I-D *Cyanotheca sp. 8802*  
Cys8802\_0627-  
Cys8802\_0620
- I-E *Escherichia coli* K12  
yjc9-yjgF
- I-F *Yersinia pseudotuberculosis*  
YPK\_1644-YPK\_1649
- I-F variant *Shewanella putrefaciens* CN-32  
Spulen32\_1819-Spulen32\_1823
- IV *Thioalkalivibrio* sp. K90mik  
TK90\_2493-TK90\_2703
- IV variant *Rhodococcus jostii* RHA1  
RHA1\_no10069-RHA1\_no10072
- III-A *Staphylococcus epidermidis*  
SERP2463-SERP2465
- III-D *Synechocystis* sp. 6803  
sl7067-sl7063
- III-C *Methanothermobacter*  
*thermautotrophicus*  
MTH329-MTH323
- III-B *Pyrococcus furiosus*  
PF1131-PP1134
- III-B variant *Martinomusa mediterranea* MMEL\_1  
Mama\_0669-Mama\_0677

## Classe 2



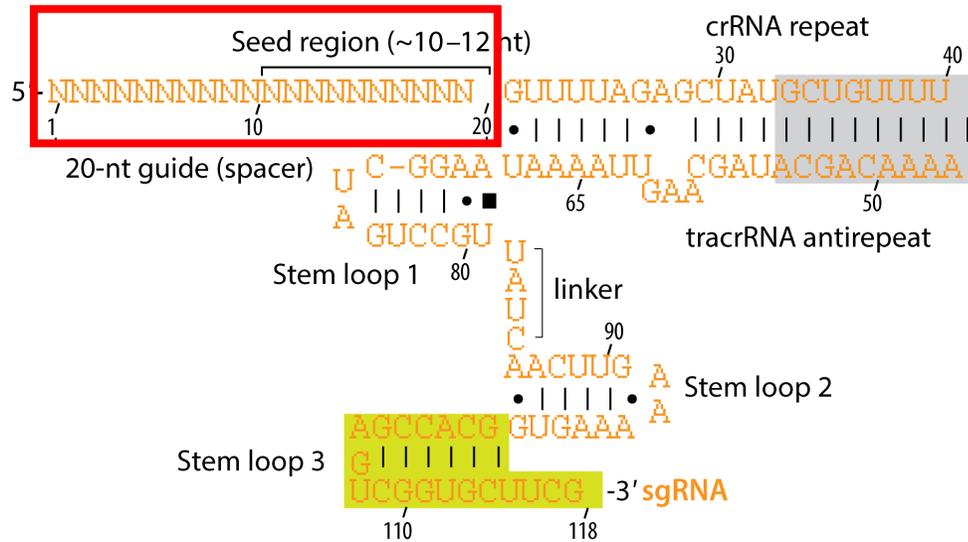
Cas9

# Componentes do sistema CRISPR-Cas9

- sgRNA: *single-guide* RNA – direciona a Cas9 para a sequência que será cortada
- Cas9: endonuclease que causa a quebra na dupla fita de DNA

# sgRNA: o pequeno guia

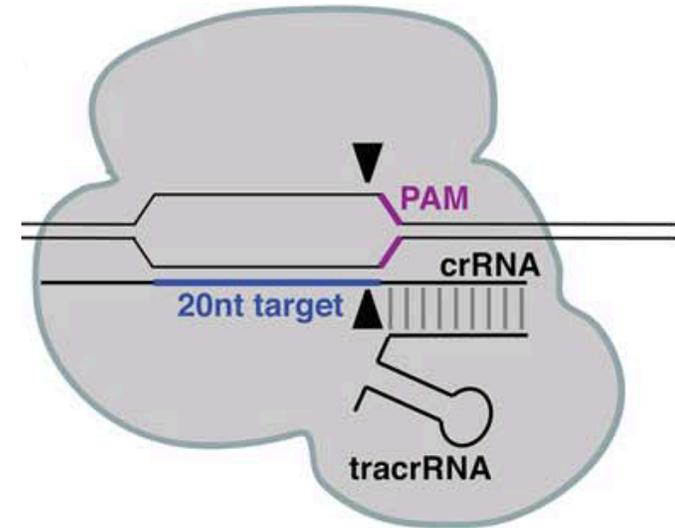
## Estrutura secundária do sgRNA



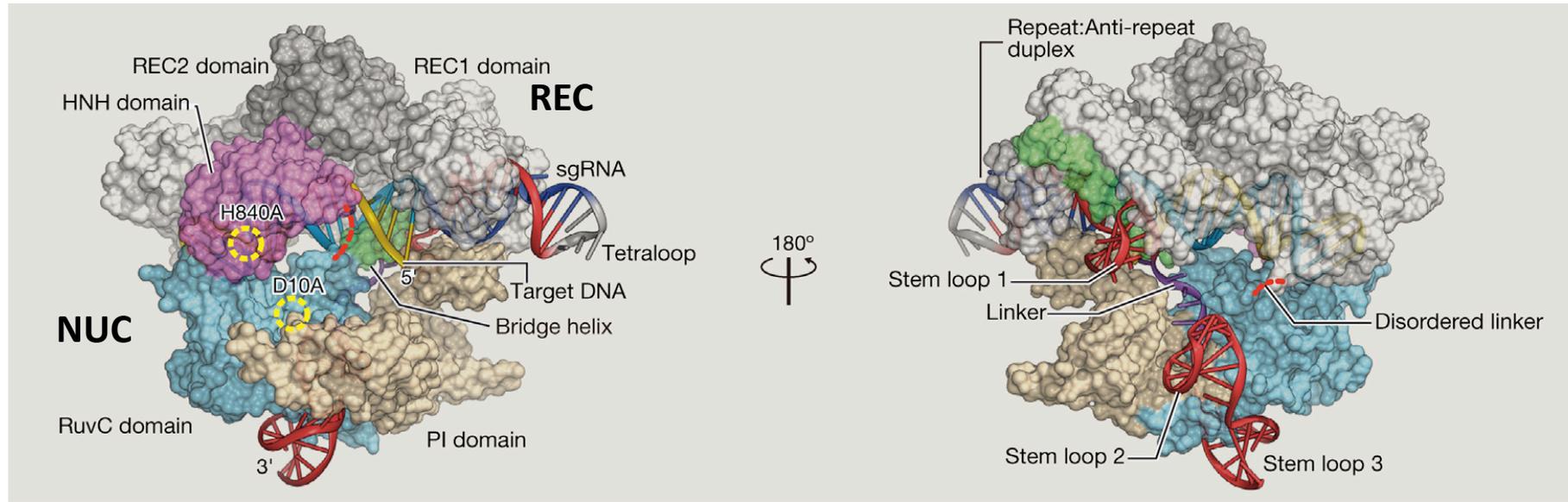
Motivo PAM: **P**roto-spacer **A**djacent **M**otifs (Motivo adjacente ao proto-espaçador)

Sequência PFS: **P**rotopspacer-**F**lanking **S**equence (Sequência flanquadora do proto-espaçador)

Corte cego em 3 bases antes do PAM



# Cas9: a "tesoura molecular"

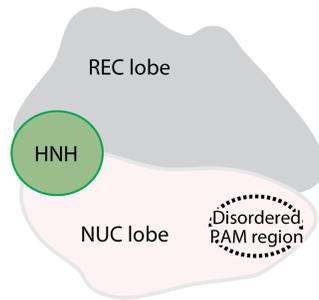


**REC e NUC:** lóbulos de REConhecimento e NUClease

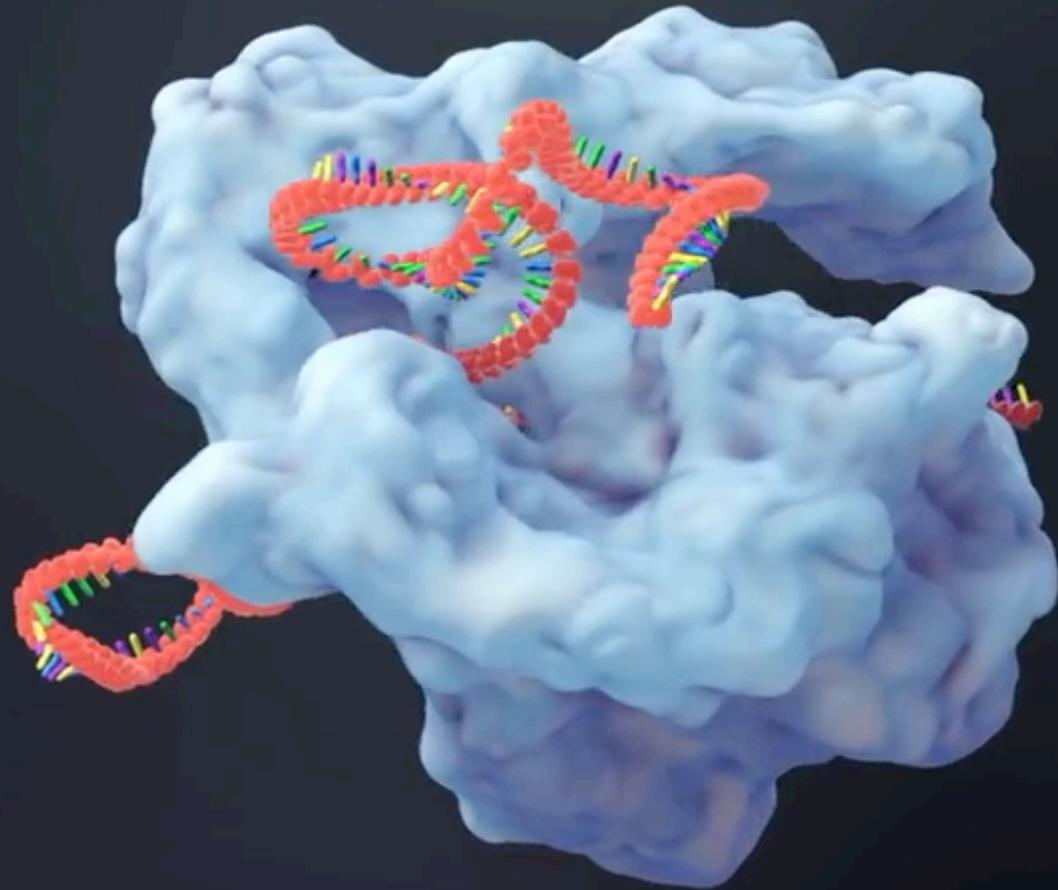
**HNH e RuvC:** domínios nuclease

**PI:** domínio de interação com o PAM

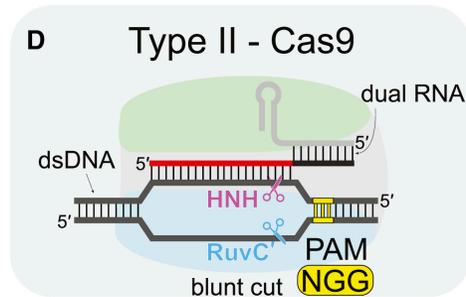
# Cas9: mecanismo de ação



Cas9 in the apo state  
(inactive)



# Os diferentes tipos de sistema CRISPR-Cas Classe 2

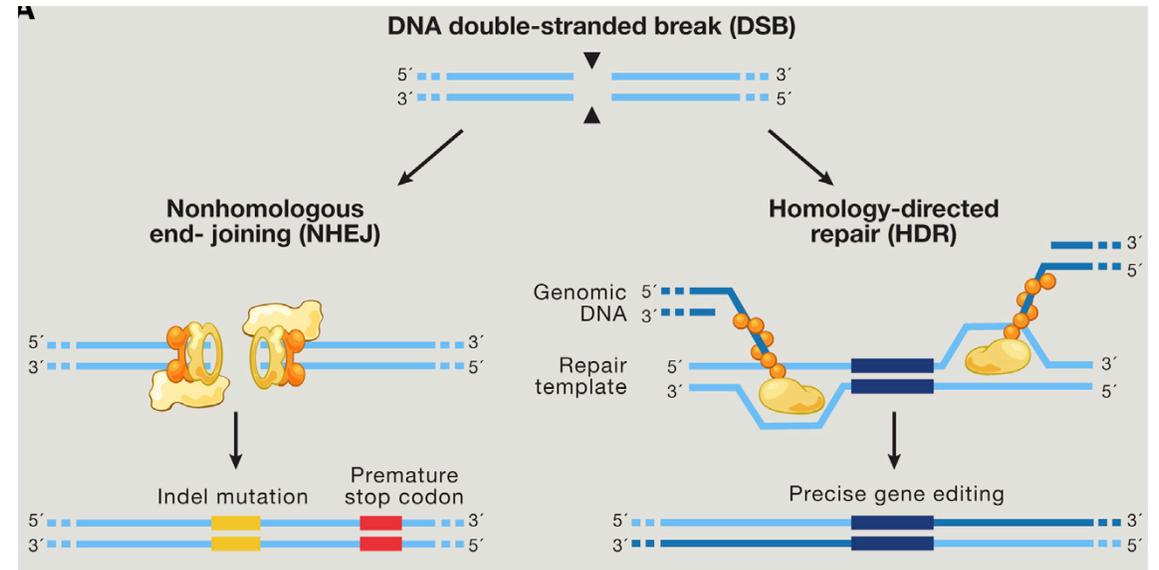
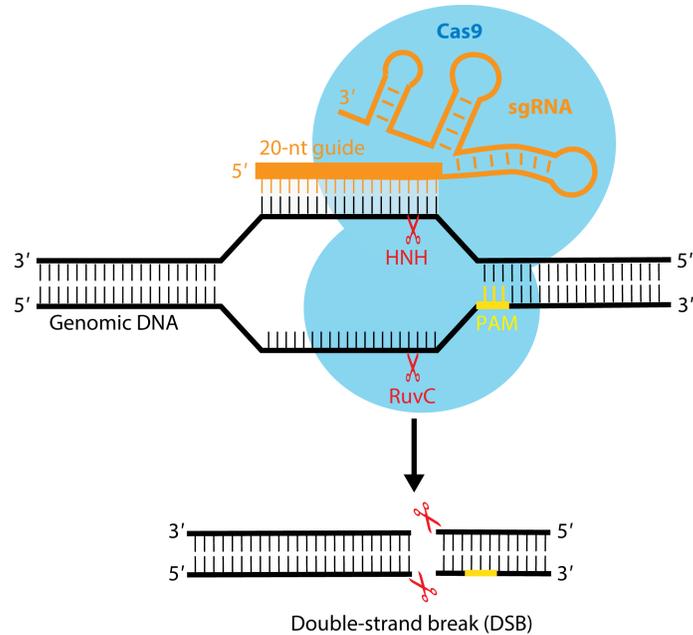


## Tipo II – Cas9

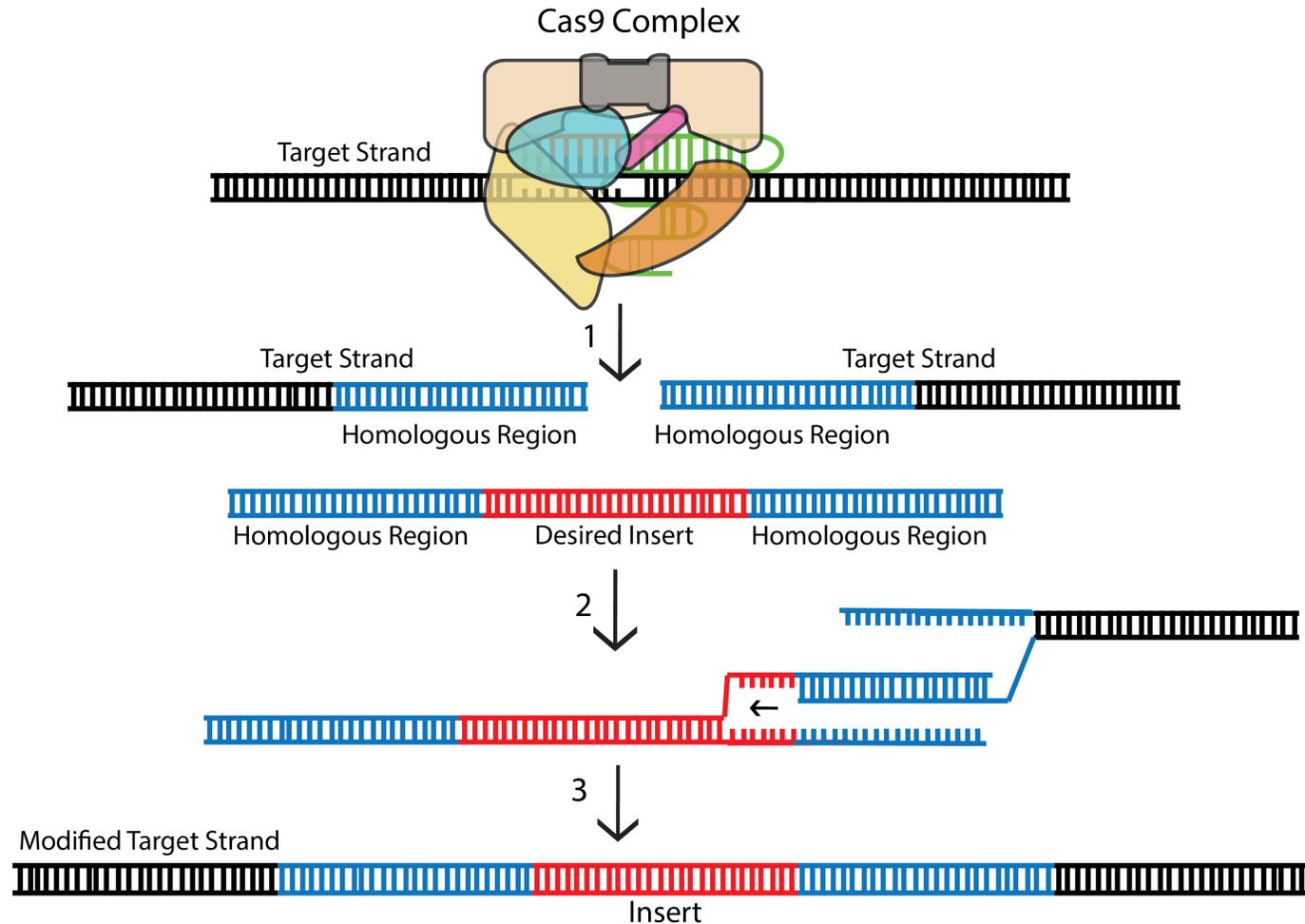
<i>Domínios nuclease</i>	RuvC, HNH
<i>Alvo</i>	dsDNA
<i>Estrutura do corte</i>	Corte cego
<i>Requer tcraRNA?</i>	Sim
<i>PAM</i>	3' GC-rich PAM

Como o sistema CRISPR-Cas pode ser utilizado para edição gênica?

# CRISPR-Cas9 e os mecanismos de reparo



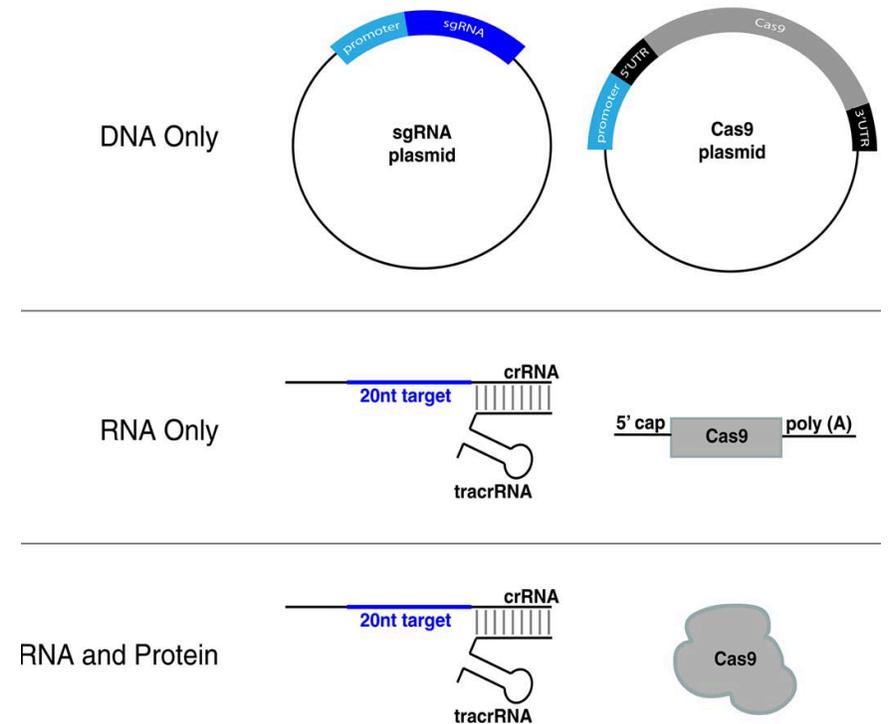
# CRISPR-Cas9 e HDR: a combinação perfeita



# Como editar o genoma utilizando o sistema CRISPR-Cas9:

- Disponibilizar para a célula:

1. Enzima Cas9
2. sgRNA
3. DNA molde para reparo desejado



# CRISPR na prática

# Protocolo – Geração das linhagens expressoras da Cas9

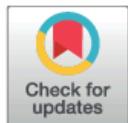
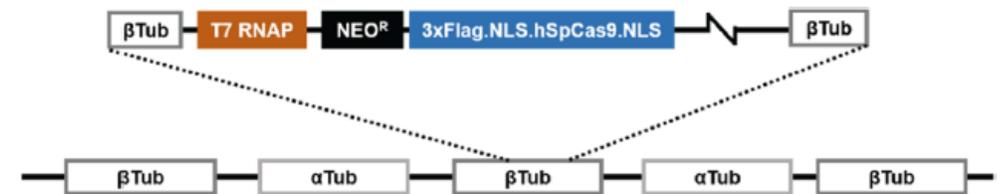
RESEARCH ARTICLE

Expanding the toolbox for *Trypanosoma cruzi*: A parasite line incorporating a bioluminescence-fluorescence dual reporter and streamlined CRISPR/Cas9 functionality for rapid *in vivo* localisation and phenotyping

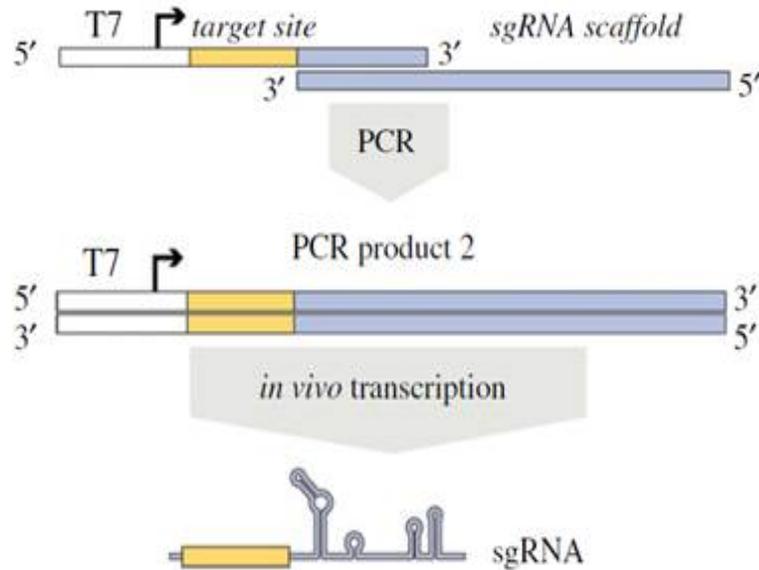
Fernanda Cristina Costa<sup>1,2</sup>, Amanda Fortes Francisco<sup>1</sup>, Shiromani Jayawardhana<sup>1</sup>, Simone Guedes Calderano<sup>1,3</sup>, Michael D. Lewis<sup>1</sup>, Francisco Olmo<sup>1</sup>, Tom Beneke<sup>4</sup>, Eva Gluenz<sup>4</sup>, Jack Sunter<sup>5</sup>, Samuel Dean<sup>4</sup>, John Morrison Kelly<sup>1\*</sup>, Martin Craig Taylor<sup>1\*</sup>

**1** Faculty of Infectious and Tropical diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, **2** Institute of Physics of São Carlos, University of São Paulo, São Carlos, Brazil, **3** Laboratory of Parasitology, Instituto Butantan, São Paulo, Brazil, **4** Sir William Dunn School of Pathology, University of Oxford, Oxford, United Kingdom, **5** Department of Biological and Medical Sciences, Oxford Brookes University, Oxford, United Kingdom

**A**



# Protocolo – Geração dos sgRNAs



**T7:** região promotora de T7 RNA polimerase.

**Target site:** região do gene alvo reconhecida pelo sgRNA.

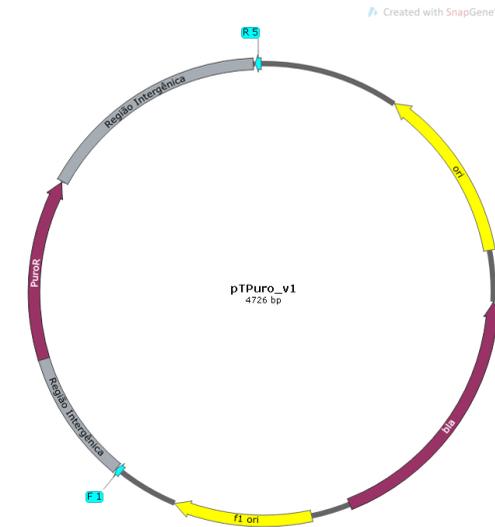
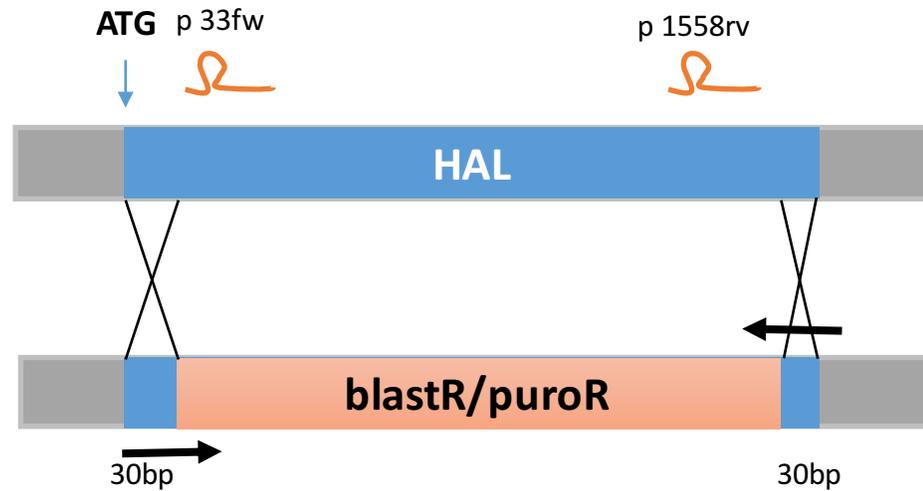
**Foward:** específico para cada região alvo de quebra de dupla fita

**GAAATTAATACGACTCACTATAGGCCTCTTCGACCCGATCATCG****GTTTTAGAGCTAGAAATAGC**

**Reverse:** igual para todas as reações

**AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTAACTTGCTA  
TTTCTAGCTCTAAAAC**

# Protocolo – Geração dos cassetes de recombinação



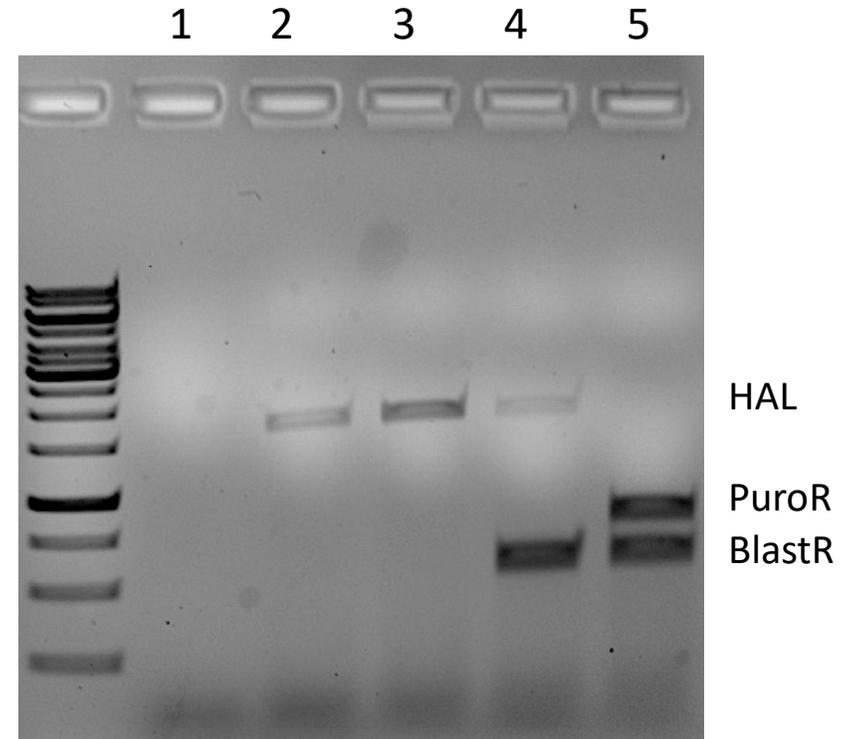
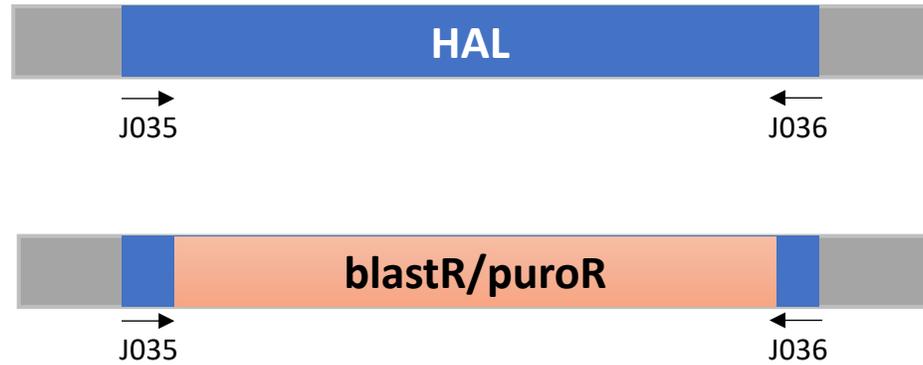
**Fwd-puro** *CTGAGGGTTATCCTTGACGGCTGCTCCTTGACGCCATGACTGAATACAAGCCAACGGTTC*

**Rev-puro** *TCACATCTTGATTCAGCTCAAATGGTTTCTTAACCTTAGGCTCCCGGCTTACGTGTCAT*

**Fwd-blast** *CTGAGGGTTATCCTTGACGGCTGCTCCTTGACGCCATGCCTTTGTCTCAAGAAGAATCCA*

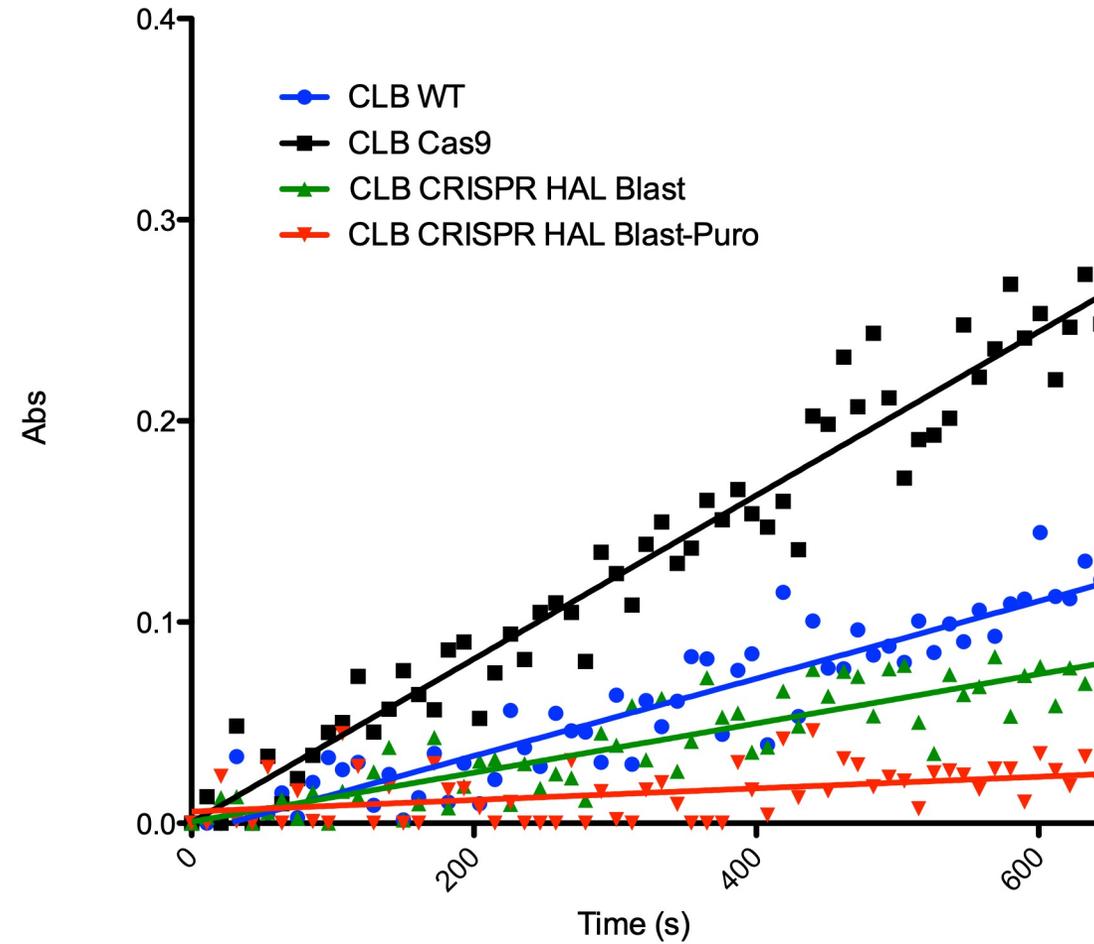
**Rev-blast** *TCACATCTTGATTCAGCTCAAATGGTTTCTTAACCTTAGCCCTCCACACATAACCAGA*

# PCR controle de deleção 2

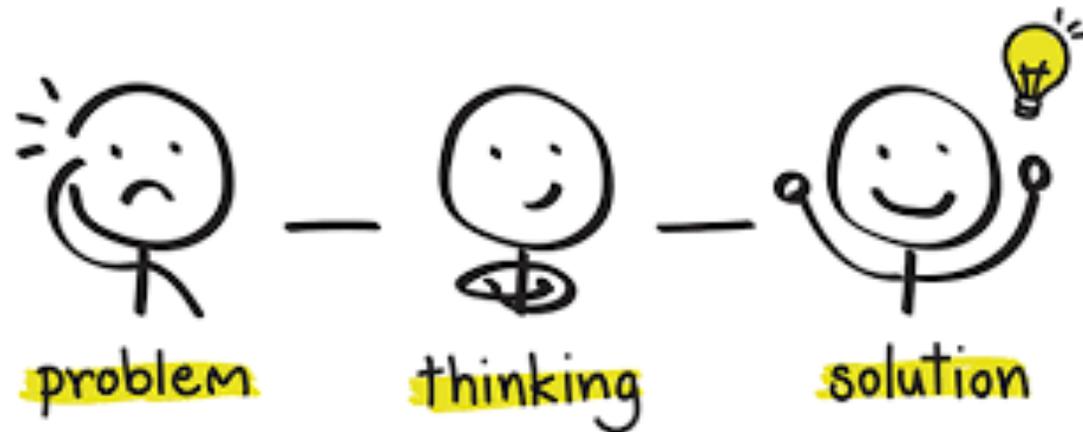


1. C-
2. TcCLB wt
3. TcCLB Cas9
4. TcCLB Cas9 CRISPR HAL blast
5. TcCLB Cas9 CRISPR HAL blast-puro

# Ensaio de atividade enzimática



É hora de pensar...



1. Escolha o seu parasita favorito e justifique
2. Escolha o seu gene putativo favorito e justifique ([veupathdb.org](http://veupathdb.org))
3. Estabeleça a hipótese para a função desse gene
4. Escolha duas técnicas que foram discutidas hoje (ou outras!)
5. Explique como você utilizaria essas técnicas para investigar a função desse gene e os resultados esperados
6. Coloque tudo no papel – máximo de 10 páginas!
7. Não esqueça as referências.

 **VEuPathDB**  
*Eukaryotic Pathogen, Vector & Host Informatics Resources*

 **AmoebaDB**  
*Amoeba Informatics Resources*

 **CryptoDB**  
*Cryptosporidium Informatics Resources*

 **FungiDB**  
*Fungal & Oomycete Informatics Resources*

 **GiardiaDB**  
*Giardia Informatics Resources*

 **HostDB**  
*Pathogen Host Informatics Resources*

 **MicrosporidiaDB**  
*Microsporidia Informatics Resources*

 **OrthoMCL DB**  
*Ortholog Groups of Protein Sequences*

 **PiroplasmaDB**  
*Piroplasma Informatics Resources*

 **PlasmoDB**  
*Plasmodium Informatics Resources*

 **ToxoDB**  
*Toxoplasma Informatics Resources*

 **TrichDB**  
*Trichomonas Informatics Resources*

 **TriTrypDB**  
*Kinetoplastid Informatics Resources*

 **VectorBase**  
*Bioinformatics Resources for  
Invertebrate Vectors of Human Pathogens*

Janaina de Freitas Nascimento

[janainafn@usp.br](mailto:janainafn@usp.br)

 @Jana\_Freitas

