



Desarrollo de una herramienta software de identificación de secuencias patógenas candidatas para el diseño de RNAs guía en el sistema SHERLOCK de diagnóstico.

 **UOC** **Universitat Oberta de Catalunya**

*Samantha López – Trabajo de Fin de Máster
Máster Universitario en Bioestadística y Bioinformática
Profesor consultor: Amadís Pagès*

Contenidos

- CRISPR.....	02
- El sistema CRISPR-Cas9.....	03
- El sistema CRISPR-Cas13a.....	04
- SHERLOCK.....	05
- Herramientas de diseño de secuencias CRISPR.....	06
- Herramienta: objetivos.....	07
- Herramienta: esquema general.....	08
- Búsqueda exhaustiva de secuencias comunes.....	09
- Búsqueda exhaustiva de secuencias particulares.....	10
- BLAT.....	11
- Acople con CRISPR-RT para el diseño de crRNAs.....	12
- Virus del Papiloma Humano.....	13
- Demostración.....	14
- Resultado.....	15
- Aplicación.....	16
- Conclusiones.....	17
- Discusión.....	18
- Referencias principales.....	20

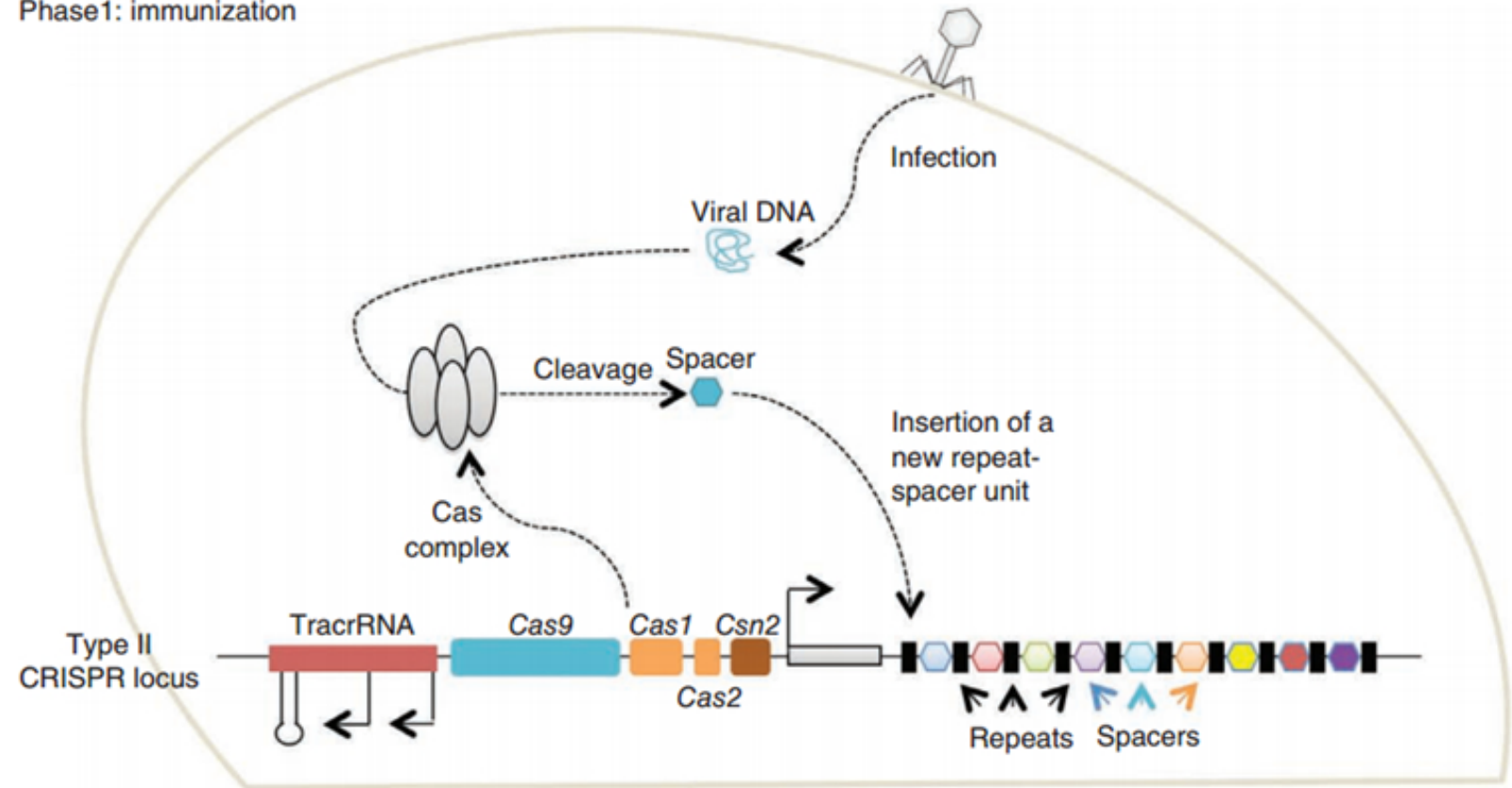
Desarrollo de una herramienta software de identificación de secuencias patógenas candidatas para el diseño de RNAs guía en el sistema SHERLOCK de diagnóstico.

- Clustered Regularly Interspaced Short Palindromic Repeats

- Inmunidad Adaptativa Bacteriana

- Familia de nucleasas Cas

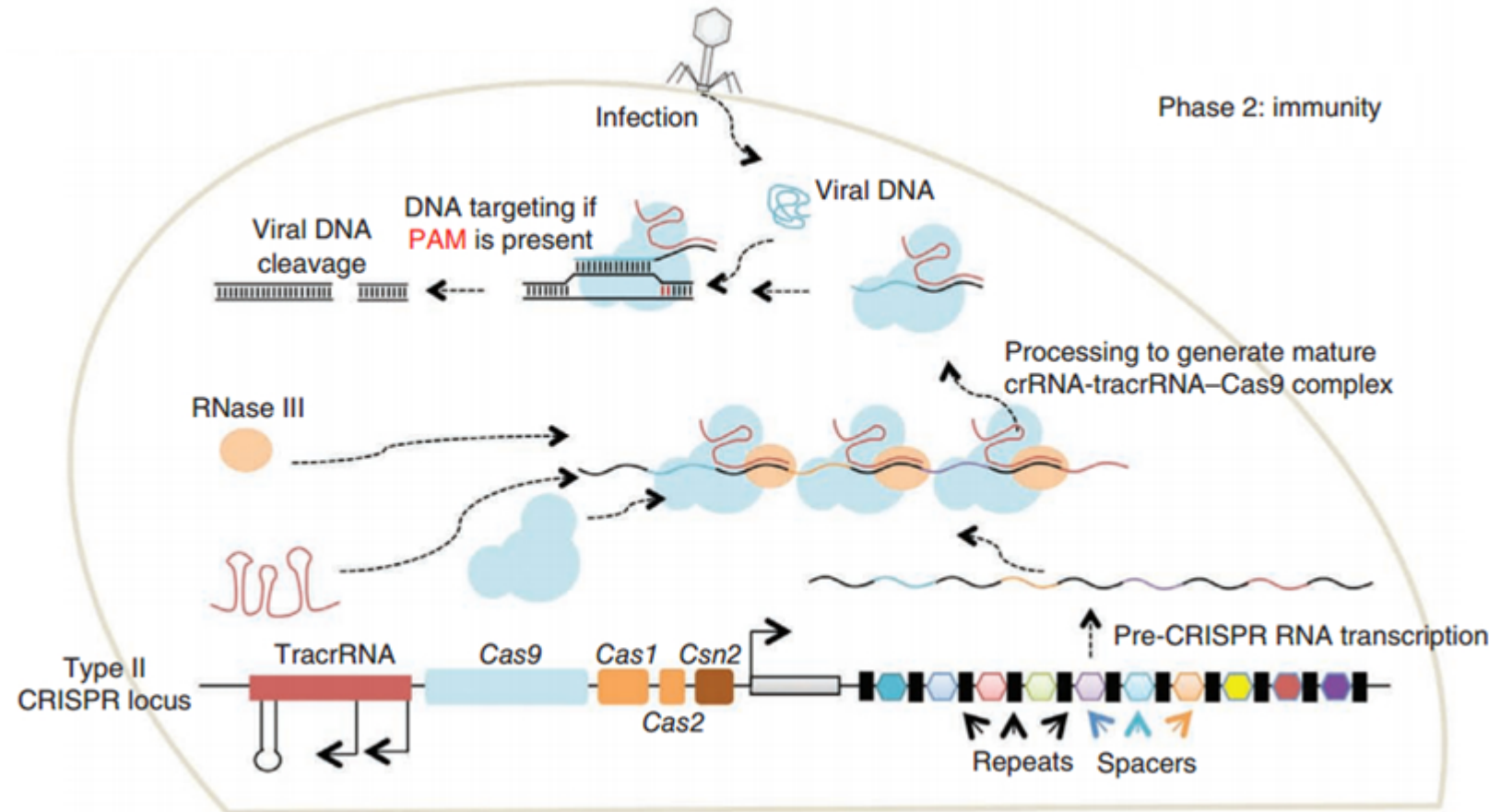
Phase 1: immunization



SUSTC-Shenzhen. Human CRISPR/Cas System against HIV with gRNA delivered by A-B toxin-based shuttle. *iGEM*. 2014.

Desarrollo de una herramienta software de identificación de secuencias patógenas candidatas para el diseño de RNAs guía en el sistema SHERLOCK de diagnóstico.

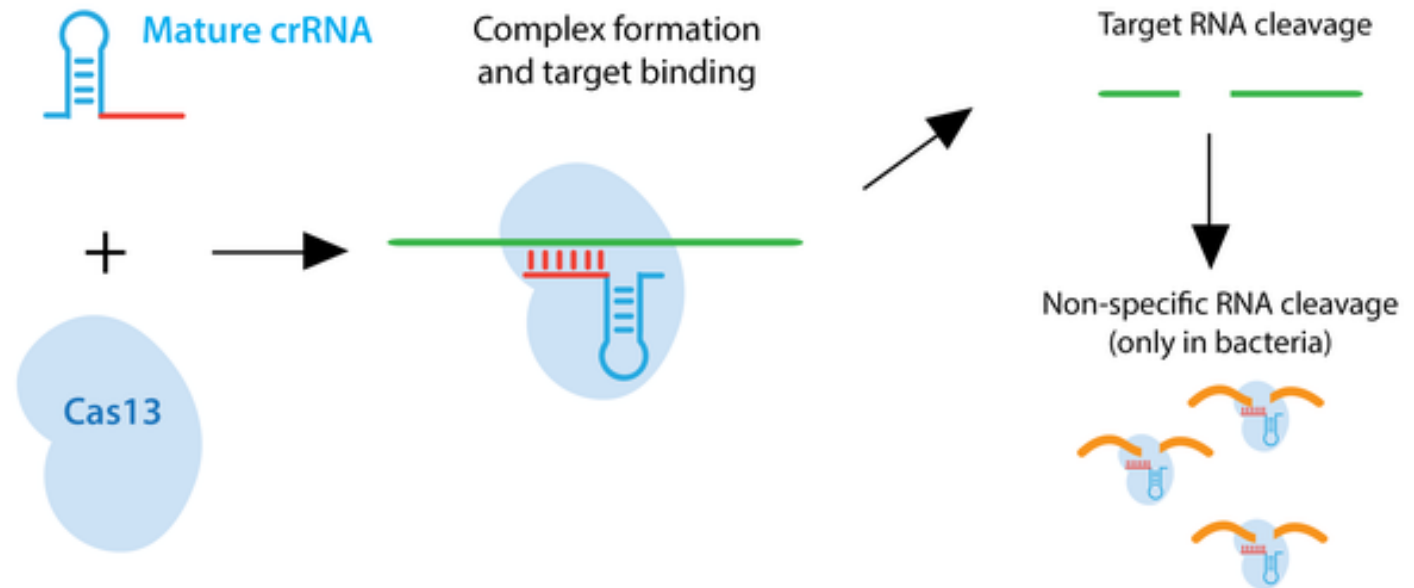
- Complejo Cas9 + crRNA + trRNA
- Aplicación *in vitro*: sgRNA sintético
- RNA diana: 20nt + secuencia PAM



SUSTC-Shenzhen. Human CRISPR/Cas System against HIV with gRNA delivered by A-B toxin-based shuttle. *iGEM*. 2014.

Desarrollo de una herramienta software de identificación de secuencias patógenas candidatas para el diseño de RNAs guía en el sistema SHERLOCK de diagnóstico.

- Complejo Cas13a + crRNA
- RNA diana + PFS
- Actividad RNAsa promiscua activa de Cas13a

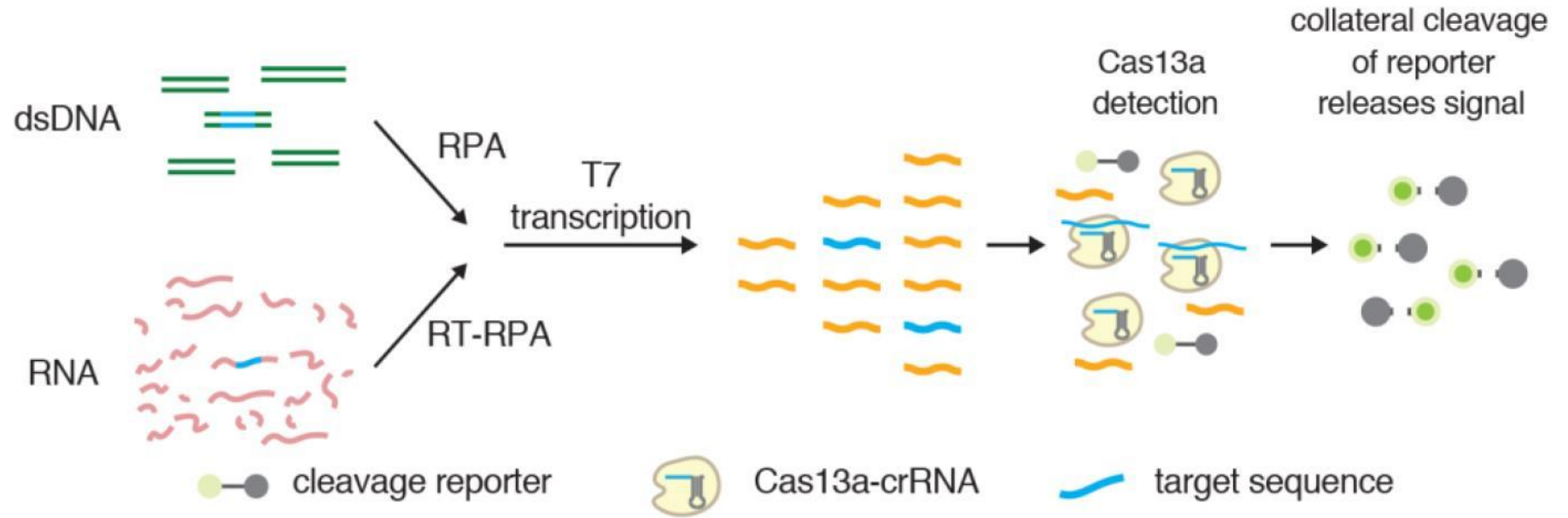


<https://www.addgene.org/crispr/guide/> - Addgene CRISPR guide

Desarrollo de una herramienta software de identificación de secuencias patógenas candidatas para el diseño de RNAs guía en el sistema SHERLOCK de diagnóstico.

- Detección de infecciones en *Point of Care*

- Actividad Cas13a sobre *beacons* fluorescentes



Gootenberg, J., Abudayyeh, O., Lee, J., Essletzbichler, P., Dy, A., Joung, J., Verdine, V., Donghia, N., Daringer, N., Freije, C., Myhrvold, C., Bhattacharyya, R., Livny, J., Regev, A., Koonin, E., Hung, D., Sabeti, P., Collins, J. and Zhang, F. (2017). Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science*, 356(6336), pp.438-442.

Desarrollo de una herramienta software de identificación de secuencias patógenas candidatas para el diseño de RNAs guía en el sistema SHERLOCK de diagnóstico.

- Diseño de sgRNAs o crRNAs

- Matching con RNA / DNA diana

- CasFinder, CRISPR Design, CHOPCHOP

CRISPR DESIGN / [Help](#) [Forum](#)

2018/05/16 BACK ONLINE Greetings CRISPR ENGINEERS!... [show more](#)

Submit

[Batch Mode](#) [Single Sequence](#)

Submit a single sequence for CRISPR design and analysis.

search name *

email address *

sequence type other region (23-500 nt) [[demo](#)]
 unique genomic region (23-500 nt) [[demo](#)]

target genome human (hg38)
 mouse (mm10)
 zebrafish (danRer11)
 c. elegans (ce10)
 fly (dm6)
 rabbit (oryCun2)
 pig (susScr11)
 possum (monDom5)

sequence

<http://crispr.mit.edu/> - CRISPR Design: Interfaz web



CHOPCHOP

Target

RefSeq/ENSEMBL/gene ID or genomic coordinates.

In

[Add new species.](#)

Using

Change default PAM and guide length in Options.

[Fasta Target](#) [Options](#) [Reset Options](#)

[Find Target Sites!](#)

<http://chopchop.cbu.uib.no/> - CHOPCHOP: Interfaz web



SHERLOCK

- Secuencias comunes y particulares
- **No-presencia** en el huésped
- **crRNAs** a partir de ambos tipos de secuencias
- Aptos para diseño de test **SHERLOCK**



Cas13a

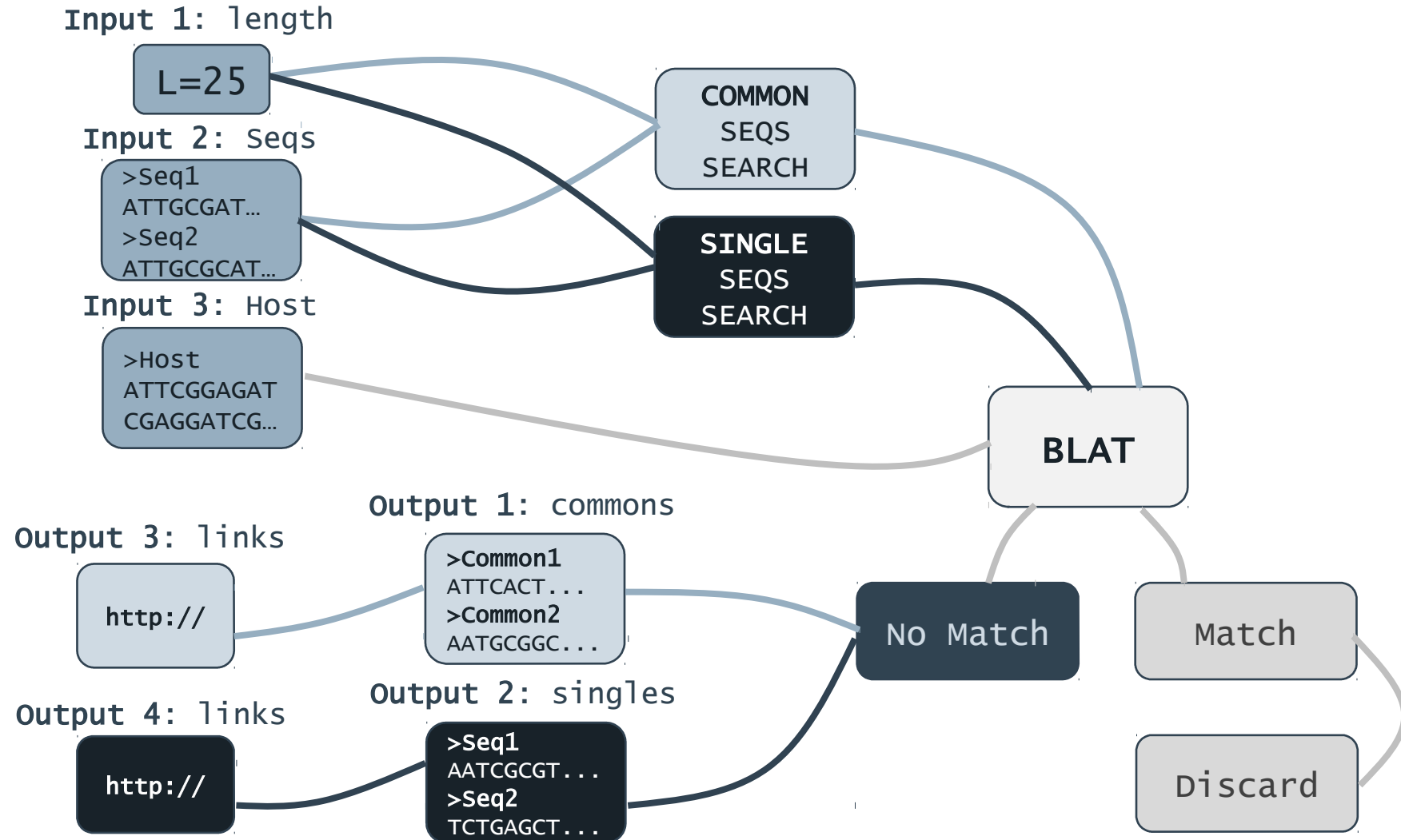
Desarrollo de una herramienta software de identificación de secuencias patógenas candidatas para el diseño de RNAs guía en el sistema SHERLOCK de diagnóstico.

INPUTS

- Longitud
- Secuencias patógenas
- Huésped

OUTPUTS

- Secuencias comunes
- Secuencias particulares
- Enlaces CRISPR-RT



Búsqueda exhaustiva:
secuencias comunes

- Escisión (+PFS)

- Búsqueda

- Comprobación de PFSs

Length = 7

>Seq1

TATAGAGTCATGCTAGGCGGATCGGCTAGGCTAGGCT

TATAGAGTCATGCTAGGCGGATCGGCTAGGCTAGGCT – 1st subseq (No match)

TATAGAGTCATGCTAGGCGGATCGGCTAGGCTAGGCT – 2nd subseq (No match)

TATAGAGTCATGCTAGGCGGATCGGCTAGGCTAGGCT – 3rd subseq (MATCH!)

>Seq2

AAGCTGGTCGAGCTGATAGAGTCACGCGCGTAGGCTA

>Seq3

TTCGCTAGCGCTGAGACTGAGCTTAGAGTCATCTATC

TAGAGTCA – PFS: A (Not G) – **TAGAGTCA IS A COMMON SEQ!**

Búsqueda exhaustiva: secuencias particulares

- Escisión (+PFS)
- Búsqueda
- Iteración para cada secuencia
- Comprobación de PFSs

Length = 7

>Seq1

```
TATAGAGTCATGCTAGGCGGATCGGCTAGGCTAGGCT
TATAGAGTCATGCTAGGCGGATCGGCTAGGCTAGGCT - No Match - Right PFS
TATAGAGTCATGCTAGGCGGATCGGCTAGGCTAGGCT - Match (Seq 2)
TATAGAGTCATGCTAGGCGGATCGGCTAGGCTAGGCT - Match (Seq 2, Seq 3)
```

>Seq2

```
AAGCTGGTCGAGCTGATAGAGTCACGCGCGTAGGCTA
AAGCTGGTCGAGCTGATAGAGTCACGCGCGTAGGCTA - No Match - Right PFS
AAGCTGGTCGAGCTGATAGAGTCACGCGCGTAGGCTA - No Match - Right PFS
AAGCTGGTCGAGCTGATAGAGTCACGCGCGTAGGCTA - No Match - Wrong PFS
```

>Seq3

```
TTCGCTAGCGCTGAGACTGAGCTTAGAGTCATCTATC
TTCGCTAGCGCTGAGACTGAGCTTAGAGTCATCTATC - No Match - Wrong PFS
TTCGCTAGCGCTGAGACTGAGCTTAGAGTCATCTATC - No Match - Right PFS
TTCGCTAGCGCTGAGACTGAGCTTAGAGTCATCTATC - No Match - Wrong PFS
```

Desarrollo de una herramienta software de identificación de secuencias patógenas candidatas para el diseño de RNAs guía en el sistema SHERLOCK de diagnóstico.

- BLAT: Blast-like alignment tool

- Alineamiento pareado local

- Se descartan aquellas presentes en el huésped

The screenshot shows the Ensembl BLAT/BLAT search interface. At the top, there is a navigation bar with the Ensembl logo and links for BLAST/BLAT, VEP, Tools, BioMart, and More. A search box for 'Search Human...' is located in the top right corner. Below the navigation bar, a 'Web Tools' menu is open, showing 'BLAST/BLAT' as the selected option. The main content area is titled 'BLAST/BLAT search' and contains a message: 'BLAST/BLAT for Human GRCh37. If you are looking for BLAST/BLAT for Human GRCh37, please go to GRCh37 website.' Below this message, there is a 'Sequence data' section with a text input field and a button to 'Or upload sequence file'. The input field contains the text 'Maximum of 30 sequences (type in plain text, FA'. Below the input field, there are radio buttons for 'DNA' (selected) and 'Protein'. At the bottom of the page, there is a 'Search against:' label.

http://www.ensembl.org/Homo_sapiens/Tools/Blast?db=core - BLAT: Interfaz web

- De DNA a RNA

- Generación de enlaces a CRISPR-RT

Commons :

>Common1

TAGAGTCA -> UAGAGUCA -> <http://www...UAGAGUC...PFS=A>

Singles:

>Seq1

TATAGAGT -> UAUAGAGU -> <http://www...UAUAGAG...PFS=U>

>Seq2

AAGCTGGT -> AAGCUGGU -> <http://www...AAGCUGG...PFS=U>

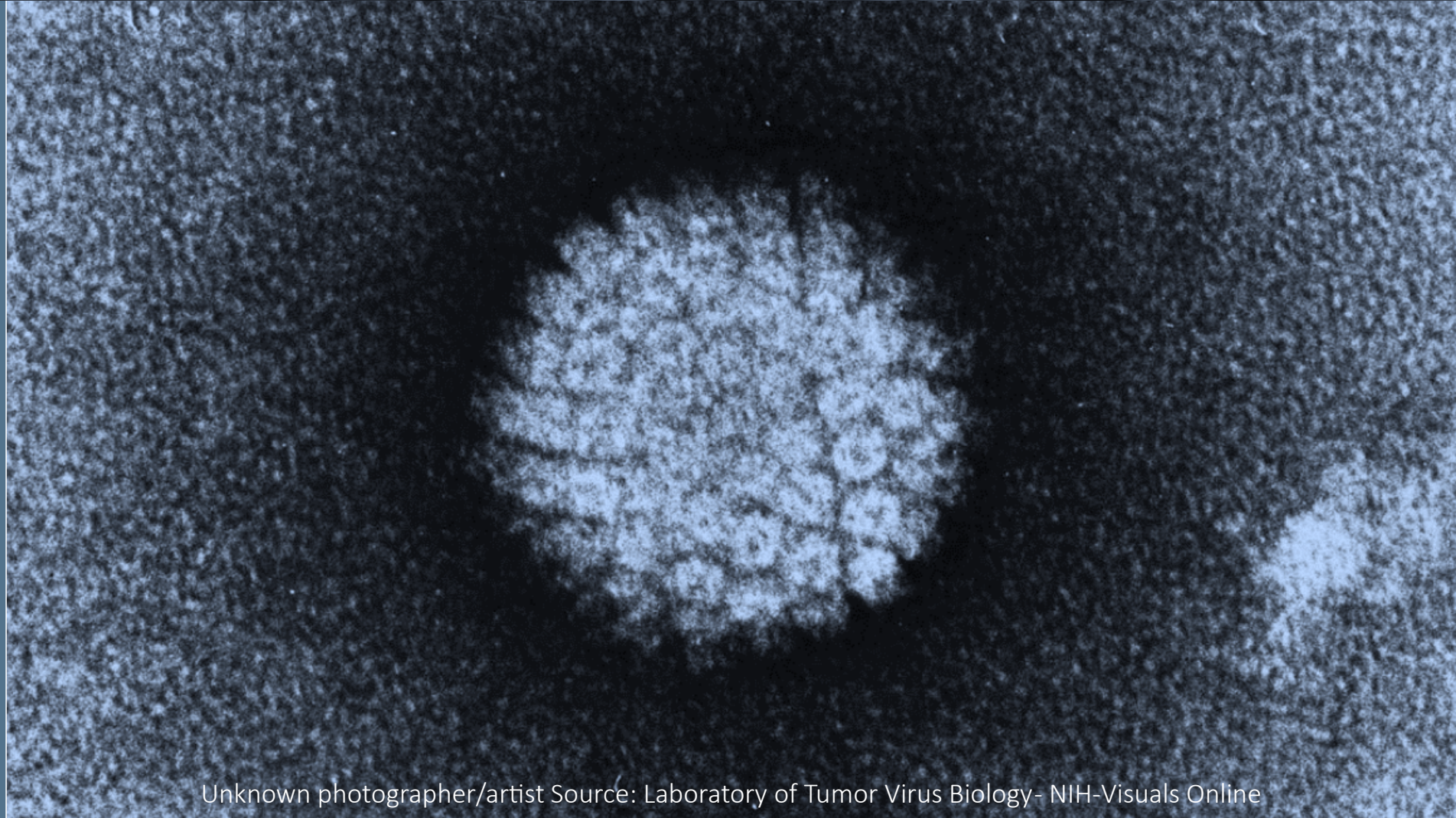
AGCTGGTC -> AGCUGGUC -> <http://www...AGCUGGU...PFS=C>

>Seq3

TCGCTAGC -> UCGCUAGC -> <http://www...UCGCUAG...PFS=C>

Desarrollo de una herramienta software de identificación de secuencias patógenas candidatas para el diseño de RNAs guía en el sistema SHERLOCK de diagnóstico.

- **HPV** (*Human Papillomavirus*)
- **Cepas** 4, 5, 9 (no oncogénicas), 16 y 18 (sí oncogénicas)
- **Objetivo:** crRNAs



Unknown photographer/artist Source: Laboratory of Tumor Virus Biology- NIH-Visuals Online

Desarrollo de una herramienta software de identificación de secuencias patógenas candidatas para el diseño de RNAs guía en el sistema SHERLOCK de diagnóstico.

DEMO

- Secuencias cortas de HPV 4, 5, 9, 16 y 18
- Fragmento de chr1 de *H.sapiens*
- Longitud outputs: 7 nt (+PFS)



RESULTADO

- Secuencia de 20 nt (+PFS)
- Particular de HPV16
- Enlace a CRISPR-RT

ID + Seq (FASTA format):

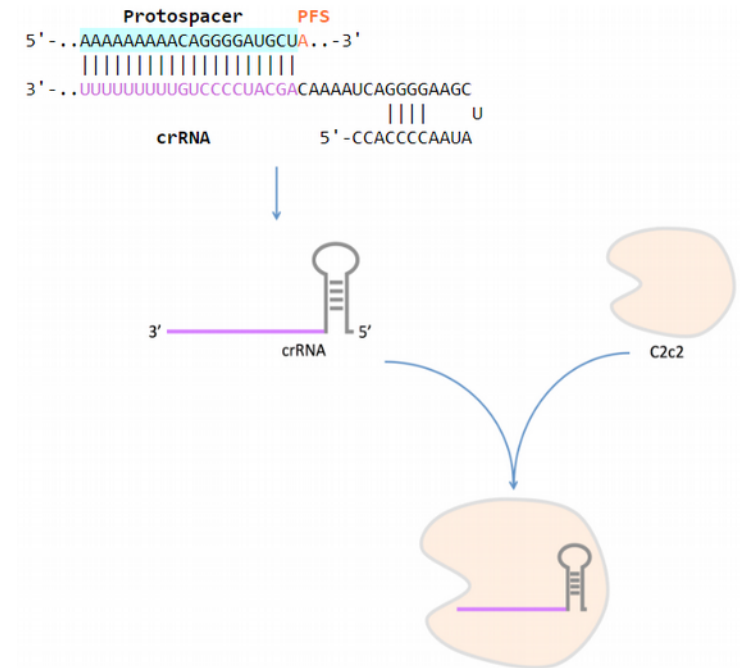
>JN565303.1 Human papillomavirus type 16 strain ZG01-258, complete genome

```
AAAAAAAAACAGGGGATGCTA
```

CRISPR-RT link:

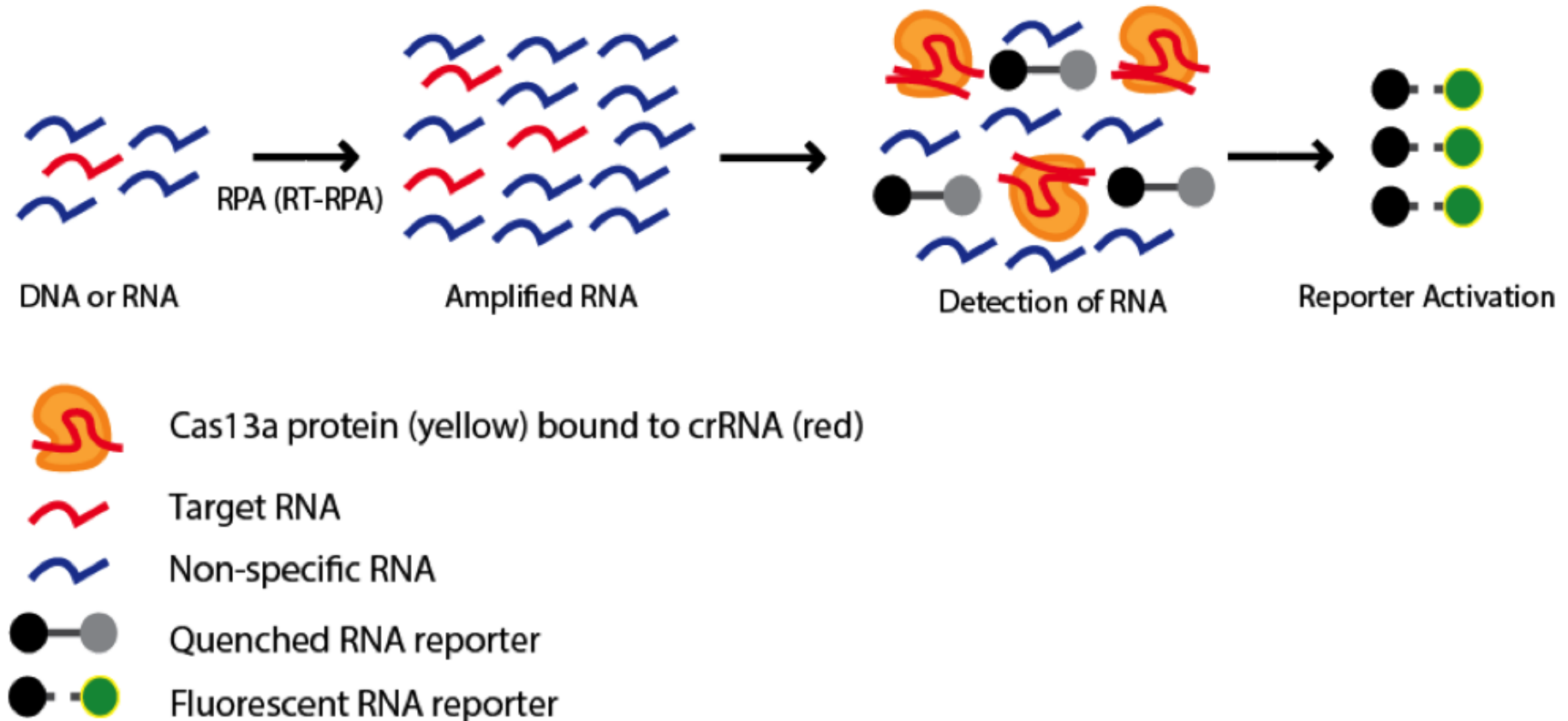
>JN565303.1 Human papillomavirus type 16 strain ZG01-258, complete genome

http://bioinfolab.miamioh.edu/CRISPR-RT/proc/crRNA_design.php
seq=AAAAAAAAACAGGGGAUGCUA&gRNA=AAAAAAAAACAGGGGAUGCU&PFS=A



Desarrollo de una herramienta software de identificación de secuencias patógenas candidatas para el diseño de RNAs guía en el sistema SHERLOCK de diagnóstico.

- Síntesis de crRNA con la secuencia obtenida
- Cas13a + crRNA + *beacons*
- Muestra biológica
- Detección cepa oncogénica



McDade, J. (2018). *CRISPR 101: Targeting RNA with Cas13a (C2c2)*. [online] Blog.addgene.org. Available at: <http://blog.addgene.org/crispr-101-targeting-rna-with-cas13a-c2c2> [Accessed 5 Mar. 2018]

Desarrollo de una herramienta software de identificación de secuencias patógenas candidatas para el diseño de RNAs guía en el sistema SHERLOCK de diagnóstico.

- Rapidez
- Precisión
- Bajo coste
- Sencillez
- Máxima especificidad
- Mínimo error
- *Point of Care*



Mejoras futuras

- Pruebas con mayor número de cepas
- Sistemas de puntuación
- Interfaz
- Validación experimental

Future...

- Trials:

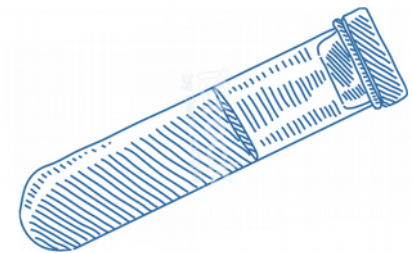
- >HPV1
- >HPV2
- ...
- >HPV103
- >HPV104
- ...

- Interface:

- Scoring:

- >Seq1
- TATAGAGT - 12 pt
- TGGTGCCC - 4 pt
- >Seq2
- AAGCTGGT - 6 pt
- AGCTGGTC - 3 pt
- >Seq3
- TCGCTAGC - 15 pt
- AGGCTAGC - 2 pt

- Final validation:





Gracias

- [1] Gootenberg, J., Abudayyeh, O., Lee, J., Essletzbichler, P., Dy, A., Joung, J., Verdine, V., Donghia, N., Daringer, N., Freije, C., Myhrvold, C., Bhattacharyya, R., Livny, J., Regev, A., Koonin, E., Hung, D., Sabeti, P., Collins, J. and Zhang, F. (2017). Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science*, 356(6336), pp.438-442.
- [2] Hsu, P., Lander, E. and Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell*, 157(6), pp.1262-1278.
- [3] Ran, F., Hsu, P., Wright, J., Agarwala, V., Scott, D. and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11), pp.2281-2308.
- [4] McDade, J. (2018). *CRISPR 101: Targeting RNA with Cas13a (C2c2)*. [online] Blog.addgene.org. Available at: <http://blog.addgene.org/crispr-101-targeting-rna-with-cas13a-c2c2> [Accessed 5 Mar. 2018].
- [5] Zhu, H., Richmond, E. & Liang, C. CRISPR-RT: A web application for designing CRISPR-C2c2 crRNA with improved target specificity. *Bioinformatics*.