


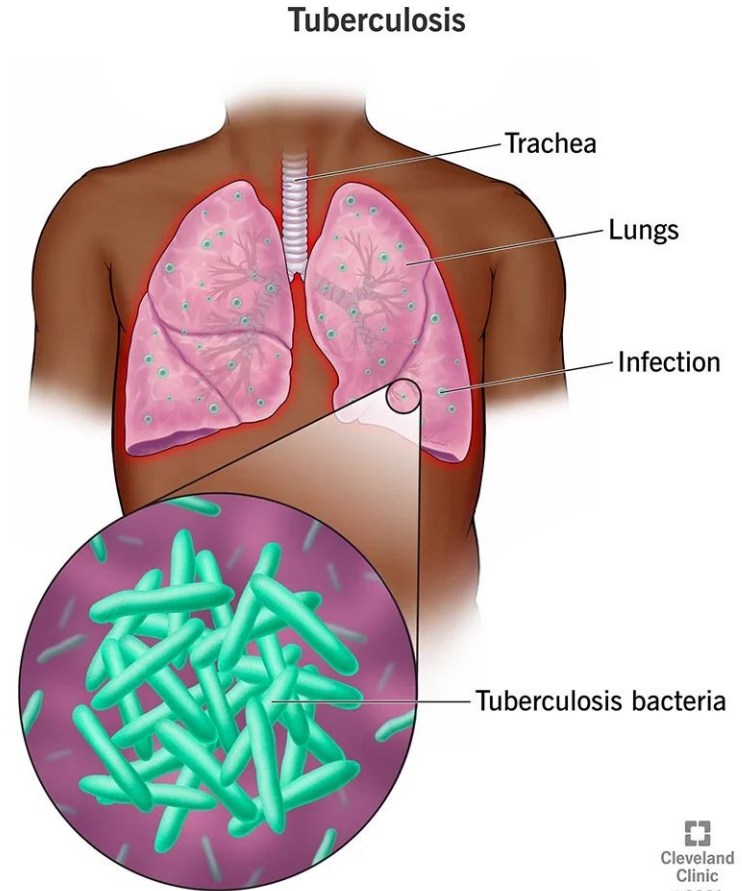
# Resistencia a antibióticos en *Mycobacterium* *tuberculosis*



Cervantes López Ximena  
Alamina Ramírez Sebastián  
Hernandez Gasca Adriana

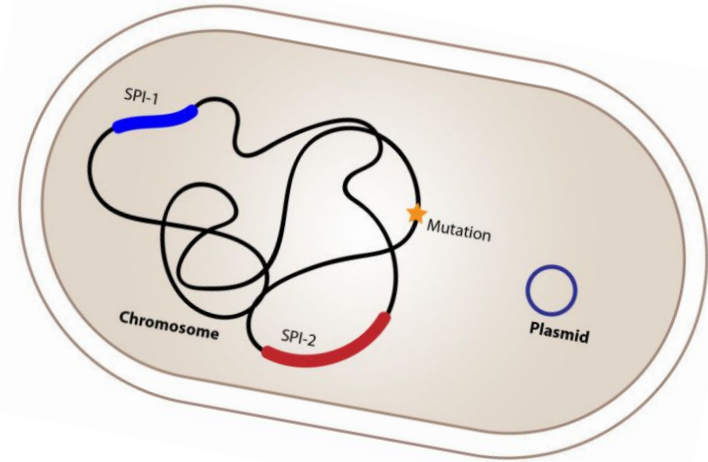
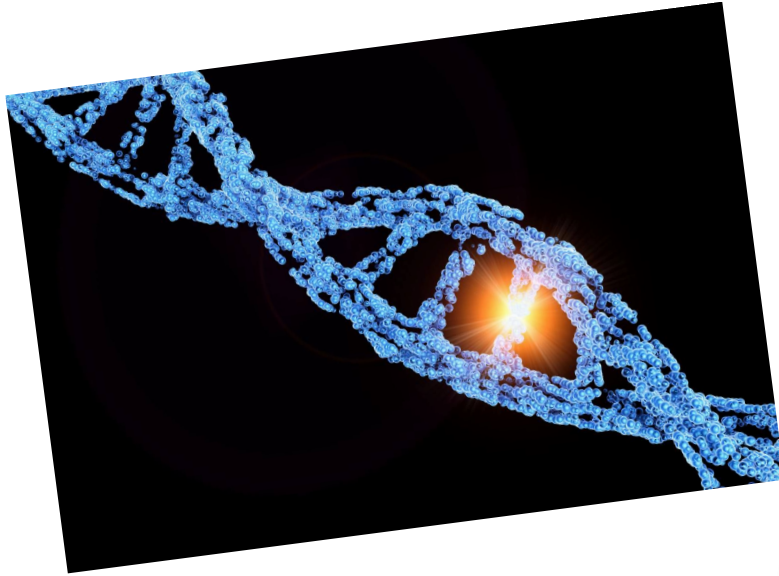
# INTRODUCCIÓN

- La tuberculosis es causada por la bacteria *Mycobacterium tuberculosis*.
- *M. tuberculosis* puede generar resistencia antibiótica.
- Se buscan herramientas de detección oportuna de esta respuesta inmune.
- La secuenciación del genoma completo (**WGS**) y la secuenciación de nueva generación (**NGS**) permiten estudiar el genoma para la identificación de variantes genéticas que le proporciona a los organismos diversas aptitudes.



# PREGUNTA DE INVESTIGACIÓN

¿Existen variantes genéticas que le permita a *M. tuberculosis* el desarrollo de resistencia a fármacos durante la primera línea de tratamiento?



# OBJETIVOS

- Identificar el cambio en los nucleótidos dentro del genoma de muestras de *M. tuberculosis* con resistencia a antibióticos.



...GTGCTGGCCCCAT...



...GTGCCGGCCCCAT...

# OBJETIVOS

- Corroborar con la literatura si la resistencia a antibióticos es debido al cambio de los nucleótidos dentro del genoma.



# METODOLOGÍA

- Búsqueda del genoma de referencia a utilizar de *M. tuberculosis* **sin** resistencia a antibióticos.

The screenshot shows the GenBank entry for the complete genome of *Mycobacterium tuberculosis* H37Rv. The header includes the NIH logo and the text "National Library of Medicine National Center for Biotechnology Information". A search bar is visible with the text "Nucleotide" and a "Search" button. The main content area displays the following information:

**Mycobacterium tuberculosis H37Rv, complete genome**  
NCBI Reference Sequence: NC\_000962.3  
[FASTA](#) [Graphics](#)

**LOCUS** NC\_000962 4411532 bp DNA linear CON 14-DEC-2017  
**DEFINITION** Mycobacterium tuberculosis H37Rv, complete genome.  
**ACCESSION** NC\_000962  
**VERSION** NC\_000962.3  
**DBLINK** BioProject: [PRJNA57777](#)  
Assembly: [GCF\\_000195955.2](#)  
**KEYWORDS** RefSeq; complete genome.  
**SOURCE** Mycobacterium tuberculosis H37Rv  
**ORGANISM** [Mycobacterium tuberculosis H37Rv](#)  
Bacteria; Actinomycetota; Actinomycetes; Mycobacteriales;  
Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis  
complex.

**REFERENCE** 1  
**AUTHORS** Lew, J.M., Kapopoulou, A., Jones, L.M. and Cole, S.T.  
**TITLE** TuberculList--10 years after  
**JOURNAL** Tuberculosis (Edinb) 91 (1), 1-7 (2011)  
**PUBMED** [20980199](#)

**REFERENCE** 2  
**AUTHORS** Camus, J.C., Pryor, M.J., Medigue, C. and Cole, S.T.  
**TITLE** Re-annotation of the genome sequence of Mycobacterium tuberculosis

On the right side, there are several tabs and links: "Change region shown", "Customize view", "Analyze this sequence" (with a sub-link "Run BLAST"), "Pick Primers", "Related information" (with sub-links: "Assembly", "BioProject", "Protein", "PubMed", "Taxonomy", "Components (Core)", "Full text in PMC", "Gene", "Genome", "Identical GenBank Sequence").

**Genoma de referencia de *Mycobacterium tuberculosis*.**

# METODOLOGÍA

- Búsqueda del genoma de referencia a utilizar de *M. tuberculosis* **con** resistencia a antibióticos.  
(SRX26381650 a SRX26381664)

The screenshot shows the SRA (Sequence Read Archive) page for SRX2638164. The header includes the NIH logo and the text "National Library of Medicine National Center for Biotechnology Information". A search bar is present with the text "SRA" and a dropdown menu showing "SRA". A "Search" button is to the right. Below the search bar, there is a "Full" dropdown and a "Send to:" dropdown. The main content area displays the title "SRX2638164: WGS of pulmonary tuberculosis isolates: Sputum" and a description: "1 ILLUMINA (Illumina HiSeq 2500) run: 1.4M spots, 285.7M bases, 107.3Mb downloads". It also shows the "Submitted by: Institute of Bioinformatics" and a "Study" description: "Whole genome sequencing of Mycobacterium tuberculosis clinical isolates from India reveals genetic heterogeneity and region specific variations that might affect drug susceptibility". There are links for "PRJNA379070", "SRP101836", "All experiments", and "All runs". A "hide Abstract" link is also present. The abstract text reads: "Genomic landscape of Mycobacterium tuberculosis pulmonary clinical isolates from India reveals high molecular heterogeneity". Below the abstract, there is a "Sample:" section with links for "SAMN06578951", "SRS2046186", "All experiments", and "All runs". The "Organism:" is listed as "Mycobacterium tuberculosis". The "Library:" section includes "Name: LIB26053", "Instrument: Illumina HiSeq 2500", "Strategy: WGS", "Source: GENOMIC", "Selection: RANDOM", and "Layout: PAIRED". The "Runs:" section shows "1 run, 1.4M spots, 285.7M bases, 107.3Mb". A table with 5 columns (Run, # of Spots, # of Bases, Size, Published) is displayed, showing one run: "SRR5341471" with 1,414,236 spots, 285.7M bases, 107.3Mb size, and published on 2019-01-21. The ID "3816236" is shown at the bottom. On the right side, there is a "Related information" section with links for "BioProject", "BioSample", and "Taxonomy". Below that is a "Recent activity" section with links for "PRJNA379070. (200)" and "SRA Links for BioProject (Select 379070) (200)". There are also links for "Mycobacterium tuberculosis H37Rv, complete genome" and "Tuberculosis - StatPearls". A "See more..." link is at the bottom right.

NIH National Library of Medicine  
National Center for Biotechnology Information

SRA SRA Search Advanced Help

Full Send to:

**SRX2638164: WGS of pulmonary tuberculosis isolates: Sputum**  
1 ILLUMINA (Illumina HiSeq 2500) run: 1.4M spots, 285.7M bases, 107.3Mb downloads

**Submitted by:** Institute of Bioinformatics

**Study:** Whole genome sequencing of Mycobacterium tuberculosis clinical isolates from India reveals genetic heterogeneity and region specific variations that might affect drug susceptibility  
[PRJNA379070](#) • [SRP101836](#) • [All experiments](#) • [All runs](#)  
[hide Abstract](#)  
Genomic landscape of Mycobacterium tuberculosis pulmonary clinical isolates from India reveals high molecular heterogeneity

**Sample:**  
[SAMN06578951](#) • [SRS2046186](#) • [All experiments](#) • [All runs](#)  
[Organism:](#) [Mycobacterium tuberculosis](#)

**Library:**  
[Name:](#) LIB26053  
[Instrument:](#) Illumina HiSeq 2500  
[Strategy:](#) WGS  
[Source:](#) GENOMIC  
[Selection:](#) RANDOM  
[Layout:](#) PAIRED

**Runs:** 1 run, 1.4M spots, 285.7M bases, [107.3Mb](#)

Run	# of Spots	# of Bases	Size	Published
<a href="#">SRR5341471</a>	1,414,236	285.7M	107.3Mb	2019-01-21

ID: 3816236

**Related information**  
[BioProject](#)  
[BioSample](#)  
[Taxonomy](#)

**Recent activity**  
[Turn Off](#) [Clear](#)  
[PRJNA379070. \(200\)](#) SRA  
[SRA Links for BioProject \(Select 379070\) \(200\)](#) SRA  
[Mycobacterium tuberculosis H37Rv, complete genome](#) Nucleotide  
[Tuberculosis - StatPearls](#)  
[See more...](#)

**Genoma con variantes genéticas de *Mycobacterium tuberculosis*.**



# METODOLOGÍA

- Búsqueda de los archivos en formato FASTA (SRA).

## **Mycobacterium tuberculosis H37Rv, complete genome**

NCBI Reference Sequence: NC\_000962.3

[GenBank](#) [Graphics](#)

```
>NC_000962.3 Mycobacterium tuberculosis H37Rv, complete genome
TTGACCGATGACCCGGTTACAGGCTTCACACAGTGTGGAACGGGTCTGCCAACTTAACGGCGACC
CTAAGGTTGACGACGGACCCAGCAGTGTATGCTAATCTCAGCGCTCCGCTGACCCCTCAGCAAAGGGCTTG
GCTCAATCTCGTCCAGCCATTGACCATCGTCGAGGGGTTTGCTCTGTTATCCGTCGCCGAGCAGCTTTGTC
CAAAACGAAATCGAGCGCCATCTGCGGGCCCCGATTACCGACGCTCTCAGCCGCGACTCGGACATCAGA
TCCAACCTCGGGTCCGCTATCGTCCGCGCGGACGACGAAAGCCGACGACACTACCGTCCGCGCTTCCGA
AAATCTCGTCTACCCATCGCCAGACACCAACCGACAACGACGAGATTGATGACAGCGCTGCGGCACGG
GGCGATAACGACGACAGTTGGCCAAGTTACTTACCAGCGCGCCGACAATACCGATTCCGCTACCGCTG
GCGTAACGAGCTTAACCGTCGCTACACCTTTGATACGTTCTGTTATCGGCGCTCCAACCGGTTCCGCGCA
CGCCGCGCTTGGCGATCGCAGAAGCACCCGCGCGCTTACAACCCCTGTTTCATCTGGGGCGAGTCC
GGTCTCGGCAAGACACCTGCTACACGCGGACGGCAACTATGCCAACGGTTGTTCCCGGGAATGCGGG
TCAAAATATGCTCCACCGAGGAATTCACCAACGACTTCATTAACTCGCTCCGCGATGACCGCAAGGTCGC
ATTCAAACGCGAGCTACCGGACGTAGACGTGCTGTTGGTCGACGACATCCAATTCAATTGAAGGCAAGAG
GGTATTCAAGAGGAGTCTTCCACACCTTCAACACCTTGACAATGCCAACAAGCAAAATCGTCATCTCAT
CTGACCGCCCAACCAAGCAGCTCGCCACCTCGAGGACCGGCTGAGAACCCTGTTGAGTGGGGGCTGAT
CACTGACGTACAACCACCGAGCTGGAGACCCGATCGCCATCTTGCGCAAGAAAGCACAGATGGAACGG
CTCGGGTCCCCGACGATGTCTCGAACTCATCGCCAGCAGTATCGAACGCAATATCCGTAACCTCGAGG
GCGCGCTGATCCGGGTACCGCGTTTCGCTCATTGAACAAAACCAATCGACAAAGCGCTGGCCGAGAT
TGTGCTTCGCGATCTGATCGCCGACGCCAACCATGCAAACTAGCGGCGGCGACGATCATGGCTGCCACC
GCCGAATACTTCGACACTACCGTCAAGAGCTTCGCGGGCCCGCAAGACCAGGACTGGCCAGTAC
GACAGATTGGGATGTACCTGTGTCGTGAGCTACCGATCTTTCTGTTGCCAAAATCGGCCAAGCGTTCCG
CCGTGATCACACAACCGTATGTACGCCCAACGCAAGATCTGTCCGAGATGGCCGAGCGCGTGAAGTC
TTTGATCAGCTCAAAGAATCACCCTCGATCCGTGAGCGCTCCAAGCGCTAGCACGGCGTGTCTTCC
GACAACGTTCTTAAAAAACTTCTCTCTCCAGGTACACCAAGTACAGAGATTGGCTGTGAGTGTGCT
GTGCAACAAACCGGCACAGACTCATACAGTCCCGCGGTTCCGTTACAACCCACGCTCATCCCCACCG
ACCCAACACACACCCACAGTCATCGCCACCGTATCCACAATCCGACCGACGTGACCTGCACCAAGA
CCAGACTGCTCCCAAACTGCACACCTCTAATACTGTTACCGAGATTCTTCGTGTTGTTCTTGGAAA
GACAGCGCTGGGGATCGTTGCTGGATACCAACCCGATAAATGGCTCGTCGCGTGGTTCAGAGGTCAAT
GATGAATTTCAAGTTGACGTGAGAAGCTTACGGTTGTTGTTGACTGCTGTTGCGGCGCTGTTGGCGG
GTCACGCGTCATGGGATTCGTGTTGGCAGTCCACGCTAGCGGGGCGTAGCCACGGGATCGAACTC
ATCGTGAGGTGAAAGGCGCAATGGACGCGGCTACGACAAGAGTTGGCTCACCGACTTGACGTTTCGTT
TGCTACGAGAGTCTTTTCGCGATGCGGTGTCGTGGGTGGCTAAAAATCTGCAGCGCAGGCGCGGTTGCC
GGTGCTCTCCGGCGTGTGTTGACCGGCTCGGACAACGGTCTGACGATTTCGGGATTCGACTACGAGGTT
TCCGCCGAGGCCAGGTTGGCGCTGAAATGTTTCTCCTGGAAGCGTTTATGTTTTCGGCCGATTGTTGT
```

ence: 4.16MB (COMPLETE)



- Carga de los datos, tanto del genoma de referencia como de las muestras a comparar.

**Tools**

**Upload Data**

- Get Data
- Send Data
- Collection Operations
- GENERAL TEXT TOOLS**
- Text Manipulation
- Filter and Sort
- Join, Subtract and Group
- Datamash
- GENOMIC FILE MANIPULATION**
- FASTA/FASTQ
- FASTQ Quality Control
- SAM/BAM
- BED
- VCF/BCF
- Nanopore
- Convert Formats
- Lift-Over
- COMMON GENOMICS TOOLS**
- Interactive tools

! This dataset is large and only the first megabyte is shown below.  
[Show all](#) | [Save](#)

```
@1/1
ATAGGCGTTGGTGTATTGTTAAATACATCAACATTGGCCTAAGTGATTAAATTTTTTCCTATGTCGCACACC#
+
AAFF<FKKKFKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKFKKKKKFKKKKK7<
@2/1
GTGGCCGGTCAGGTACCAGAGCATCGGGCGTCCACCTGACATTCTCTGA CTGGTCACATCGGTGACAGTAGCC
+
<A, FFKFKKKKKKKKKFKKKKKKKFKKKKKKKFKAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKFKKKKF7FA>
@3/1
ACATCGAATCGATGACGTGCCNCNTNAGACTTGACCTGAGGCTGAGCACGATGGTCTCGCCGACATGCGGTC1
+
AAAFFKKKFKFKKKFKKKKKKKKK#K##K#KKKKKKKKKKKKAFKKKKKKKKKKFKKKKKKFA<FK( FKKKKFF
@4/1
TGCGCAGCGTCCGGAAACCATGNGNNNAGNAAC TGACACCCAGGTGACCCAGTCCGCCGATGCCAAGACGC
+
AAAFFKKKKKKKKFKKKKKKKKK#K###K#KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK
@5/1
GTGCCCTCGGCCGCCGATGTGNCNNNANGTGCTGTCATCGCCCAGCCTCACCCAACATCGTGGCGGCCAAA1
+
AA, AFKKFFFFKKFFFFKKKKKKK#K###KK#<FFKKKKFF7<, , AF(<KKAKAKKK<FKK, F, FF<, <7F
@6/1
GCGGCGAGCAGGCCCGGTGTTNCNNNNGNCGGCCGATGAGTTCAGAGTGCGCGCCACCGGATGCCACCC
+
AAFFFFFFFKKKFKFFFFFKK#K###K#KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK
@7/1
GGCCGAGCTCGACACCGGTGNCNNNNGNCGCCGATACCACTCACAACCGCATGTCAGCATCACCC#
+
AAAFFK, FKK, AKKKFKFKKKK#K###K#AFA<AFF, FKKAKKA, , AFF7<AFF<F, AF7AFKKKF7<, FF
@8/1
GGTGATGTTCCCGCGGGCTTCGNNNNNNGAGCGGATGAGCGGTGACTGAGCACGCCGACAGAGCACCA
+
AAFFAFAKKKKKFF<FFFKKKK#F####F#KKKKFKKKFKFFKKFKFK7FA, FKKFKKAKKK7FKK<<F<, <F
@9/1
CGACAGCCACAGCTCGGAGATGACCCGGCGGATGTTGTA CTGGTCTCGGTGCGGTCACCGTCTGCTGGCGCC
+
AAAFFKKFKFKF<FKFK, <FFKKKKKKAKAKAF7FFKKKKKKFKFKFFKFAFFKKFKKFF7AA(AF(K7(,
@10/1
```

**History**

**Projecto final**

27.2 GB    506    532    3

Name	Size	Views	Downloads	Actions
71 : SRX2638157:reverse				[eye] [pencil] [trash]
70 : SRX2638157:forward				[eye] [pencil] [trash]
69 : SRX2638160:reverse				[eye] [pencil] [trash]
68 : SRX2638160:forward				[eye] [pencil] [trash]
65 : SRX2638150:reverse				[eye] [pencil] [trash]
64 : SRX2638150:forward				[eye] [pencil] [trash]
61 : SRX2638159:reverse				[eye] [pencil] [trash]
60 : SRX2638159:forward				[eye] [pencil] [trash]
57 : SRX2638161:reverse				[eye] [pencil] [trash]
56 : SRX2638161:forward				[eye] [pencil] [trash]
51 : SRX2638162:reverse				[eye] [pencil] [trash]
50 : SRX2638162:forward				[eye] [pencil] [trash]
47 : SRX2638163:reverse				[eye] [pencil] [trash]
46 : SRX2638163:forward				[eye] [pencil] [trash]

# METODOLOGÍA

- Uso de la herramienta *FastQC* para el análisis de calidad de las muestras.

**Galaxy** Workflow Visualize Shared Data Help User Using 10%

Tools search tools Upload Data

Get Data  
Send Data  
Collection Operations  
**GENERAL TEXT TOOLS**  
Text Manipulation  
Filter and Sort  
Join, Subtract and Group  
Datamash  
**GENOMIC FILE MANIPULATION**  
FASTA/FASTQ  
FASTQ Quality Control  
SAM/BAM  
BED  
VCF/BCF  
Nanopore  
Convert Formats  
Lift-Over  
**COMMON GENOMICS TOOLS**  
Interactive tools  
Create a Galaxy Instance

**FastQC Report** Wed 31 May 2023 SRX2638154\_forward.gz

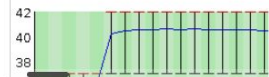
**Summary**

- ✓ Basic Statistics
- ✓ Per base sequence quality
- ✓ Per sequence quality scores
- ✗ Per base sequence content
- ✗ Per sequence GC content
- ✓ Per base N content
- ✓ Sequence Length Distribution
- ✓ Sequence Duplication Levels
- ✗ Overrepresented sequences
- ✓ Adapter Content

**Basic Statistics**

Measure	Value
Filename	SRX2638154_forward.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	4400016
Sequences flagged as poor quality	0
Sequence length	75
%GC	63

**Per base sequence quality**



Produced by **FastQC** (version 0.11.9)

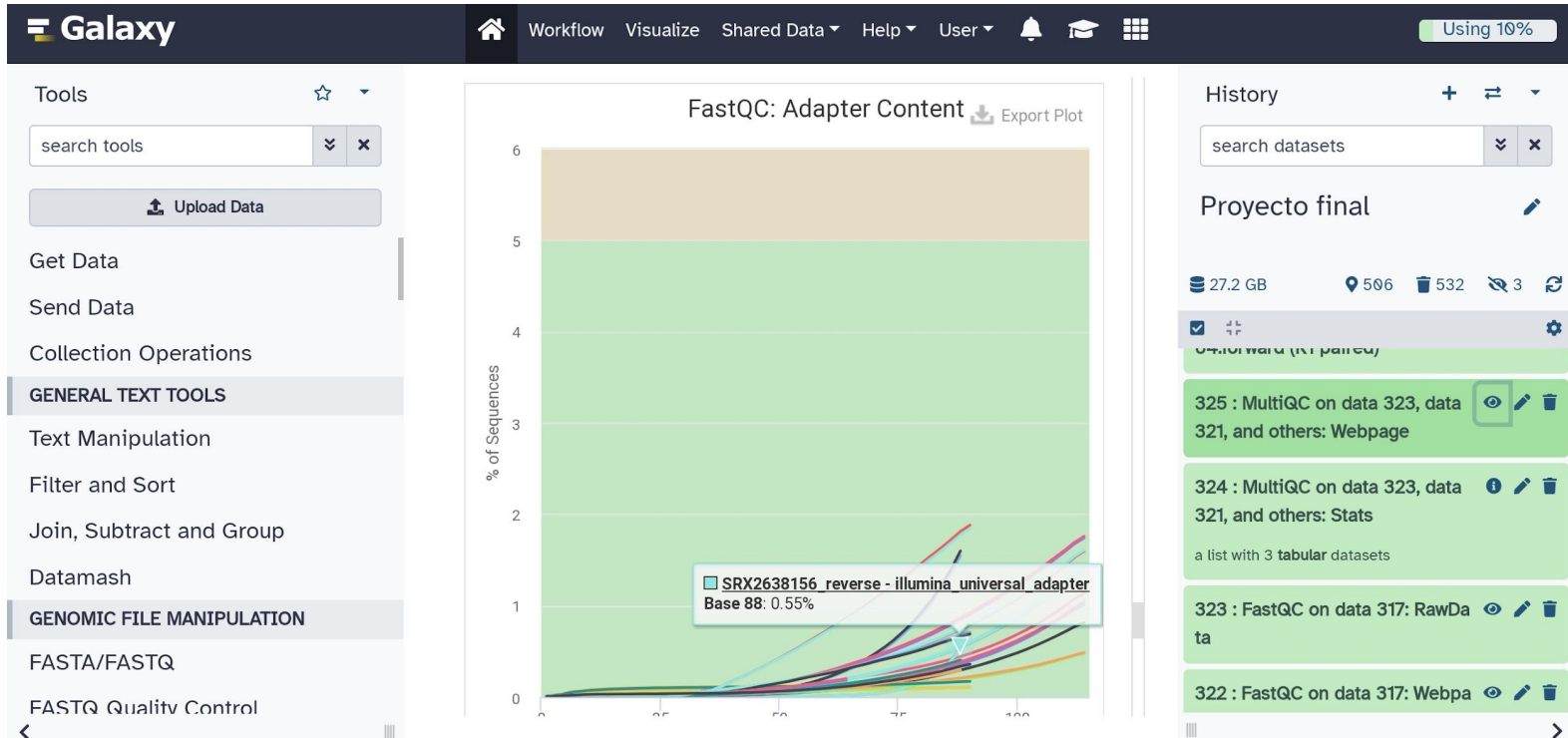
History search datasets

Proyecto final

- 27.2 GB 506 532 3
- 237 : FastQC on data 106: RawData
- 236 : FastQC on data 106: Webpage
- 235 : FastQC on data 103: RawData
- 234 : FastQC on data 103: Webpage
- 233 : FastQC on data 102: RawData
- 232 : FastQC on data 102: Webpage
- 231 : FastQC on data 97: RawData
- 230 : FastQC on data 97: Webpage
- 229 : FastQC on data 96: RawData

# METODOLOGÍA

- Uso de la herramienta *MultiQC* para observar la presencia de adaptadores y el análisis de calidad en conjunto.



# METODOLOGÍA

- Uso de la herramienta *Trimmomatic* para el corte de adaptadores en las muestras secuenciadas.

The screenshot displays the Galaxy web interface. The top navigation bar includes the Galaxy logo, a home icon, and links for Workflow, Visualize, Shared Data, Help, User, a notification bell, a graduation cap icon, and a grid icon. A status bar on the right indicates 'Using 10%'. The left sidebar contains a 'Tools' section with a search bar and an 'Upload Data' button. Below this are categories: 'Get Data', 'Send Data', 'Collection Operations', 'GENERAL TEXT TOOLS' (highlighted), 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Datemash', and 'GENOMIC FILE MANIPULATION'. Under 'GENOMIC FILE MANIPULATION', 'FASTA/FASTQ' and 'FASTQ Quality Control' are listed. The main workspace shows a dataset view with a warning: 'This dataset is large and only the first megabyte is shown below.' Below the warning, the dataset content is displayed in a text editor, showing FASTQ format data with headers like '@1/2', '@3/2', '@4/2', '@6/2', '@7/2', and '@8/2'. The right sidebar features a 'History' section with a search bar and a list of datasets. The datasets listed are: '160:reverse (R2 unpaired)', '499 : Trimmomatic on SRX2638 160:forward (R1 unpaired)', '498 : Trimmomatic on SRX2638 160:reverse (R2 paired)', '497 : Trimmomatic on SRX26381 60:forward (R1 paired)', and '492 : Trimmomatic on SRX26381 50:reverse (R2 unpaired)'. Each dataset entry includes icons for viewing, editing, and deleting.

# METODOLOGÍA

- Uso de la herramienta *Kraken2* para el análisis de contaminación de las muestras.

The screenshot displays the Galaxy web interface. The top navigation bar includes the Galaxy logo, a home icon, and links for Workflow, Visualize, Shared Data, Help, User, and a notification bell. A status bar on the right indicates 'Using 10%'. The left sidebar contains a 'Tools' section with a search bar and an 'Upload Data' button. Below this are categories: 'Get Data', 'Send Data', 'Collection Operations', 'GENERAL TEXT TOOLS' (highlighted), 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Datemash', 'GENOMIC FILE MANIPULATION' (highlighted), 'FASTA/FASTQ', and 'FASTQ Quality Control'.

The main workspace shows a table of results from a Kraken2 workflow. The table has six columns: Column 1, Column 2, Column 3, Column 4, Column 5, and Column 6. The data rows are as follows:

Column 1	Column 2	Column 3	Column 4	Column 5	Column 6
8.85	267325	267325	U	0	unclassified
91.15	2754702	320487	R	1	root
80.20	2423686	1326	R1	131567	cellular organisms
79.60	2405623	30487	D	2	Bacteria
78.38	2368542	1764	D1	1783272	Terrabacteria group
75.55	2283205	1951	P	201174	Actinobacteria
75.48	2281075	74135	C	1760	Actinomycetia
72.73	2197984	33961	O	85007	Corynebacteriales
71.45	2159341	210668	F	1762	Mycobacteriaceae
59.49	1797754	196415	G	1866885	Mycolicibacterium
51.30	1550208	1550208	S	1766	Mycolicibacterium fortuitum
0.69	20936	20936	S	146017	Mycolicibacterium boenickei

The right sidebar shows the 'History' section with a search bar and a list of datasets. The top entry is 'Proyecto final'. Below it, a summary shows '27.2 GB', '506' locations, '532' files, and '3' links. A list of datasets follows, including 'ata 417: Classification', '644 : Report: Kraken2 on data 418 and data 417', '643 : Kraken2 on data 450 and data 449: Classification', '642 : Report: Kraken2 on data 450 and data 449', and '641 : Kraken2 on data 426 and data 425: Classification'.

# METODOLOGÍA

- Uso de la herramienta *Snippy* para el alineamiento y el llamado de variantes en las muestras.

The screenshot displays the Galaxy web interface. The top navigation bar includes the Galaxy logo, a home icon, and links for Workflow, Visualize, Shared Data, Help, User, and a notification bell. A 'Using 10%' indicator is visible on the right.

On the left, the 'Tools' sidebar is open, showing a search bar and a list of tool categories: Get Data, Send Data, Collection Operations, GENERAL TEXT TOOLS (highlighted), Text Manipulation, Filter and Sort, Join, Subtract and Group, Datamash, GENOMIC FILE MANIPULATION, FASTA/FASTQ, and FASTQ Quality Control.

The central workspace contains a table with 8 columns labeled Column 1 through Column 8. The table lists genomic data for various chromosomes and positions, including variant types (snps, ins) and reference/alternative alleles.

On the right, the 'History' panel is visible, showing a search bar and a list of datasets. The datasets are titled 'Proyecto final' and include details such as size (27.2 GB), location (506), and a list of operations performed, including 'snippy on data 386, data 385, and data 318 mapped reads (bam)'.

Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8
CHROM	POS	TYPE	REF	ALT	EVIDENCE	FTYPE	STRAND
Chromosome	157333	snp	C	T	T:14 C:0	CDS	-
Chromosome	218599	snp	T	C	C:11 T:0		
Chromosome	309765	snp	T	C	C:11 T:0	CDS	+
Chromosome	336005	snp	A	G	G:13 A:1	CDS	-
Chromosome	339230	snp	G	C	C:15 G:0		
Chromosome	428921	snp	G	A	A:13 G:0	CDS	-
Chromosome	830868	ins	G	GGC	GGC:12 G:0	CDS	+



# METODOLOGÍA

- Uso de la herramienta *TB-Variant Filter* para la perfilación de las diferentes variantes que se encontraron en las muestras.

The screenshot displays the Galaxy web interface. The top navigation bar includes the Galaxy logo, a home icon, and links for Workflow, Visualize, Shared Data, Help, User, and a notification bell. A status bar on the right indicates 'Using 10%'. The left sidebar contains a 'Tools' section with a search bar and an 'Upload Data' button. Below this are categories: 'Get Data', 'Send Data', 'Collection Operations', 'GENERAL TEXT TOOLS' (highlighted), 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Datemash', 'GENOMIC FILE MANIPULATION' (highlighted), 'FASTA/FASTQ', and 'FASTQ Quality Control'.

The main workspace shows a workflow with two tools. The first tool is '##INFO=<ID=LOF,Number=.,Type=String,Description="Predicted loss of function"'. The second tool is '##INFO=<ID=NMD,Number=.,Type=String,Description="Predicted nonsense medi'. Below these tools is a table with columns: #CHROM, POS, ID, REF, ALT, QUAL, FILTER, and INFO. The table contains 12 rows of genomic data.

The right sidebar shows the 'History' section with a search bar and a list of datasets. The first dataset is 'Proyecto final' with a size of 27.2 GB, 506 locations, 532 files, and 3 views. Below this are several workflow steps, each represented by a green box with a tool icon, a name, and a number. The steps are: '720 : TB Variant Filter on data 6 62', '719 : TB Variant Filter on data 65 9', '718 : TB Variant Filter on data 65 6', '717 : TB Variant Filter on data 65 3', and '716 : snippy on data 422, data 4'.

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO
Chromosome	1849	.	C	A	2577.47	.	AB=0.0;AO=
Chromosome	2532	.	C	T	3604.01	.	AB=0.0;AO=
Chromosome	3446	.	C	T	4007.0	.	AB=0.0;AO=
Chromosome	9143	.	C	T	3223.11	.	AB=0.0;AO=
Chromosome	11976	.	G	C	3172.34	.	AB=0.0;AO=
Chromosome	12163	.	G	A	2718.59	.	AB=0.0;AO=
Chromosome	12204	.	G	A	2625.39	.	AB=0.0;AO=
Chromosome	13460	.	G	A	2860.88	.	AB=0.0;AO=
Chromosome	13926	.	C	T	2229.84	.	AB=0.0;AO=
Chromosome	14401	.	G	A	2795.14	.	AB=0.0;AO=
Chromosome	21819	.	C	A	2991.41	.	AB=0.0;AO=
Chromosome	23174	.	A	C	2667.61	.	AB=0.0;AO=

# METODOLOGÍA

- Uso de la herramienta *TB-Profiler-Profile* para la identificación de las variantes relacionadas a la resistencia en fármacos.

The screenshot displays the Galaxy web interface. The top navigation bar includes the Galaxy logo, a home icon, and links for Workflow, Visualize, Shared Data, Help, User, and a notification bell. A 'Using 10%' status indicator is on the right. The left sidebar contains a 'Tools' section with a search bar and an 'Upload Data' button. Below this are categories: 'Get Data', 'Send Data', 'Collection Operations', 'GENERAL TEXT TOOLS' (highlighted), 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Datamash', and 'GENOMIC FILE MANIPULATION' (highlighted), which includes 'FASTA/FASTQ' and 'FASTQ Quality Control'. The main workspace shows a workflow titled 'Resistance variants report'. It lists input datasets: Clofazimine, Para-aminosalicylic\_acid, Delamanid, Bedaquiline, and Linezolid. The workflow steps are shown in a table with columns: Genome, Position, Locus, Tag, Gene, Variant, Type, and Change. The first section, 'Resistance variants report', shows three steps for rifampicin, isoniazid, and ethambutol. The second section, 'Other variants report', shows two steps for moxifloxacin and levofloxacin. The right sidebar shows a 'History' section with a search bar and a list of previous jobs, including '1019 : TB-Profiler Profile VCF on data 700', '1018 : TB-Profiler Profile on data 700: Results.json', '1017 : TB-Profiler Profile report on data 697', '1016 : TB-Profiler Profile VCF on data 697', and '1015 : TB-Profiler Profile on data'.

**Galaxy**

Workflow Visualize Shared Data Help User Using 10%

**Tools**

search tools

Upload Data

Get Data

Send Data

Collection Operations

**GENERAL TEXT TOOLS**

Text Manipulation

Filter and Sort

Join, Subtract and Group

Datamash

**GENOMIC FILE MANIPULATION**

FASTA/FASTQ

FASTQ Quality Control

Clofazimine  
Para-aminosalicylic\_acid  
Delamanid  
Bedaquiline  
Linezolid

Resistance variants report

Genome	Position	Locus	Tag	Gene	Variant	Type	Change
Estimated	Fraction			Drug	Annotation		
761161	Rv0667	rpoB		missense_variant			p.Leu452Pro
1.000		rifampicin					
type=who_confidence drug=rifampicin who_confidence=Assoc w R							
2155168	Rv1908c	katG		missense_variant			p.Ser315Thr
1.000		isoniazid					
type=who_confidence drug=isoniazid who_confidence=Assoc w R							
4247574	Rv3795	embB		missense_variant			p.Asp354Ala
1.000		ethambutol					
type=who_confidence drug=ethambutol who_confidence=Assoc w R							

Other variants report

Genome	Position	Locus	Tag	Gene	Variant	Type	Change
Estimated	Fraction			Annotation			
6112	Rv0005	gyrB		missense_variant			p.Met291Ile
1.000		type=who_confidence drug=moxifloxacin who_confidence=Not					
assoc w R							
R; type=who_confidence drug=levofloxacin who_confidence=Not assoc w R							
7362	Rv0006	gyrA		missense_variant			p.Glu216Gln
1.000		type=who_confidence drug=moxifloxacin who_confidence=Not					

**History**

search datasets

Projecto final

27.2 GB 506 532 3

1019 : TB-Profiler Profile VCF on data 700

1018 : TB-Profiler Profile on data 700: Results.json

1017 : TB-Profiler Profile report on data 697

1016 : TB-Profiler Profile VCF on data 697

1015 : TB-Profiler Profile on data

# RESULTADOS

- Fármacos a los que se encuentran asociadas nuestras 12 muestras de las 14 ya que 2 no mostraron variantes de resistencia.

- Estreptomicina (Str)
- Rifampicina (Rif)
- Ofloxacina (Ofi)
- Kanamicina (Kan)
- Moxifloxacino (Mox)
- Isoniazida (Iso)
- Levofloxacino (Lev)
- Bedaquilina (Bed)
- Fluoroquinolonas (Flu)
- Clofazimina (Clo)
- Ciprofloxacina (Cip)
- Etionamida (Eti)
- Etambutol (Etam)
- Capreomicina (Cap)
- Amikacina (Ami)
- Pirazinamida (Pir)
- Capreomicina (Cap)
- Aminoglucósidos (Ami)

Muestra	Tipo de variante	Fármacos a los que es resistente.
SRX26381650	Sensitive	Ninguno
SRX26381651	Missense	Rif, Iso, Etam
SRX26381652	Missense	Mox, Lev, Rif, Etam, Bed, Clo, Str, Eti, Kan
SRX26381653	Frameshift	Str
SRX26381655	Missense	Str, Iso
SRX26381656	Sensitive	Ninguno
SRX26381657	Non coding transcript exon	Str, Kan, Cap, Ami
SRX26381658	Missense	Ofi, Mox, Lev, Flu, Cip, Str, Rif, Iso, Eta, Pir
SRX26381659	Missense	Rif, Iso
SRX26381660	Non coding transcript exon	Str
SRX26381661	Missense	Ofi, Mox, Lev, Flu, Cip
SRX26381662	Non coding transcript exon	Str.
SRX26381663	Missense	Etam
SRX26381664	Missense y Frameshift	Oflo, Mox, Levo, Flu, Cip, Ami, Cap, Ami

# CONCLUSIONES

- 12/14 muestras presentan mutaciones significativas.
- Estas últimas relacionadas con respuestas (principalmente respecto a la resistencia) ante fármacos.
- Resistencia a estreptomicina.
- Con el paso de los años las bacterias han obtenido cada vez más una resistencia a fármacos que hoy en día puede ser muy preocupante.

# FUENTES DE INFORMACIÓN

- [Arrigoni et al., 2022] Arrigoni, R., Ballini, A., Topi, S., Bottalico, L., Jirillo, E., and Santacroce, L. (2022). Antibiotic resistance to mycobacterium tuberculosis and potential use of natural and biological products as alternative anti-mycobacterial agents. *Antibiotics (Basel, Switzerland)*, 11(10), 1431. <https://doi.org/10.3390/antibiotics11101431>.
- [Chai et al., 2018] Chai, Q., Zhang, Y., and Liu, C. H. (2018). Mycobacterium tuberculosis: An adaptable pathogen associated with multiple human diseases. *Frontiers in cellular and infection microbiology*, 8, 158. <https://doi.org/10.3389/fcimb.2018.00158>.
- [Dees et al., 2017] Dees, J., Momsen, J. L., Niemi, J., and Montplaisir, L. (2017). Student interpretations of phylogenetic trees in an introductory biology course. *CBE—Life Sciences Education* Vol. 13, No. 4. <https://doi.org/10.1187/cbe.14-01-0003>.
- [Ghosh et al., 2020] Ghosh, A., N., S., and Saha, S. (2020). Survey of drug resistance associated gene mutations in mycobacterium tuberculosis, escape and other bacterial species. *Scientific reports*, 10(1), 8957. <https://doi.org/10.1038/s41598-020-65766-8>.
- [Hiltemann et al., 2023] Hiltemann, S., Rasche, H., and Soranzo, N. (2023). A short introduction to galaxy. <https://training.galaxyproject.org/training-material/topics/introduction/tutorials/galaxy-intro-short/slides-plain.html>. Galaxy Training!



# FUENTES DE INFORMACIÓN

- [Martinez and Baquero, 2000] Martinez, J. L. and Baquero, F. (2000). Mutation frequencies and antibiotic resistance. *Antimicrobial Agents and Chemotherapy*, 44(7):1771–1777. DOI: 10.1128/aac.44.7.1771-1777.2000.
- [NCBI, 2017] NCBI (2017). Mycobacterium tuberculosis h37rv, complete genome. [http://www.ncbi.nlm.nih.gov/nuccore/NC\\_000962.3](http://www.ncbi.nlm.nih.gov/nuccore/NC_000962.3). (accessed 2023-06-01).
- [NCBI, 2019] NCBI (2019). WGS of pulmonary tuberculosis isolates: Sputum - sra. [https://www.ncbi.nlm.nih.gov/sra/SRX2638164\[accn\]](https://www.ncbi.nlm.nih.gov/sra/SRX2638164[accn]). (accessed 2023-06-01).
- [Qin, 2019] Qin, D. (2019). Next-generation sequencing and its clinical application. *Cancer biology & medicine*, 16(1), 4–10. <https://doi.org/10.20892/j.issn.2095-3941.2018.0055>.
- [van Heusden et al., 2023] van Heusden, P., Gladman, S., and Lose, T. (2023). M. tuberculosis variant analysis. <https://training.galaxyproject.org/training-material/topics/variant-analysis/tutorials/tb-variant-analysis/tutorial.html>. Galaxy Training!