


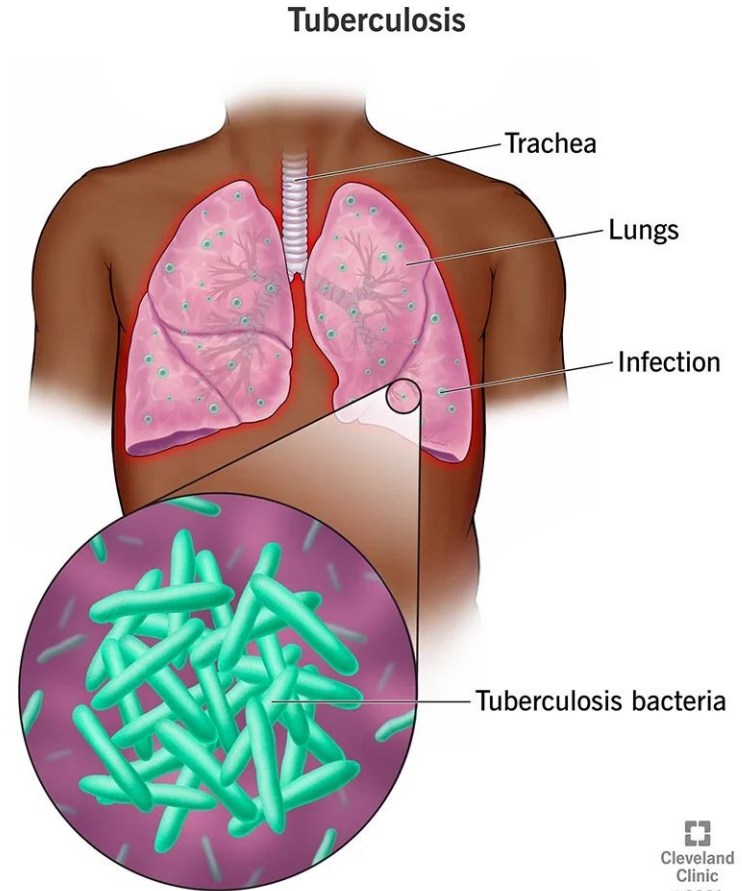
Resistencia a antibióticos en *Mycobacterium tuberculosis*



Cervantes López Ximena
Alamina Ramírez Sebastián
Hernandez Gasca Adriana

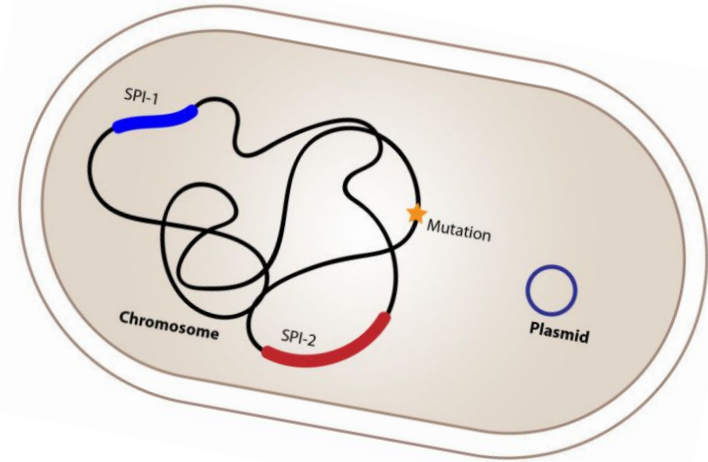
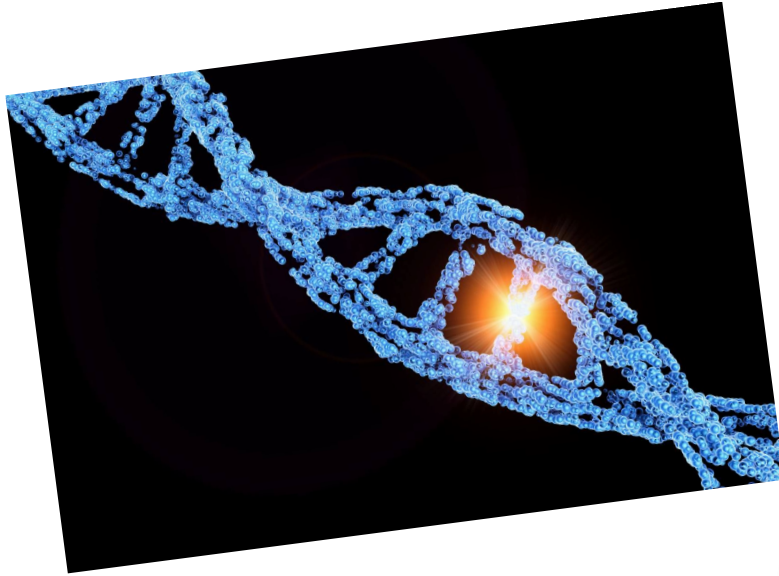
INTRODUCCIÓN

- 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.



...GTGCTGGCCCAT...



...GTGCCGGCCCAT...

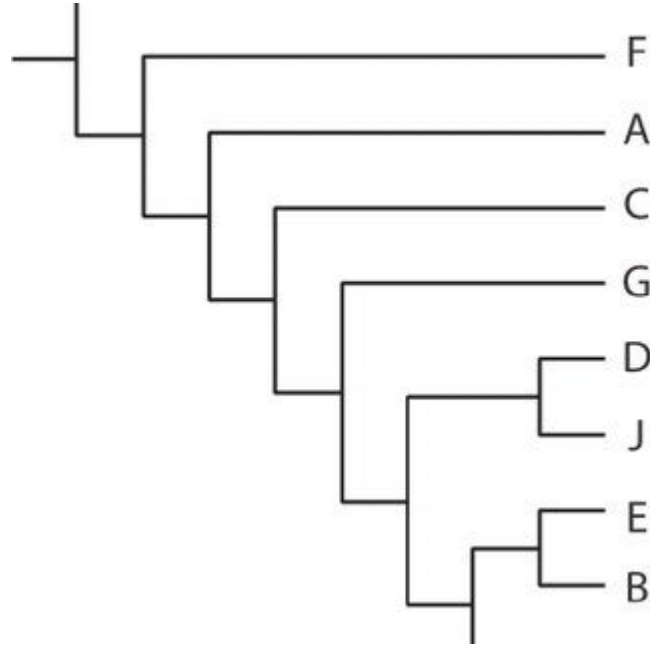
OBJETIVOS

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



OBJETIVOS

- Identificar si las muestras analizadas tienen una relación filogenética alejada entre ellas.



METODOLOGÍA

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

The screenshot shows the GenBank entry for *Mycobacterium tuberculosis* H37Rv, complete genome. The page is from the National Library of Medicine (NIH) National Center for Biotechnology Information. The top navigation bar includes the NIH logo, the name of the center, and a 'Log in' button. Below the navigation bar is a search bar with a 'Nucleotide' dropdown menu and a 'Search' button. The main content area displays the entry details for 'Mycobacterium tuberculosis H37Rv, complete genome'. It includes the NCBI Reference Sequence (NC_000962.3), links to 'FASTA' and 'Graphics', and a 'Go to:' dropdown menu. The entry details are organized into sections: 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), REFERENCE (1, 2), and AUTHORS (Lew, J.M., Kapopoulou, A., Jones, L.M. and Cole, S.T.; Camus, J.C., Pryor, M.J., Medigue, C. and Cole, S.T.). The right sidebar contains links to 'Change region shown', 'Customize view', 'Analyze this sequence' (Run BLAST, Pick Primers), and 'Related information' (Assembly, BioProject, Protein, PubMed, Taxonomy, Components (Core), Full text in PMC, Gene, Genome, Identical GenBank Sequence).

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

Nucleotide Nucleotide Search

Advanced Help

GenBank Send to:

Due to the large size of this record, sequence and annotated features are not shown. Use the "Customize view" panel to change the display.

Mycobacterium tuberculosis H37Rv, complete genome

NCBI Reference Sequence: NC_000962.3

[FASTA](#) [Graphics](#)

Go to: ▾

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

Change region shown

Customize view

Analyze this sequence

Run BLAST

Pick Primers

Related information

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 menu and a "Send to:" dropdown menu. 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 lists the submitter as "Institute of Bioinformatics" and provides 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". Links for "PRJNA379070", "SRP101836", "All experiments", and "All runs" are provided. 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". The sample information includes "Sample: SAMN06578951 • SRS2046186 • All experiments • All runs" and "Organism: Mycobacterium tuberculosis". The library information includes "Library: Name: LIB26053", "Instrument: Illumina HiSeq 2500", "Strategy: WGS", "Source: GENOMIC", "Selection: RANDOM", and "Layout: PAIRED". The runs information states: "Runs: 1 run, 1.4M spots, 285.7M bases, 107.3Mb". A table with 5 columns (Run, # of Spots, # of Bases, Size, Published) shows 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 to "BioProject", "BioSample", and "Taxonomy". Below that is a "Recent activity" section with links to "PRJNA379070. (200)" and "SRA Links for BioProject (Select 379070) (200)".

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
SRR5341471	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
TTGACCGATGACCCGGTTTCAGGCTTCACACAGTGTGGAACGGGTCTCCGAACTTAACGGCGACC
CTAAGGTTGACGACGGACCCAGCAGTGTATGCTAATCTCAGCGCTCCGCTGACCCCTCAGCAAAAGGGCTTG
GCTCAATCTCGTCCAGCCATTGACCATCGTCGAGGGGTTTGCTCTGTTATCCGTCGCCGAGCAGCTTTGTCT
CAAAACGAAATCGAGCGCCATCTGCGGGCCCCGATTACCGACGCTCTCAGCCGCGACTCGGACATCAGA
TCCAACCTCGGGTCCGCTATCGTCCGCGCGGACGACGAAAGCCGACGACACTACCGTCCGCCCTTCCGA
AAATCTCGTCTACCCATCGCCAGACACCAACCGACAACGACGAGATTGATGACAGCGCTGCGGCACGG
GGCGATAACGACGACAGTTGGCCAAGTTACTTACCAGCGCGCCGACAATACCGATTCCGCTACCGCTG
GCGTAACGAGCTTAACCGTCGCTACACCTTTGATACGTTCTGTTATCGGCGCTCCAACCGGTTCCGCGCA
CGCCGCGCTTGGCGATCGCAGAAGCACCCGCGCGCTTACAACCCCTGTTTCATCTGGGGCGAGTCC
GGTCTCGGCAAGACACCTGCTACACGCGGACGGCAACTATGCCAACGGTTGTTCCCGGGAATGCGGG
TCAAAATATGTCTCACCAGGAATTCACCAACGACTTCATTAACTCGCTCCGCGATGACCGCAAGGTCCG
ATTCAAACGCGAGCTACCGGACGTAGACGTGCTGTTGGTCGACGACATCCAATTCAATTGAAGGCAAGAG
GGTATTCAAGAGGAGTCTTCCACACCTTCAACACCTTGACAATGCCAACAAGCAAAATCGTCATCTCAT
CTGACCGCCCAACCAAGCAGCTCGCCACCTCGAGGACCGGCTGAGAACCCTGTTGAGTGGGGGCTGAT
CACTGACGTACAACCACCGAGCTGGAGACCCGATCGCCATCTTGCGCAAGAAAGCACAGATGGAACGG
CTCGGGTCCCCGACGATGCTCTGAACTCATCGCCAGCAGTATCGAACGCAATATCCGTAACCTCGAGG
GCGCGCTGATCCGGGTACCGCGTTTCGCTCATTGAACAAAACCAATCGACAAAGCGCTGGCCGAGAT
TGTGCTTCGCGATCTGATCGCCGACGCCAACCATGCAAACTAGCGGCGGCGACGATCATGGCTGCCACC
GCCGAATACTTCGACACTACCGTCAAGAGCTTCGCGGGCCCGCAAGACCAGGACTGGCCAGTAC
GACAGATTGGGATGTACCTGTGCTGAGCTACCGATCTTTCTGTTGCCAAAATCGGCCAAGCGTTCCG
CCGTGATCACACAACCGTATGTACGCCCAACGCAAGATCTGTCCGAGATGGCCGAGCGCGTGAAGTC
TTTGATCAGCTCAAAGAATCACCCTCGATCCGTGAGCGCTCCAAGCGCTAGCACGGCGTGTCTTCC
GACAACGTTCTTAAAAAACTTCTCTCTCCAGGTACACCAAGTACAGAGATTGGCTGTGAGTGTGCT
GTGCAACAAACCGGCACAGACTCATACAGTCCCGCGGTTCCGTTTACAACCCACGCTCATCCCCACCG
ACCCAACACACACCCACAGTCATCGCCACCGTATCCACAATCCGACCGACGTGACCTGCACCAAGA
CCAGACTGCTCCCAAACTGCACACCTCTAATACTGTTACCGAGATTCTTCGTGTTGTTCTTGGAAA
GACAGCGCTGGGGATCGTTGCTGGATACCAACCCGATAAATGGCTCGTCGCGTGGTTCAGAGGTCAAT
GATGAATTTCAAGTTGACGTGAGAAGCTTACGGTTGTTGTTGACTGCTGTTGCGGCGCTGTTGCGG
GTCACGCGTCATGGGATTCGTGTTGGCAGTCCACGCTAGCGGGGCGTAGCCACGGGATCGAACTC
ATCGTGAGGTGAAAGGCGCAATGGACGCGGCTACGACAAGAGTTGGCTCACCGACTTGACGTTTCGTT
TGCTACGAGAGTCTTTTCGCGATGCGGTGCTGTTGGTGGCTAAAAATCTGCAGCGCAGGCGCGGTTGCC
GGTGCTCTCCGGCGTGTGTTGACCGGCTCGGACAACGGTCTGACGATTTCCGGATTGACTACGAGGTT
TCCCGCGAGGCCAGGTTGGCGCTGAAATGTTTCTCCTGGAAGCGTTTATGTTTTCGGCCGATTGTTGT
```

ence: 4.16MB (COMPLETE)

METODOLOGÍA

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

The screenshot shows the Galaxy web interface's 'Tools' sidebar. At the top, there is a search bar and an 'Upload Data' button. Below these are categories: 'Get Data', 'Send Data', 'Collection Operations', 'GENERAL TEXT TOOLS' (highlighted), 'Text Manipulation', 'Filter and Sort', 'Join, Subtract and Group', 'Datamash', 'GENOMIC FILE MANIPULATION' (highlighted), 'FASTA/FASTQ', 'FASTQ Quality Control', 'SAM/BAM', 'BED', 'VCF/BCF', 'Nanopore', 'Convert Formats', 'Life-Over', 'COMMON GENOMICS TOOLS' (highlighted), and 'Interactive tools'. At the bottom, the URL 'usegalaxy.org' is visible.

The screenshot shows a warning banner at the top: 'This dataset is large and only the first megabyte is shown below. Show all | Save'. Below the banner is a list of genomic data entries, each with a tabular representation of a sequence. The entries are labeled with IDs like @1/1, @2/1, @3/1, @4/1, @5/1, @6/1, @7/1, @8/1, @9/1, and @10/1. The sequences are composed of nucleotide bases (A, C, G, T) and some entries include quality scores or additional annotations.

The screenshot shows the 'History' panel on the right side of the Galaxy web interface. It has a search bar and a list of datasets. The datasets are labeled with IDs and names, such as '71 : SRX2638157:reverse', '70 : SRX2638157:forward', '69 : SRX2638160:reverse', '68 : SRX2638160:forward', '65 : SRX2638150:reverse', '64 : SRX2638150:forward', '61 : SRX2638159:reverse', '60 : SRX2638159:forward', '57 : SRX2638161:reverse', '56 : SRX2638161:forward', '51 : SRX2638162:reverse', '50 : SRX2638162:forward', '47 : SRX2638163:reverse', and '46 : SRX2638163:forward'. Each entry has icons for viewing, editing, and deleting the dataset.

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

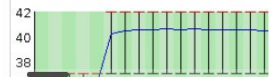
Summary

- ✓ 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)

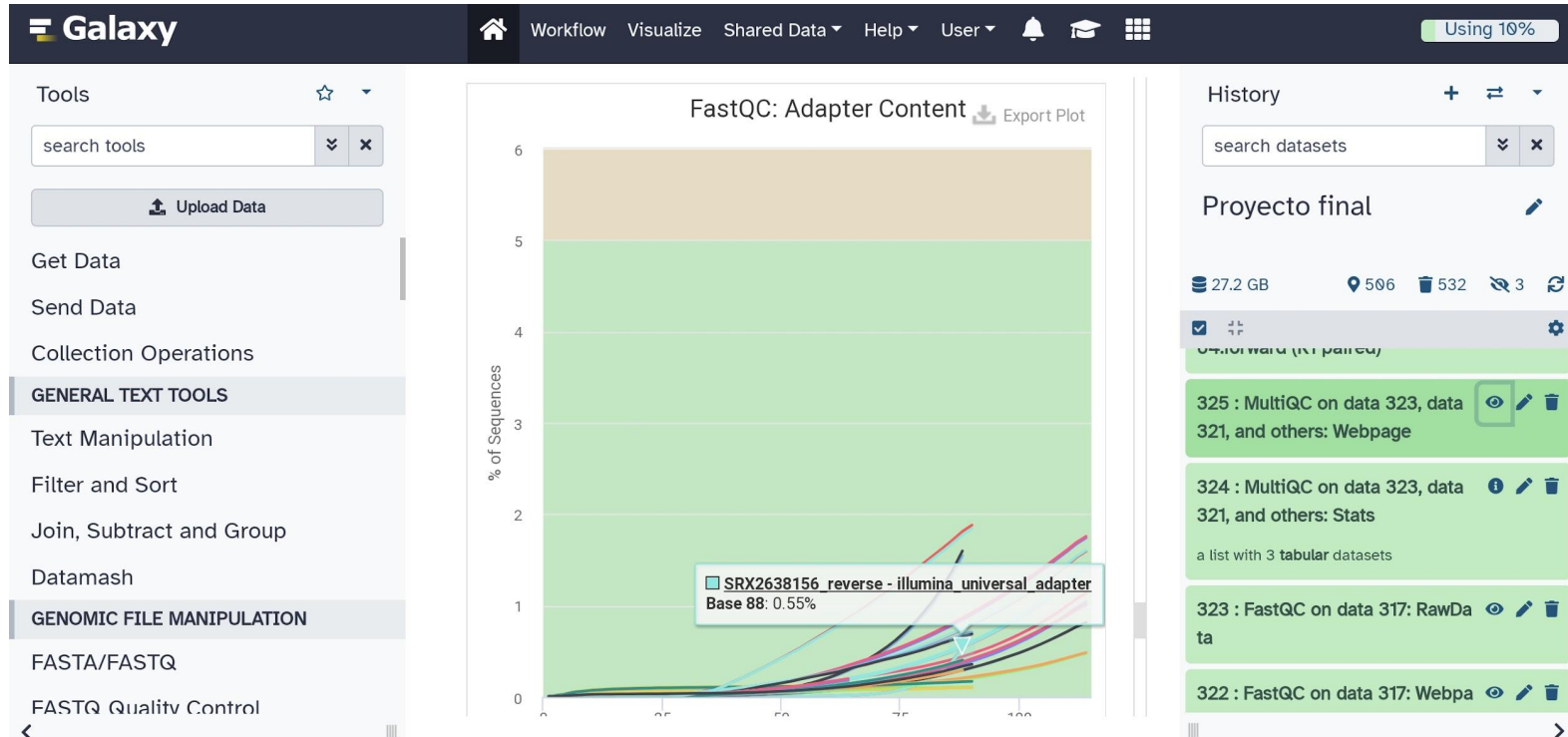
History search datasets

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, sequence data is displayed in columns, with headers like '@1/2', '@3/2', '@4/2', '@6/2', '@7/2', and '@8/2'. The right sidebar shows a 'History' section with a search bar and a list of datasets. The top dataset is 'Proyecto final' (27.2 GB, 506 locations, 532 deletions, 3 views). Below it are several Trimmomatic datasets: '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 has 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 central area shows a table with 6 columns: Column 1, Column 2, Column 3, Column 4, Column 5, and Column 6. The table contains 10 rows of data, including values like 8.85, 267325, 267325, U, 0, and unclassified.

The right sidebar features a 'History' section with a search bar and a list of datasets. The top entry is 'Proyecto final' with a size of 27.2 GB, 506 locations, 532 files, and 3 views. Below it are several entries for 'Kraken2 on data 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'.

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

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 menu items 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 for '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), with sub-items for 'FASTA/FASTQ' and 'FASTQ Quality Control'. The main workspace features a table with 8 columns: Column 1, Column 2, 3, 4, 5, Column 6, 7, and Column 8. The table contains genomic data rows with headers like CHROM, POS, TYPE, REF, ALT, EVIDENCE, FTYPE, and STRAND. The right sidebar shows a 'History' panel with a search bar and a list of datasets. The top dataset is 'Proyecto final' (27.2 GB, 506 locations, 532 deletions, 3 views). Below it are four green-highlighted dataset entries, each with a description of a Snippy workflow step and icons for viewing, editing, and deleting.

Column 1	Column 2	3	Column 4	Column 5	Column 6	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	+

History

search datasets

Proyecto final

27.2 GB 506 532 3

700 : snippy on data 386, data 385, and data 318 mapped reads (bam)

699 : snippy on data 386, data 385, and data 318 snps table

698 : snippy on data 386, data 385, and data 318 snps vcf file

697 : snippy on data 402, data 401, and data 318 mapped reads (bam)

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 data objects. The first object is a table with the following header and data:

##INFO=<ID=LOF,Number=.,Type=String,Description="Predicted loss of function							
##INFO=<ID=NMD,Number=.,Type=String,Description="Predicted nonsense medi							
#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=

The second object is a table with the following header and data:

##INFO=<ID=LOF,Number=.,Type=String,Description="Predicted loss of function							
##INFO=<ID=NMD,Number=.,Type=String,Description="Predicted nonsense medi							
#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=

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 it are several datasets created using the 'TB Variant Filter' tool on data 6, 65, and 422, with sizes ranging from 62 to 716. The 'TB Variant Filter' tool is highlighted in the history list.

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', 'Datemash', 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

Datemash

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

Proyecto 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. (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!