


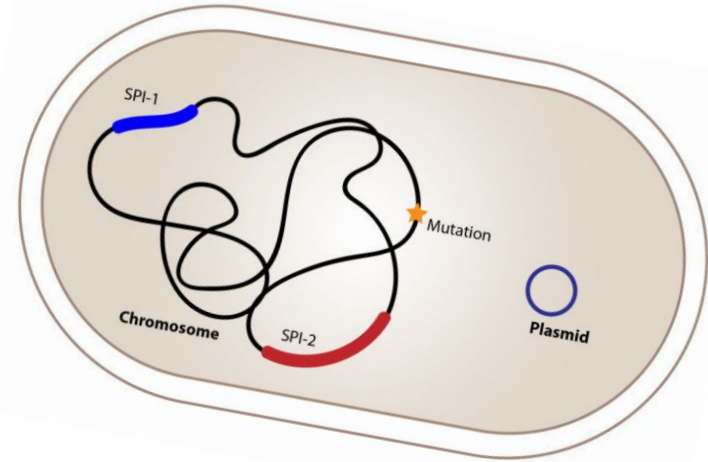
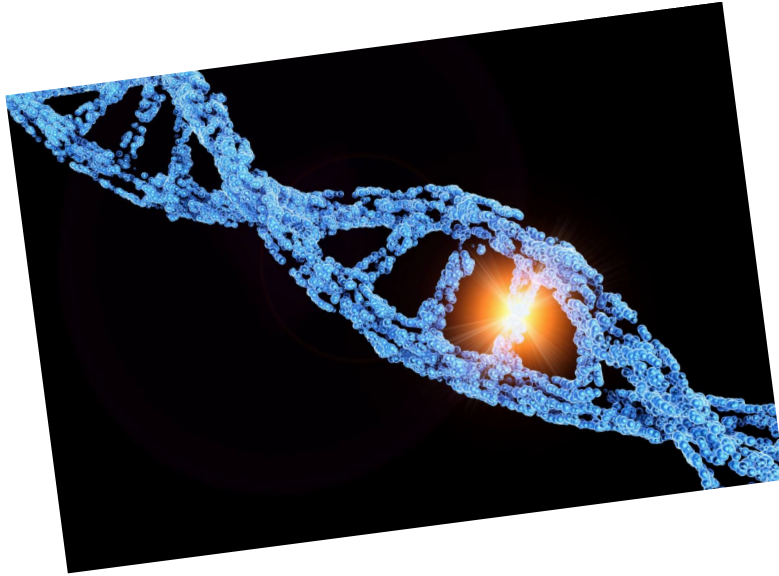
Resistencia a antibióticos en *Mycobacterium tuberculosis*



Cervantes López Ximena
Alamina Ramírez Sebastián
Hernandez Gasca Adriana

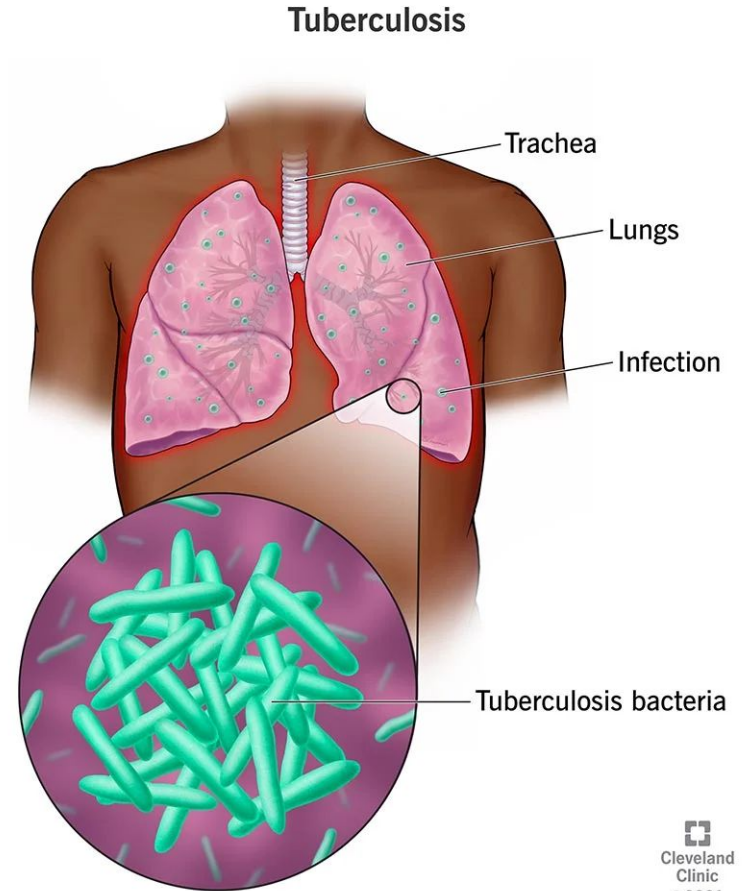
PREGUNTA DE INVESTIGACIÓN

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



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.



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.



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), and REFERENCE (1. Lew, J.M., Kapopoulou, A., Jones, L.M. and Cole, S.T. TuberculList--10 years after Tuberculosis (Edinb) 91 (1), 1-7 (2011). 2. Camus, J.C., Pryor, M.J., Medigue, C. and Cole, S.T. Re-annotation of the genome sequence of Mycobacterium tuberculosis). The right sidebar contains a 'Change region shown' dropdown, a 'Customize view' dropdown, an 'Analyze this sequence' section with links to 'Run BLAST' and 'Pick Primers', and a 'Related information' section with links to 'Assembly', 'BioProject', 'Protein', 'PubMed', 'Taxonomy', 'Components (Core)', 'Full text in PMC', 'Gene', 'Genome', and '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)

 **National Library of Medicine**
National Center for Biotechnology Information

Log in

SRA

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
TTGACCGATGACCCGGTTACAGGCTTCACACAGTGTGGAACGGGTCTGCCAACTTAACGGCGACC
CTAAGGTTGACGACGGACCCAGCAGTGTATGCTAATCTCAGCGCTCCGCTGACCCCTCAGCAAAGGGCTTG
GCTCAATCTCGTCCAGCCATTGACCATCGTCGAGGGGTTTGCTCTGTTATCCGTCGCCGAGCAGCTTTGTC
CAAAACGAAATCGAGCGCCATCTGCGGGCCCCGATTACCGACGCTCTCAGCCGCGACTCGGACATCAGA
TCCAACCTCGGGTCCGCATCGTCCGCGCGCGACGACGAAGCCGACGACACTACCGTCCGCCCTTCCGA
AAATCTCGTCTACCCATCGCCAGACACCAACCGACAACGACGAGATTGATGACAGCGCTGCGGCACGG
GGCGATAACGACGACAGTTGGCCAAGTTACTTACCAGCGCGCCGACAATACCGATTCCGCTACCGCTG
GCGTAACGAGCTTAACCGTCGCTACACCTTTGATACGTTCTGTTATCGGCGCTCCAACCGGTTCCGCGCA
CGCCGCGCTTGGCGATCGCAGAAGCACCCGCGCGCTTACAACCCCTGTTTCATCTGGGGCGAGTCC
GGTCTCGGCAAGACACCTGCTACACGCGGACGGCAACTATGCCAACGGTTGTTCCCGGGAATGCGGG
TCAAAATATGTCTCACCGAGGAATTCACCAACGACTTCATTAACTCGCTCCGCGATGACCGCAAGGTCGC
ATTCAAACGCGAGCTACCGGACGTAGACGTGCTGTTGGTCGACGACATCCAATTCAATTGAAGGCAAGAG
GGTATTCAAGAGGAGTCTTCCACACCTTCAACACCTTGACAATGCCAACAAGCAAAATCGTCATCTCAT
CTGACCGCCCAACCAAGCAGCTCGCCACCTCGAGGACCGGCTGAGAACCCTGTTGAGTGGGGGCTGAT
CACTGACGTACAACCACCGAGCTGGAGACCCGATCGCCATCTTGCGCAAGAAAGCACAGATGGAACGG
CTCGGGTCCCCGACGATGTCTCGAACTCATCGCCAGCAGTATCGAACGCAATATCCGTAACCTCGAGG
GCGCGCTGATCCGGGTACCGCGTTGCTCTATTGAACAAAACCAATCGACAAAGCGCTGGCCGAGAT
TGTGCTTCGCGATCTGATCGCCGACGCCAACCATGCAAACTAGCGGCGGCGACGATCATGGCTGCCACC
GCCGAATACTTCGACACTACCGTCAAGAGCTTCGCGGGCCCGCAAGACCAGGACTGGCCAGTAC
GACAGATTGGGATGTACCTGTGCTGAGCTACCGATCTTTCTGTTGCCAAAATCGGCCAAGCGTTCCG
CCGTGATCACACAACCGTATGTACGCCCAACGCAAGATCTGTCCGAGATGGCCGAGCGCGTGAAGTC
TTTGATCAGCTCAAAGAATCACCCTCGATCCGTGAGCGCTCCAAGCGCTAGCACGGCGTGTCTTCC
GACAACGTTCTTAAAAAACTTCTCTCTCCAGGTACACCAAGTACAGAGATTGGCTGTGAGTGTGCT
GTGCAACAAACCGGCACAGACTCATACAGTCCCGCGGTTCCGTTACAAACCCACGCTCATCCCCACCG
ACCCAACACACACCCACAGTCATCGCCACCGTATCCACAATCCGACCGACGTGACCTGCACCAAGA
CCAGACTGTCCTCAAACTGCACACCTCTAATACTGTTACCGAGATTCTTCTGCTGTTGTTCTTGGAAA
GACAGCGCTGGGGATCGTTGCTGGATACCAACCCGATAAATGGCTCGTCGCGGTGGTCAGAGGTCAAT
GATGAATTTCAAGTTGACGTGAGAAGCTTACGGTTGTTGTTGACTGCTGTTGCGGCGCTGTTGCGG
GTCACGCGTCATGGGATTCGTGTTGGCAGTCCACGCTAGCGGGGCGTAGCCACGGGATCGAACTC
ATCGTGAGGTGAAAGGCGCAATGGACGCGGCTACGACAAGAGTTGGCTCACCGACTTGACGTTTCGTT
TGCTACGAGAGTCTTTCCGCGATGCGGTGCTGTTGGTGGCTAAAAATCTCGACGACGGCCGCGGTGCC
GGTGCTCTCCGGCGTGTGTTGACCGGCTCGGACAACGGTCTGACGATTTCCGGATTGACTACGAGGTT
TCCCGCGAGGCCAGGTTGGCGCTGAAATGTTTCTCCTGGAAGCGTTTATGTTTTCGGCCGATTGTTGT
```

ence: 4.16MB (COMPLETE)

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

usegalaxy.org

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

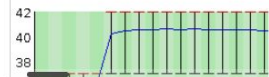
Summary

- ✓ 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)

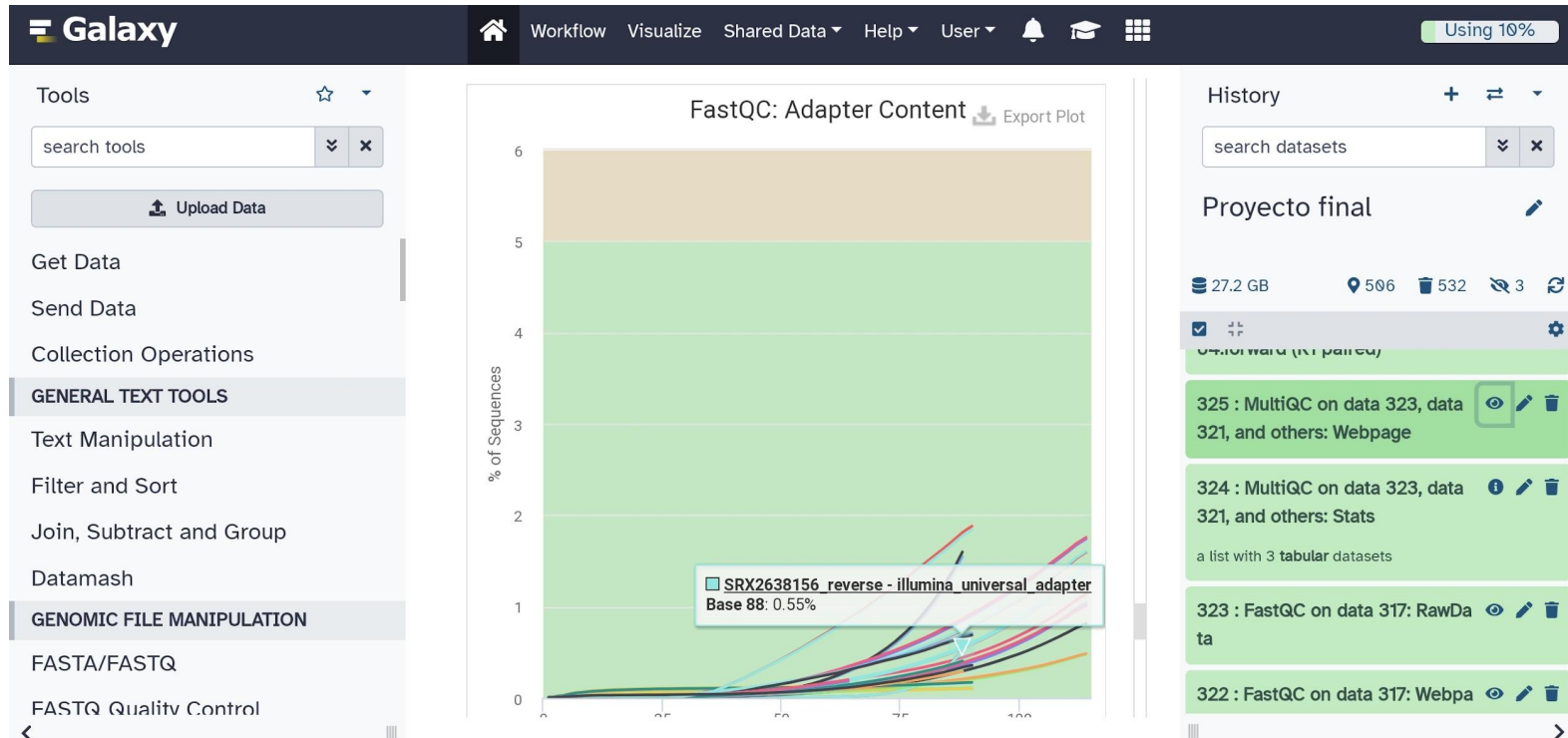
History search datasets

Proyecto final

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

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.

Galaxy

Home

Workflow

Visualize

Shared Data

Help

User

Notifications

Help

Galaxy

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

This dataset is large and only the first megabyte is shown below.

Show all | Save

```

@1/2
TTTAAAGCAGTTATGGAAGGTAAGCAGGTAGCATTCTTGCTACGACAATATTAGCGCAACAACATT
+
AAFFFK, FKKKKKKKKKKAKKK<FFFKKKKKKKKKFKKKKKKKKKKKKKKKKKKKKKKKKKKKKK
@3/2
AGACCGACGAAGAAAAGGCGATCCGCGATCGCTATTCAAAGATCCTCGGCAGCGCGGTGAACCCGTTCT
+
<AFFFFKKK<KKKKKKKKKKFKKKKKKKFKKKKKKKFKKKKKKKFKKKKKKKKKKAFKFK
@4/2
CCCGGTGACGCCCGCCATGAGGTGCGCGCGCATCGAGGCACTCGGCCGCCGCGTGTCTNACTGGCAA
+
AAFFFKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK#FKKKKKKKKKKKKKKKKKKK
@6/2
AGCACCAGCATGACCCGTCGGTGCGCGTCAGCACTTCGCACGACACCTNCTGTTGACGA
+
AAFFFKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK#KKKKKKKKKK
@7/2
GGCGATGAAGCGCTGCTTGTGCGCAGCTGGCGATCTGC
+
A<FFAA7FKKKAFFKFAFFKKKKKKFKKAFKKKKKK
@8/2

```

History

search datasets

Projecto final

27.2 GB

506

532

3

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)

492 : Trimmomatic on SRX26381

50:reverse (R2 unpaired)

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 first dataset is 'Proyecto final' (27.2 GB, 506 locations, 532 deletions, 3 views). Below it are several Kraken2 reports, including '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 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', 'Datamash', and 'GENOMIC FILE MANIPULATION' (highlighted), which includes 'FASTA/FASTQ' and 'FASTQ Quality Control'. The main workspace shows a table with 8 columns: Column 1, Column 2, 3, 4, 5, Column 6, 7, and Column 8. The table contains genomic data rows. The right sidebar features a 'History' section with a search bar and a list of datasets. The first dataset is 'Proyecto final' (27.2 GB, 506 locations, 532 deletions, 3 views). Below it are four green entries for 'snippy' workflows, each with a view icon, a pencil icon, and a trash icon.

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.

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 with a workflow titled "Proyecto final". The main panel shows the output of the TB-Profiler-Profile tool, which includes a "Resistance variants report" and an "Other variants report".

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					

The right sidebar shows the "History" section with a list of datasets, 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".

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 relacionadas con respuestas (principalmente respecto a la resistencia) ante fármacos.
- Resistencia a estreptomicina.
- Con el paso de los años, las bacterias han obtenido una resistencia cada vez más potente ante fármacos..
- Las herramientas bioinformáticas permiten el análisis exhaustivo de variantes o SNPs.

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!