

# Utilización de los datos para descubrir y diagnosticar la resistencia de las bacterias a los medicamentos

Imperial College  
London

Leonid Chindelevitch

Departamento de epidemiología de las enfermedades infecciosas

Centro por la análisis de las enfermedades infecciosas globales, MRC

# Plan de la exposición

Motivación

La situación  
actual

El trabajo  
que hacemos

Una visión  
del futuro

# El peligro de la RAM

- Se estima que la resistencia a los antimicrobianos (RAM) podría ser responsable por **10 millones de muertes anualmente** en el año 2050
- Se estima que la RAM fue directamente asociada con más de **1.27 millones de muertes** en 2019; la situación sigue empeorando (¡gracias también a Covid-19!)
- La última vez que se descubrió una nueva clase de antibióticos fue hace **más de 40 años**, posiblemente con la excepción de tuberculosis; para conservar los que tenemos, hay que utilizar los diagnósticos
- Sin antibióticos, las operaciones comunes en medicina (**transplante de órganos, quimioterapia, parto, ...**) tienen riesgo de complicaciones



# El papel de la genética



- Las causas y los mecanismos de la RAM son principalmente genéticos
- La información genética y genómica se utilice para controlar la RAM:



## Diagnósticos

¿Cuales bacterias (virus, hongos) infectan a mi paciente? ¿Cuales medicamentos puedo utilizar?



## Tratamiento

¿Cuales son las mejores dianas farmacológicas?



## Vigilancia

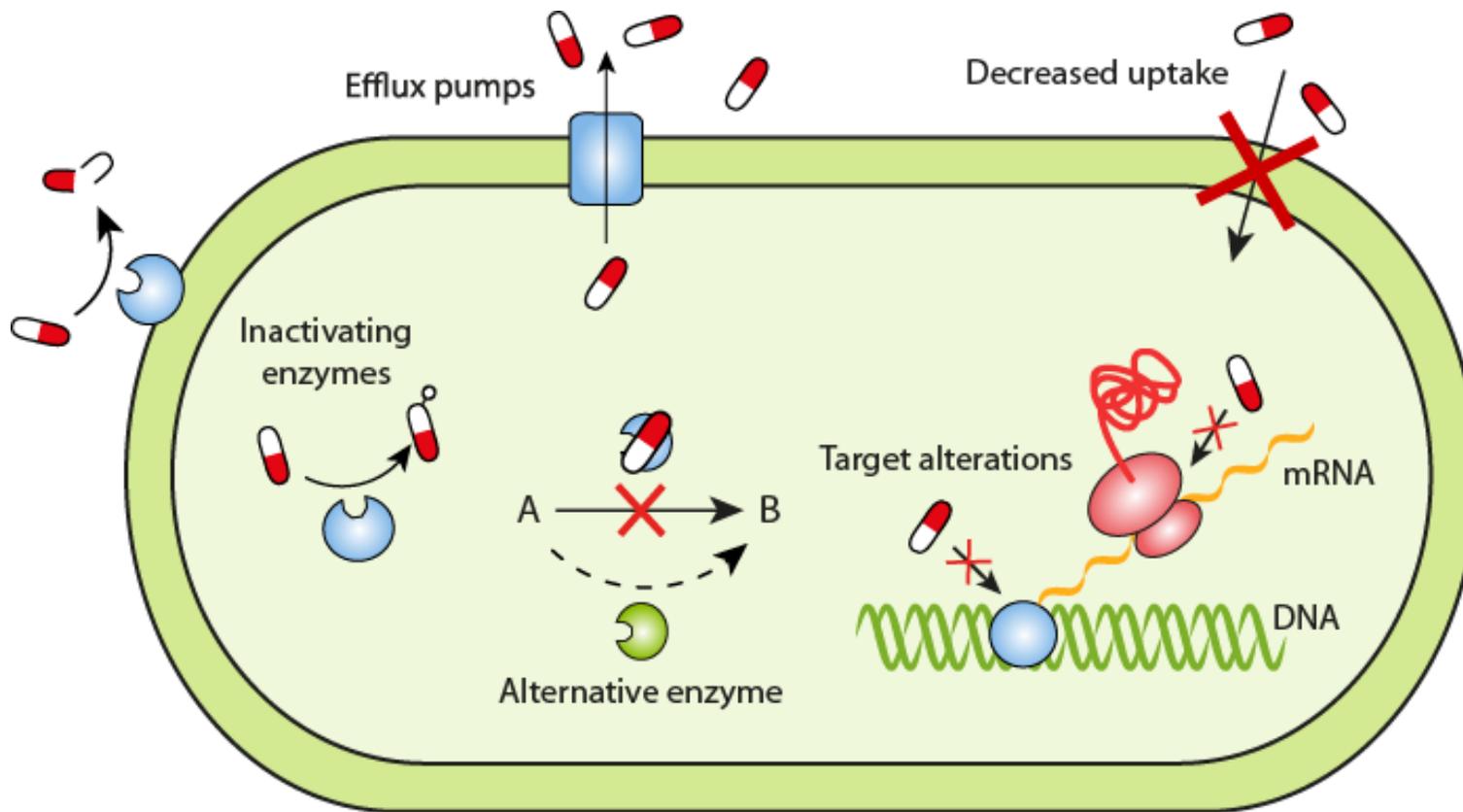
¿Cómo y dónde se transmite la RAM en mi población?



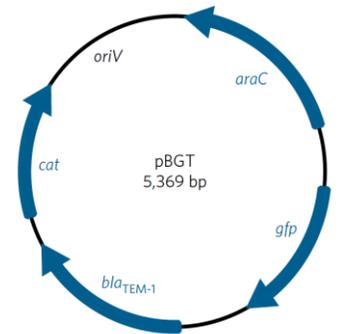
## Prevención

¿Cuales mecanismos de RAM son los más comunes en mi país?

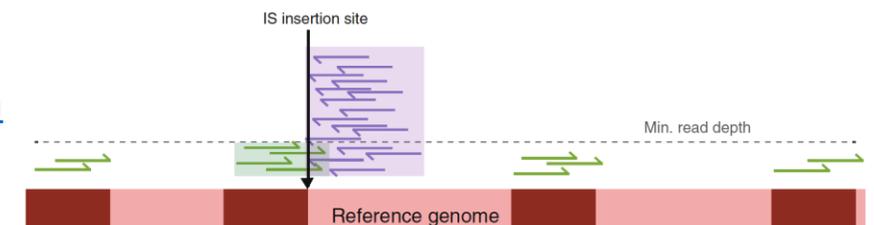
# Mecanismos de la RAM y determinantes genéticos



- Mutaciones puntuales CAT -> GAT
- Inserciones, deleciones GAT -> GATTA
- Genes o plásmidos adquiridos

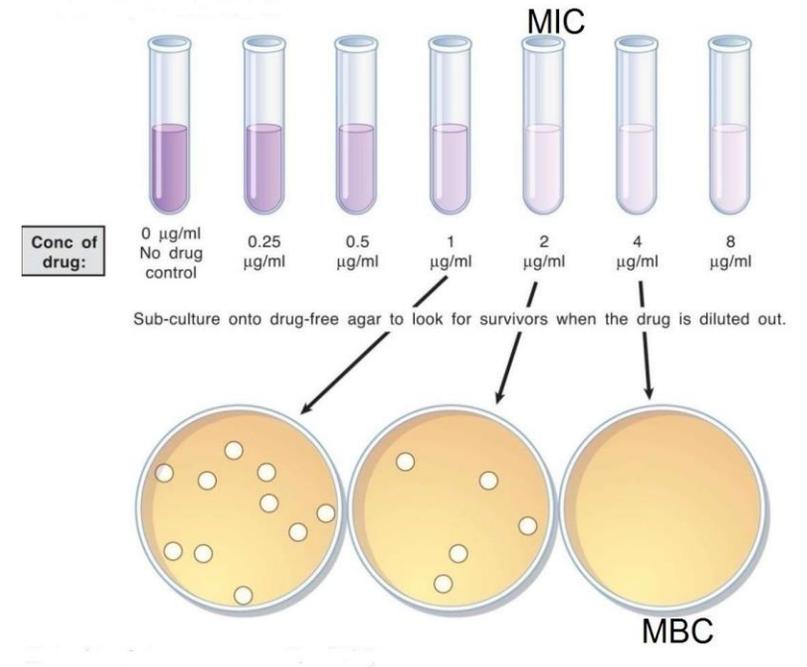


- Disrupción genética como elementos SI



[reactgroup.org/toolbox/understand/antibiotic-resistance/resistance-mechanisms-in-bacteria](https://reactgroup.org/toolbox/understand/antibiotic-resistance/resistance-mechanisms-in-bacteria)

# Medidas de la RAM: datos categoricos (R/S) y cuantitativos

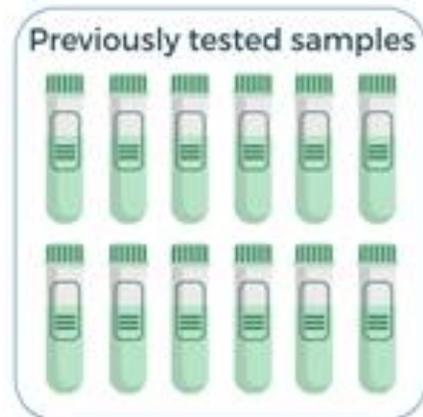


- CIM (concentración inhibitoria mínima): la cantidad de antibiótico necesitada para que **pare el crecimiento** de la población bacteriana
- CBM (concentración bactericida mínima): cantidad de antibiótico necesitada para **erradicar** la población bacteriana enteramente



# ¿Como aprende un algoritmo a diagnosticar la RAM?

Muestras probadas anteriormente

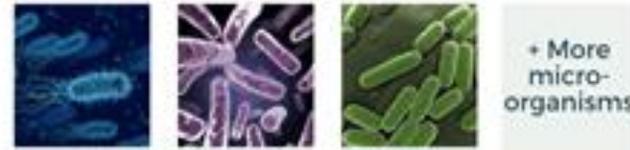


Data



Microbios

Train the Machine Learning Algorithm



		R		
	X	S	X	
		R		
+ More anti-microbials				

Antimicrobianos

Id: identificación del microbio  
AST: Prueba de susceptibilidad

Id: Identification testing result

AST: antimicrobial sensitivity testing result

Muestra probada nuevamente



Data



Id & AST results (interpreted data)

**NB:** antes de la **identificación** (Id) y de la **prueba de sensibilidad** (AST) es importante usar un **diagnóstico diferencial** para determinar si la enfermedad es causada por una bacteria, un virus, u otro microbio

# El trabajo que hacemos: aprendizaje automático

8 000 aislados



Matriz con más de 100 000 mutaciones

$$\begin{matrix}
 & \text{SNP1} & \text{SNP2} & \text{SNP3} & \text{SNP4} & \text{SNP5} & \text{SNP6} & \text{SNP7} & \text{SNP8} & \dots \\
 \text{ISO 1} & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & \dots \\
 \text{ISO 2} & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots \\
 \text{ISO 3} & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & \dots \\
 \text{ISO 4} & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & \dots \\
 \text{ISO 5} & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & \dots \\
 \vdots & \vdots
 \end{matrix}$$

$A$

Optimización automática

El objetivo es seleccionar las mutaciones que optimicen la sensibilidad y la especificidad

$$y = A \vee w, \quad w \in \{0, 1\}^n$$

Ejemplo de regla:

Si `gyrA_A90V`

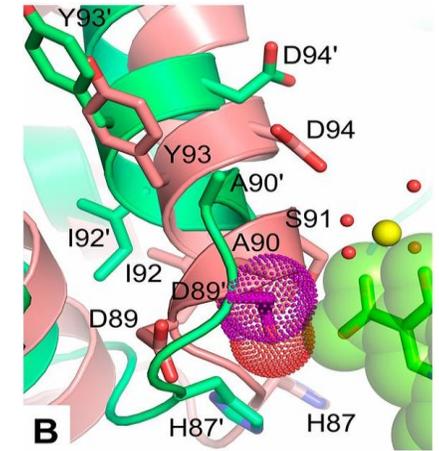
o `gyrA_S91P`

o `gyrA_D94A`

o `gyrA_D94G`

o `gyrA_D94Y`

Entonces resistente a la ciprofloxacina



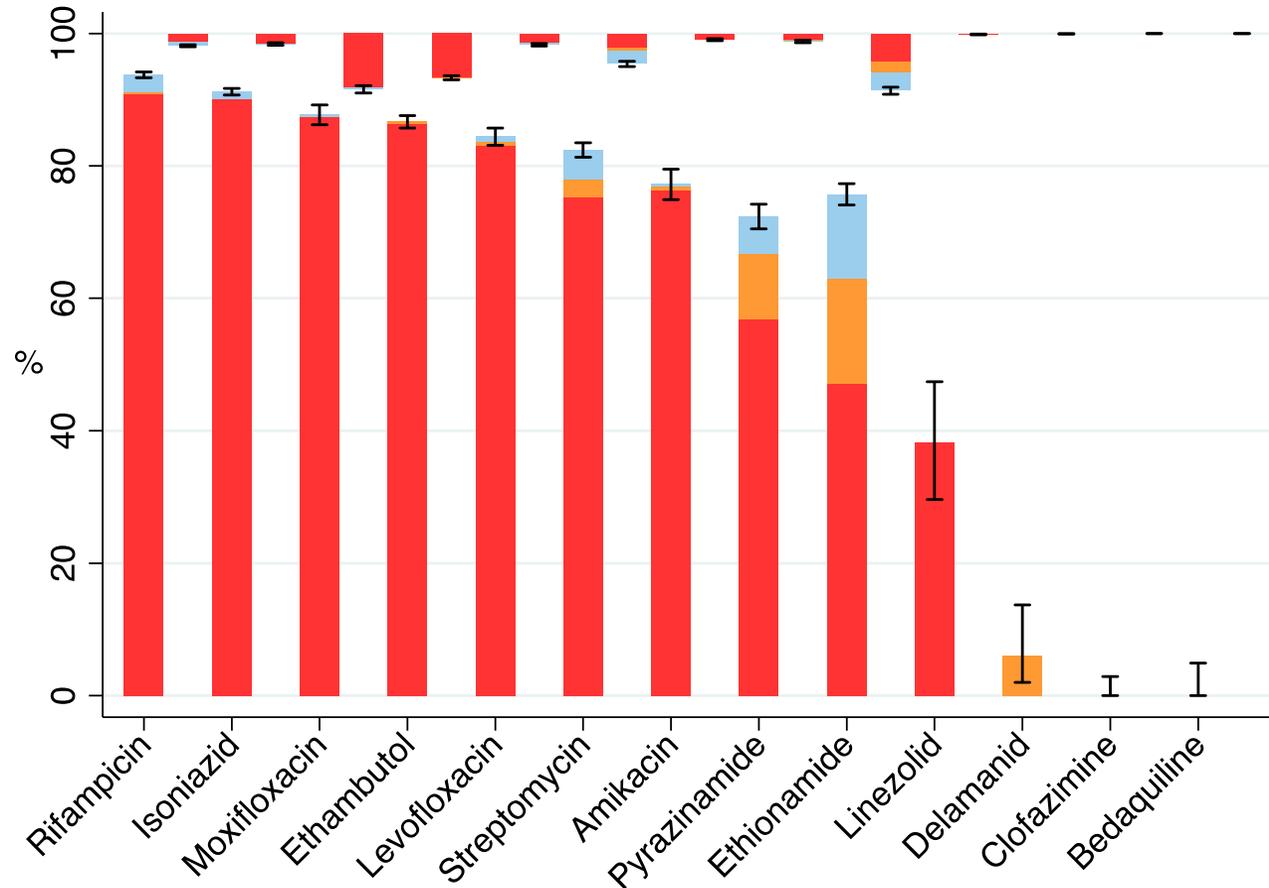
Zabeti et al. (2021). **INGOT-DR: an interpretable classifier for predicting drug resistance in *Mycobacterium tuberculosis***. *Algorithms Mol Biology*; 16:17

Blower et al (2016). Crystal structure of gyrase-fluoroquinolone complexes from *M. tb*. *PNAS* 113:7

# El trabajo que hacemos: curación de datos

- Extrajimos and procesamos todos los datos de RAM en los **bases de datos más importantes** (CDC, COMPARE-ML-AMR, microreact, NARMS, NDARO, PathogenWatch, PATRIC)
- Llevamos a cabo una búsqueda sistemática de la literatura (más de **35 000 publicaciones**; más de 150 publicaciones con datos utilizables)
- Identificamos más de **200 000 genomas completos** con un promedio de **5 fenotipos** - un total de más de **1 millón** de puntos de datos

# El trabajo que hacemos: catálogos de determinantes genéticos de la RAM



- Catalogo para la **tuberculosis**: versión 1 publicada, **versión 2** en preparación (mucho más datos)
- Catalogo para otras bacterias **prioritarias según la OMS**: en preparación
- **NB**: Las descubiertas tienen que ser confirmadas en laboratorio

# Una visión del futuro

---

Se desarrollan y se adoptan los diagnósticos rápidos (diferencial, identificación y prueba de sensibilidad)

Cada paciente con una enfermedad posiblemente infecciosa recibe un tal diagnóstico antes de iniciar un tratamiento, y también durante el tratamiento

Los antibióticos que ya tenemos se quedan eficaces durante el mayor tiempo posible... y así ganamos el tiempo para desarrollar nuevas soluciones a la RMA



# Agradecimientos; financiación



Curacion de datos genotípicos y fenotípicos

Tecnologías de datos en RAM; Diez reglas sencillas para compartir los datos de la RAM



Elita Jauneikaite



Nicole Wheeler



Maarten van Dongen



Anita Suresh



Swapna Uplekar

BenchmarkDR



Niklas Stotzem



Fernando Guntoro



John Lees



Antonio Pedrotta



Atharv Naik