



# Canadian Bioinformatics Workshops

[www.bioinformatics.ca](http://www.bioinformatics.ca)

[bioinformaticsdotca.github.io](https://bioinformaticsdotca.github.io)



# Module 6: Antimicrobial Resistant Gene (AMR) Analysis



Andrew G. McArthur

Infectious Disease Genomic Epidemiology

April 18-21, 2023



McMaster  
University  
**HEALTH SCIENCES**



Michael G. DeGroote  
INSTITUTE FOR INFECTIOUS DISEASE RESEARCH

David Braley Centre  
for Antibiotic Discovery

# Learning Objectives

- By the end of this lecture, you will:
  - Understand the molecular basis of antimicrobial resistance
  - How AMR reference sequence databases are created and used
  - How to predict antimicrobial resistance genes (ARGs) from genome assemblies
  - How to predict antimicrobial resistance genes (ARGs) from metagenomic sequencing reads
  - Understand the promise and pitfalls of AMR phenotype prediction

Bacteria are evolving drug resistance faster than we can discover new drugs

Jorge Segovia @jorgesegovr ...

La pandemia silenciosa

Translate Tweet

Cultivo standard Positivo Se aislan >100.000 UFC/mL de:

Klebsiella pneumoniae

Antibiótico	Klebsiella pneumoniae
Ampicilina	R
Amoxicilina/Clavulánico	R
Cefuroxima	R
Cefotaxima	R
Cefepime	R
Ertapenem	R
Imipenem	R
Gentamicina	R
Ciprofloxacino	R
Trimetoprim/Sulfametoazol	R
Fosfomicina	R
Nitrofurantoina	R
Amikacina	R
Ceftazidima / Avibactam	S
Ceftolozano/Tazobactam	R
Meropenem E-Test	R
Mecanismo de resistencia	Carb OXA-48

2:34 PM · Sep 13, 2021 · Twitter for Android

Review Cell PRESS

## The emerging NDM carbapenemases

Patrice Nordmann<sup>1</sup>, Laurent Poirel<sup>1</sup>, Timothy R. Walsh<sup>2</sup> and David M. Livermore<sup>3</sup>

<sup>1</sup> Service de Bactériologie-Virologie, Hôpital de Bicêtre, Institut National de la Santé et de la Recherche Médical (INSERM) Unité 914, Faculté de Médecine et Université Paris Sud, 78 rue du Général Leclerc, 94275 Le Kremlin-Bicêtre, France  
<sup>2</sup> University of Queensland, Centre for Clinical Research, University of Queensland, Brisbane, Australia  
<sup>3</sup> Health Protection Agency Microbiology Services – Colindale, London NW9 5EQ, UK

Carbapenems were the last β-lactam antibiotics to be effective against most Gram-negative bacteria. Now, carbapenemases are spreading, conferring resistance to carbapenems. These enzymes are metallo-β-lactamases (NDM) enzymes. Carbapenemases have been reported worldwide, mostly in Enterobacteriaceae, which are epidemiologically linked to the Indian subcontinent. They are where they occur widely in hospital infections, and also in contaminated food. The main type is NDM-1, but minor variants of the enzyme are present largely in Enterobacteriaceae, and also in non-fermenters and Vibrio spp. The spread of carbapenemases predominantly involves transfer among promiscuous plasmids. Bacteria with NDM-1 are typically resistant to all antibiotics, and reliable detection and identification are crucial.

**The serious threat of multidrug-resistant and untreatable gonorrhoea: the pressing need for global action to control the spread of antimicrobial resistance, and mitigate the impact on sexual and reproductive health**

Journal of Antimicrobial Chemotherapy (2009) 64, Suppl. 1, i29–i36  
doi:10.1093/jac/dkp255

Has the era of untreatable infections arrived?

David M. Livermore\*

Antibiotic Resistance Monitoring and Reference Laboratory, Health Protection Agency Centre for Infection, 61 Colindale Avenue, London NW9 5EQ, UK

were abandoned as first-line empiric treatment for gonorrhoea in the Asia-Pacific region already in the mid-to-late 1990s and, subsequently, in the USA, Europe and parts of Africa.<sup>3</sup> w2 w3–w6 Azithromycin-resistant *N. gonorrhoeae* emerged in the mid-to-late 1990s,<sup>4</sup> w7 and, subsequently, also high-level azithromycin resistance has been described from several countries.<sup>5</sup> w8–w11 Thus, since the early 2000s, third-generation cephalosporins (mainly cefixime and ceftriaxone) have been the sole class of antimicrobials recommended as first-line empiric treatment.<sup>5</sup> w2 w8 Worryingly, resistance and treatment failures to cefixime have been verified in Japan<sup>6</sup> and recently in Europe.<sup>7–9</sup> The recent report of a strain of *N. gonorrhoeae* in Japan that was highly resistant to the parenteral ceftriaxone, and associated with a probable treatment failure with ciprofloxacin,<sup>10</sup> the last remaining option for empiric treatment, sounded alarm bells. Significant future challenges to the prevention and control of gonococcal infections and their complications. This has also followed by the identification of a highly ceftriaxone-resistant strain in Japan<sup>11</sup> and in Spain.<sup>12</sup> Furthermore, many countries worldwide describe a decreasing

JAC

bioinformatics.ca

# When Antibiotics Fail

The growing cost of antimicrobial resistance in Canada



Council of  
Canadian  
Academies

Conseil des  
académies  
canadiennes



In 2018, approximately **26%** of infections were resistant to the drugs generally used to treat them.



By 2050, the rate of resistance is likely to grow to **40%**.

If AMR increases gradually from 26% to 40%, by 2050, the **cumulative** cost to Canada is estimated at:



**396,000 lives**



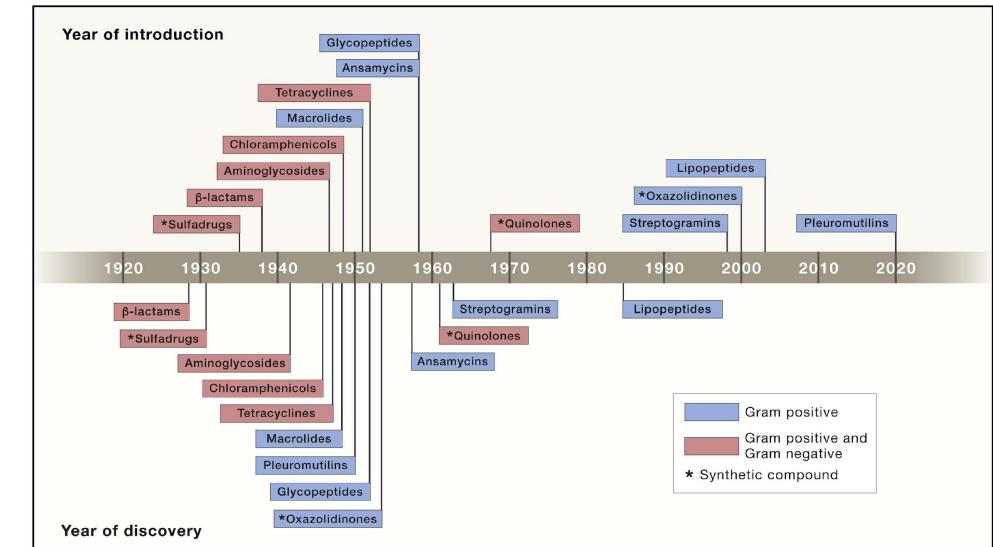
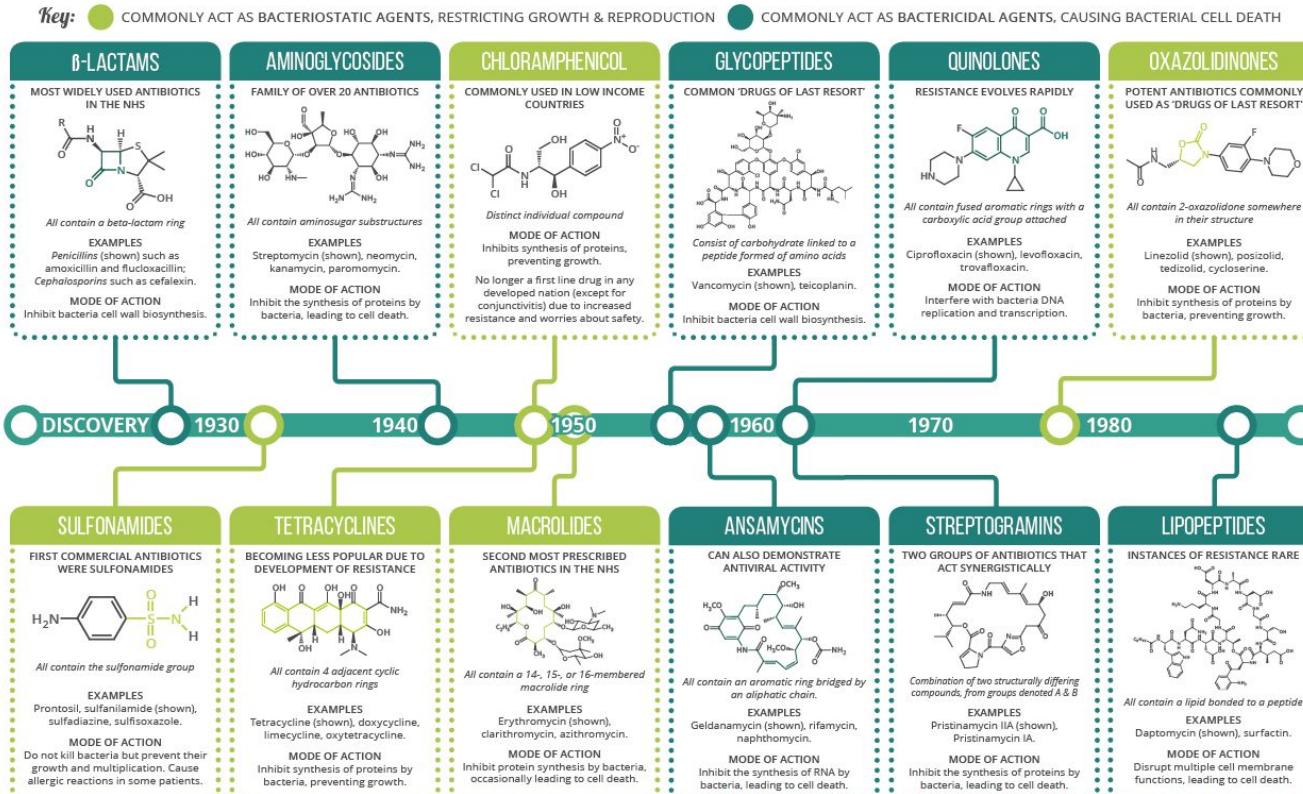
**\$120 billion in hospital costs**



**\$388 billion in GDP**

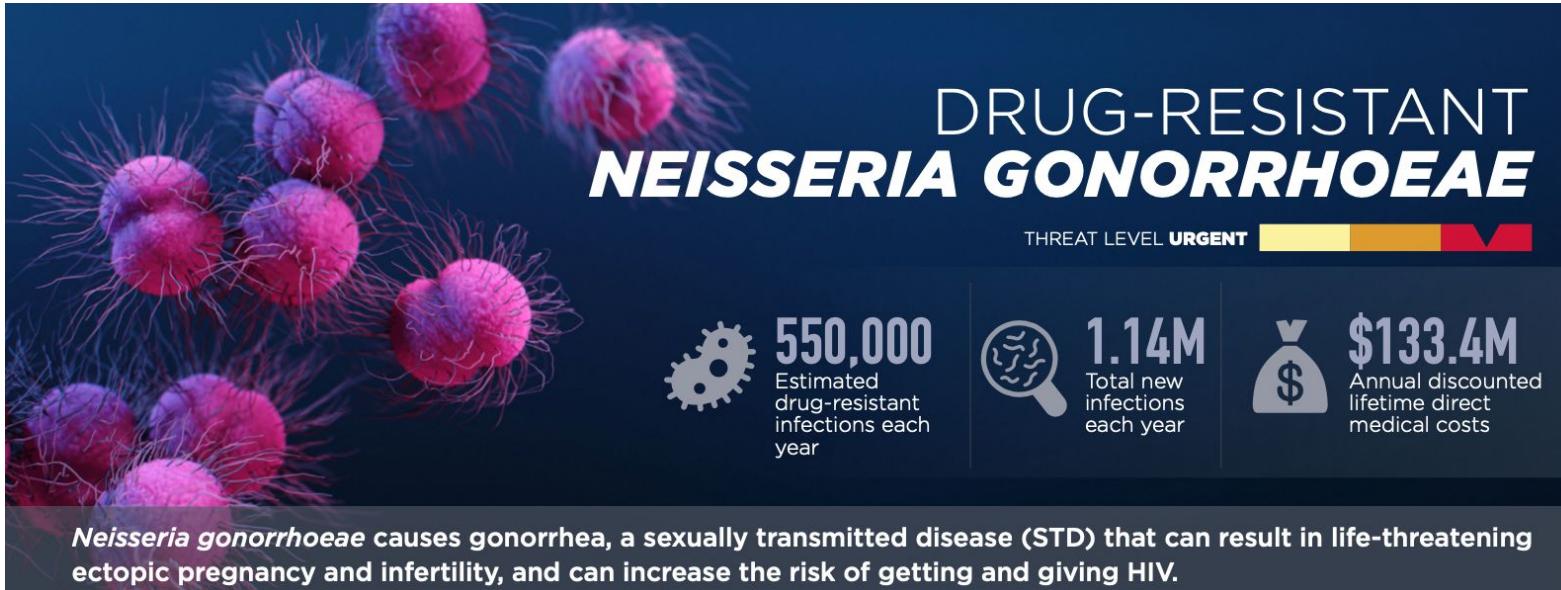
We are at risk of entering a 'post-antibiotic' era

## DIFFERENT CLASSES OF ANTIBIOTICS - AN OVERVIEW



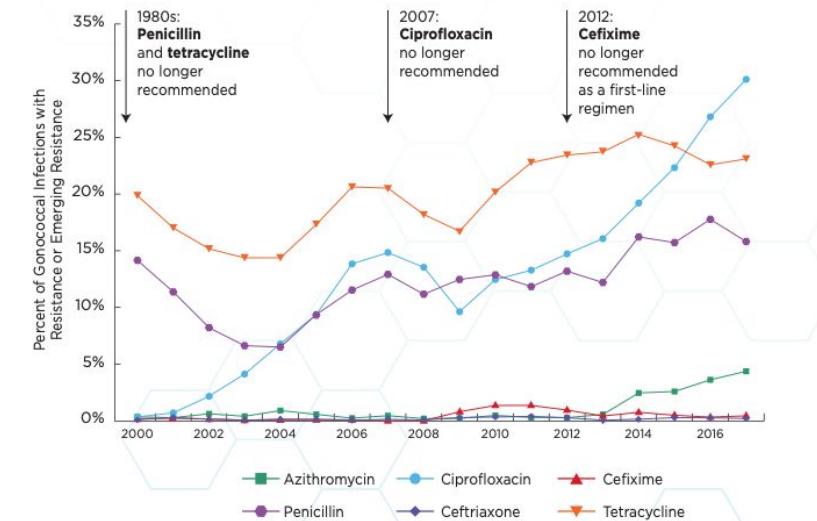
Kim Lewis. 2020. *Cell* 181(1): 29-45.

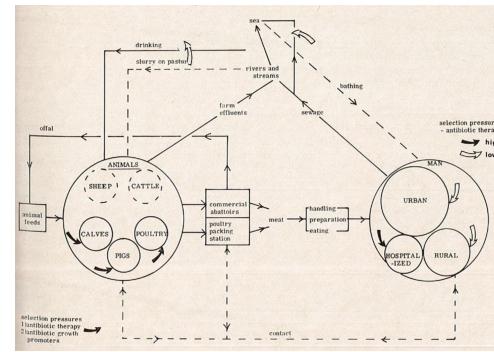
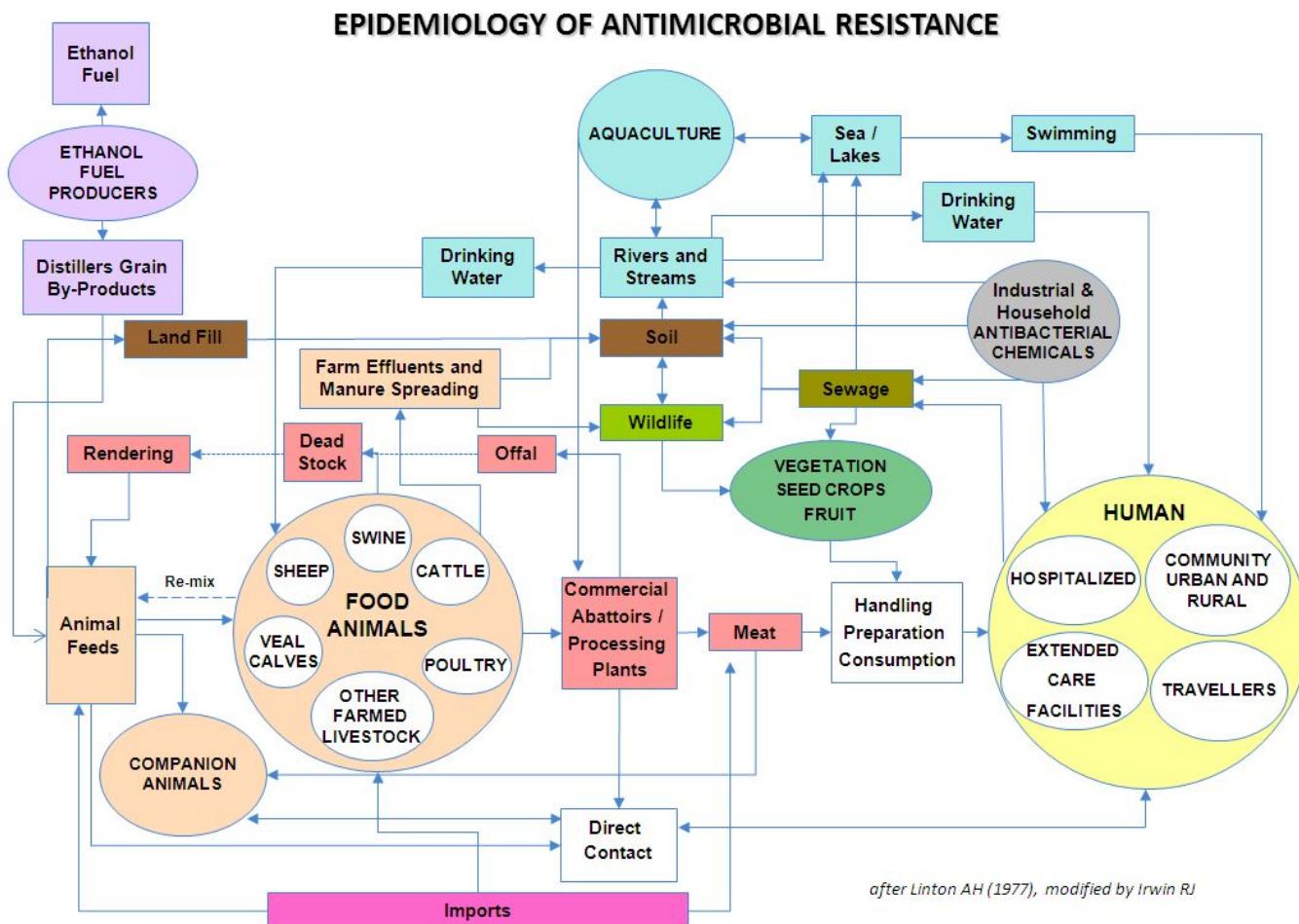
We are at risk of entering a ‘post-antibiotic’ era



## EMERGING ANTIBIOTIC RESISTANCE

Gonorrhea rapidly develops resistance to antibiotics—ceftriaxone is the last recommended treatment.





Linton. 1977. Vet Rec. 100: 354-360.

Government of Canada Gouvernement du Canada

MENU ▾  
Canada.ca > Genomics R&D Initiative > Projects

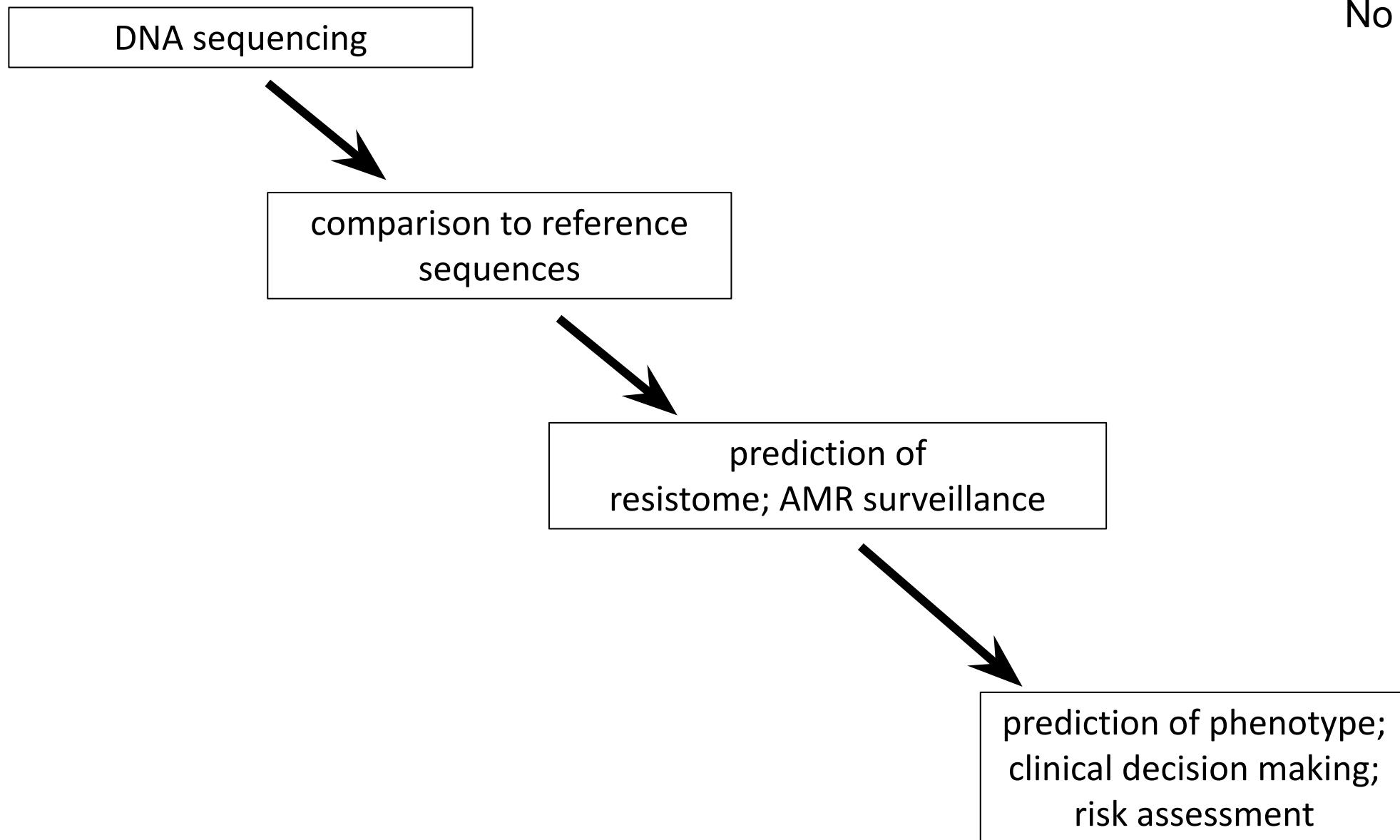
### Antimicrobial resistance (the AMR project)

From: Genomics R&D Initiative

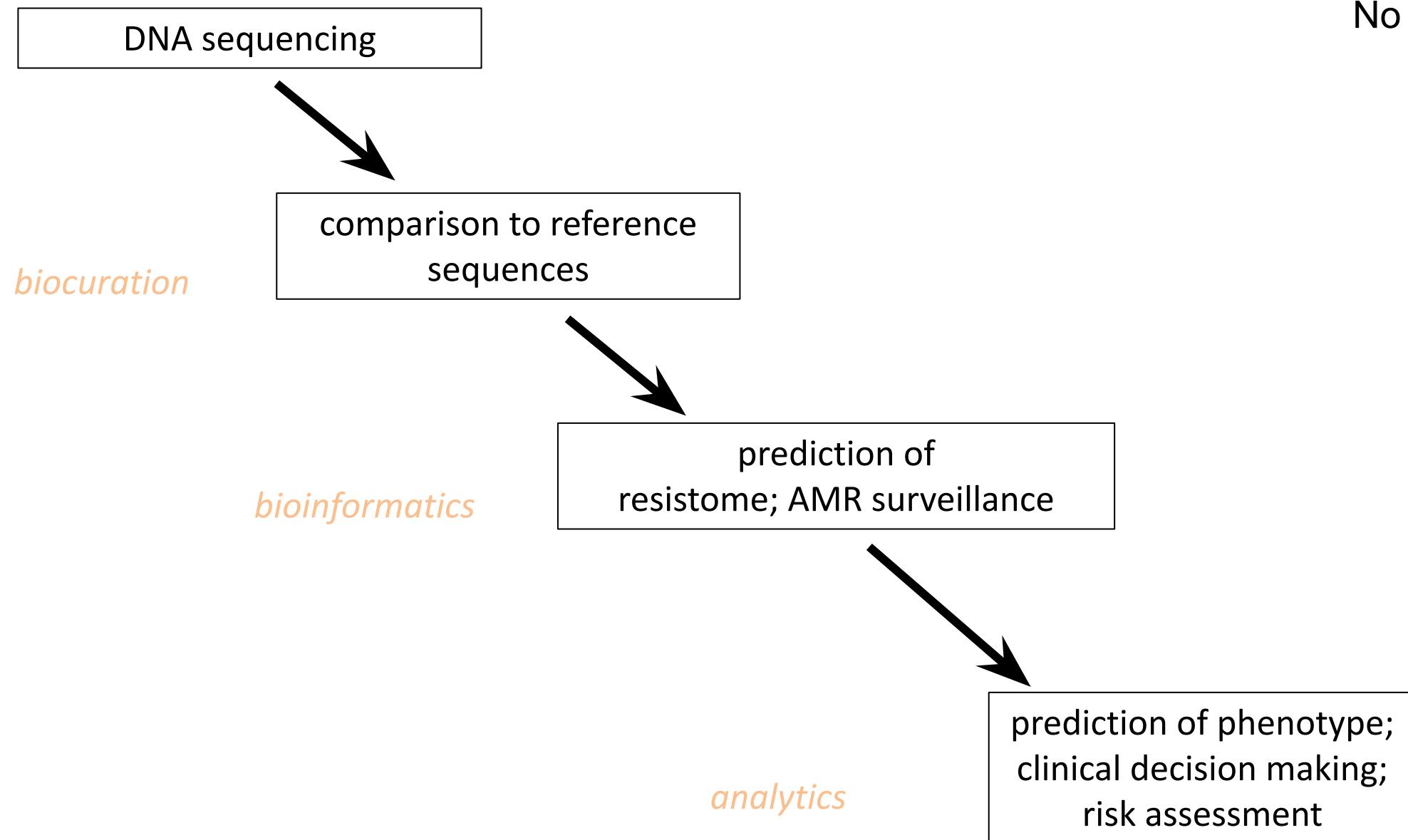
Funding period: 2016-2022  
Participating Departments and Agencies: Agriculture and Agri-Food Canada, Canadian Food Inspection Agency, Health Canada, National Research Council of Canada, Public Health Agency of Canada  
Lead: Ed Topp, AAFC  
Total GRDI funding: \$11,144,726

The development of resistance to antimicrobials by bacteria that were formerly sensitive is one of the most serious global health threats facing the world today. With no action, annual worldwide human deaths attributable to antimicrobial resistance could reach 10 million by 2050. The Antimicrobial Resistance project uses a genomics-based approach to understand how food production contributes to the development of antimicrobial resistance of human health concern, and explore strategies for reducing antimicrobial resistance in food production systems. It is a component of the Federal Action Plan for Antimicrobial Resistance and Use in Canada. The project involves scientists from 5 federal departments and agencies.

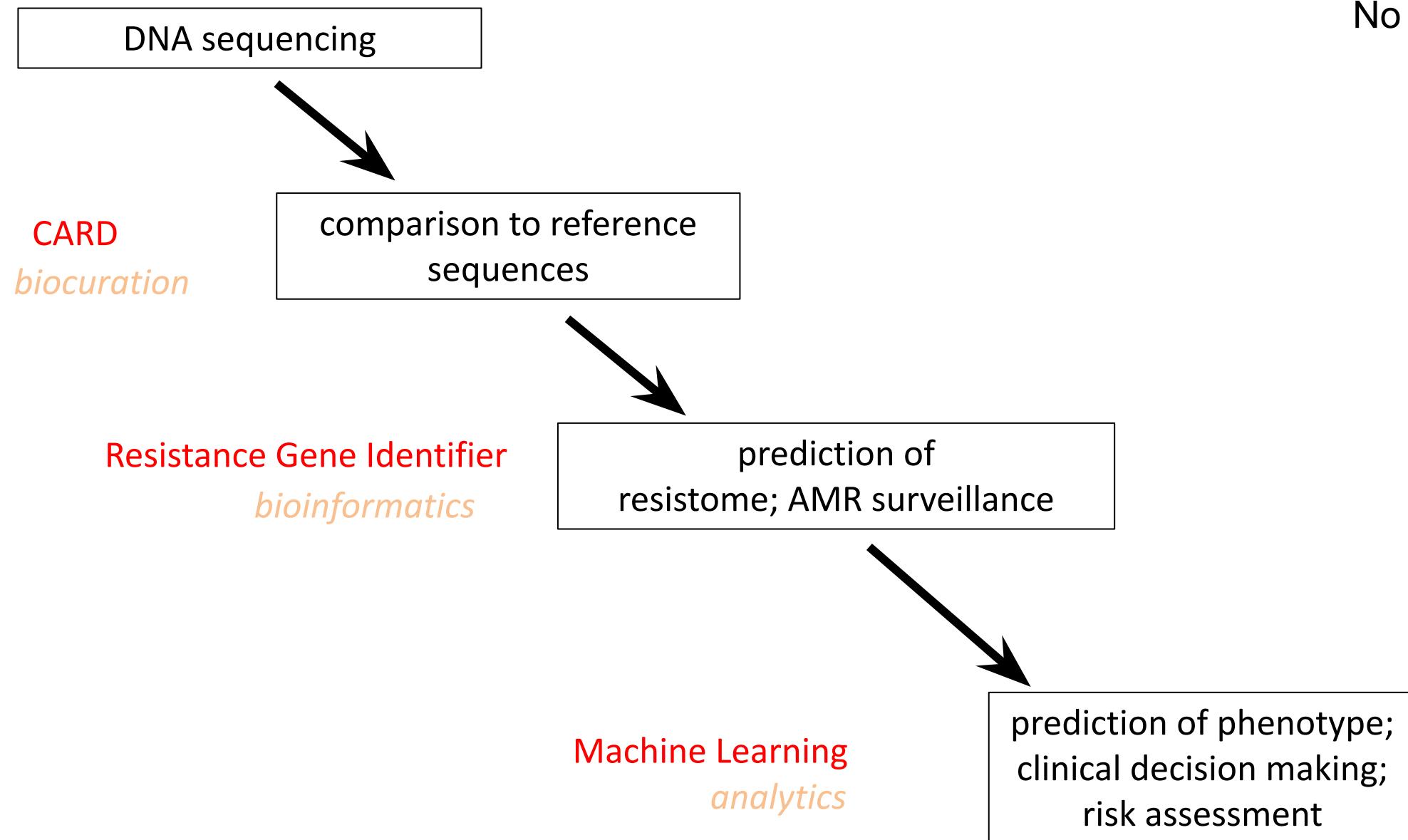
# DNA Sequencing – Resistance Genes Can No Longer Hide



# DNA Sequencing – Resistance Genes Can No Longer Hide



# DNA Sequencing – Resistance Genes Can No Longer Hide



# Mechanisms and Drivers of Resistance

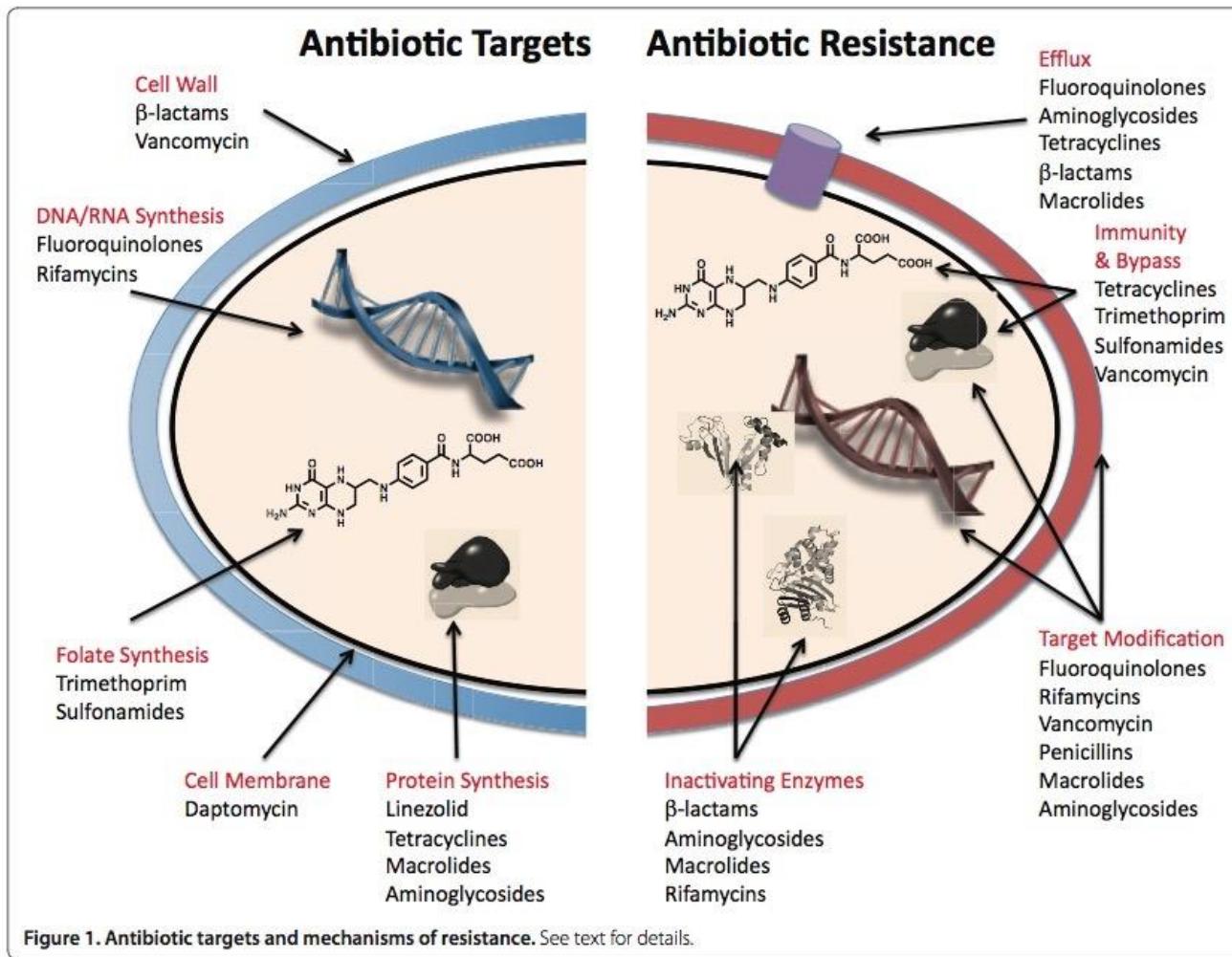
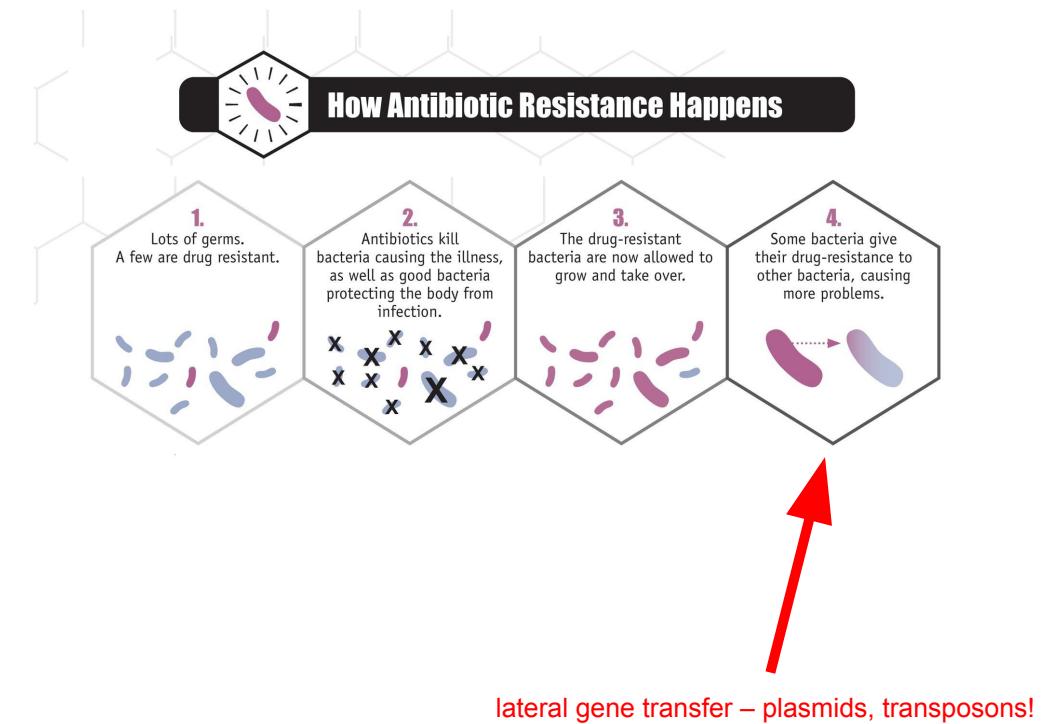


Figure 1. Antibiotic targets and mechanisms of resistance. See text for details.

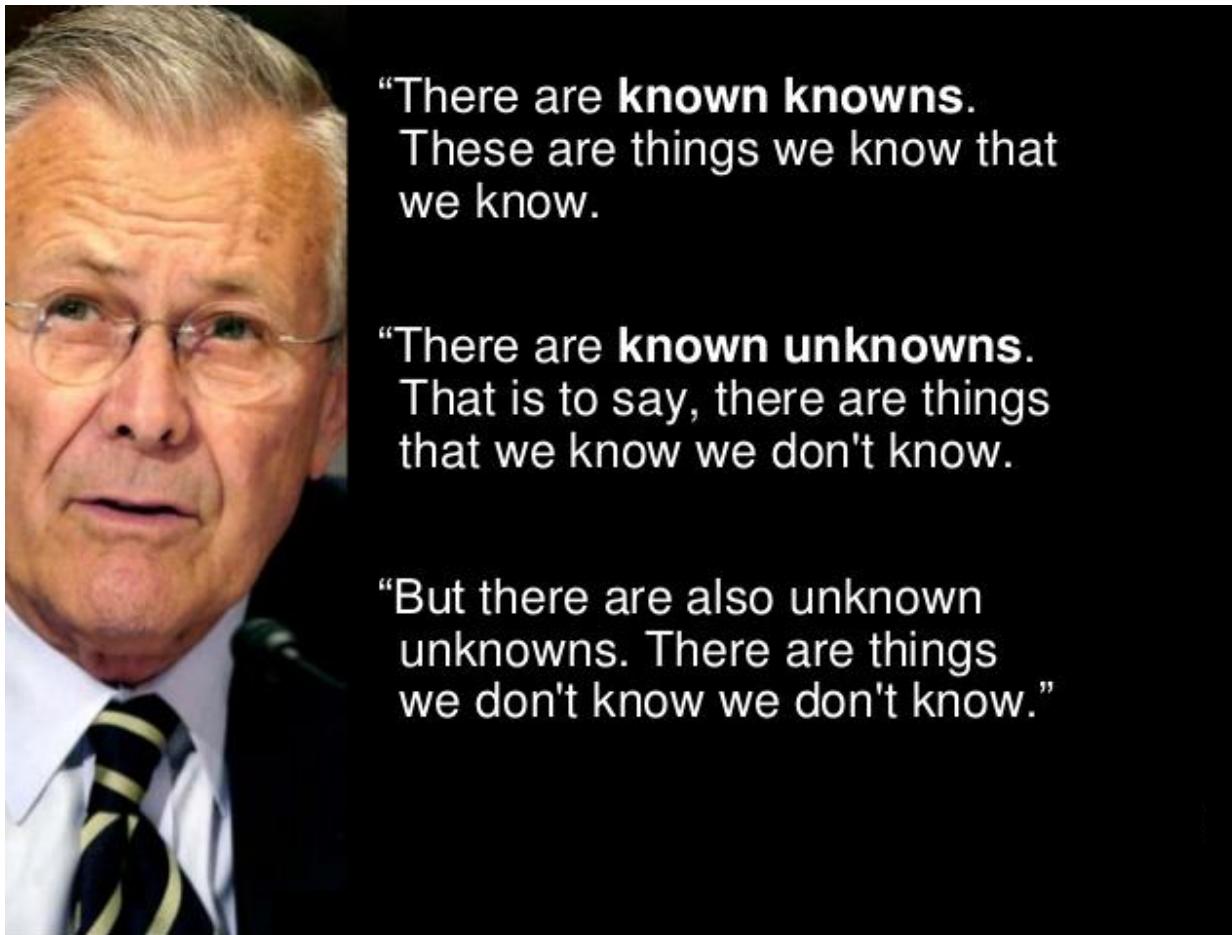
Wright, *BMC Biology* 2010 8:123



## NDM-1

[Download Sequences](#)

Accession	ARO:3000589
CARD Short Name	<i>NDM-1</i>
Definition	NDM-1 is a metallo-beta-lactamase isolated from <i>Klebsiella pneumoniae</i> with nearly complete resistance to all beta-lactam antibiotics.
AMR Gene Family	<a href="#">NDM beta-lactamase</a>
Drug Class	<a href="#">penam</a> , <a href="#">cephamycin</a> , <a href="#">carbapenem</a> , <a href="#">cephalosporin</a>
Resistance Mechanism	<a href="#">antibiotic inactivation</a>
Resistomes with Perfect Matches	Acinetobacter baumannii <sup>g+p+wgs</sup> , Acinetobacter defluvii <sup>g</sup> , Acinetobacter haemolyticus <sup>wgs</sup> , Acinetobacter indicus <sup>g+p</sup> , Acinetobacter johnsonii <sup>g+p+wgs</sup> , Acinetobacter junii <sup>g+p+wgs</sup> , Acinetobacter iwoffii <sup>wgs</sup> , Acinetobacter nosocomialis <sup>wgs</sup> , Acinetobacter pittii <sup>g+p+wgs</sup> , Acinetobacter radioresistens <sup>wgs</sup> , Acinetobacter townieri <sup>wgs</sup> , Acinetobacter wuhouensis <sup>g+wgs</sup> , Aeromonas caviae <sup>g</sup> , Bacillus subtilis <sup>wgs</sup> , Citrobacter amalonaticus <sup>wgs</sup> , Citrobacter freundii <sup>p+wgs</sup> , Citrobacter portucalensis <sup>p+wgs</sup> , Citrobacter werkmanii <sup>wgs</sup> , Citrobacter youngae <sup>wgs</sup> , Enterobacter asburiae <sup>p+wgs</sup> , Enterobacter cloacae <sup>g+p+wgs</sup> , Enterobacter hormaechei <sup>p+wgs</sup> , Enterobacter kobel <sup>p+wgs</sup> , Enterobacter rogenkampii <sup>wgs</sup> , Escherichia coli <sup>g+p+wgs</sup> , Klebsiella aerogenes <sup>p+wgs</sup> , Klebsiella michiganensis <sup>p+wgs</sup> , Klebsiella oxytoca <sup>wgs</sup> , Klebsiella pneumoniae <sup>g+p+wgs</sup> , Klebsiella quasipneumoniae <sup>p+wgs</sup> , Leclercia adecarboxylata <sup>wgs</sup> , Morganella morganii <sup>g+p+wgs</sup> , Proteus mirabilis <sup>g+p+wgs</sup> , Proteus vulgaris <sup>p</sup> , Providencia rettgeri <sup>g+p+wgs</sup> , Providencia stuartii <sup>g+p+wgs</sup> , Pseudomonas aeruginosa <sup>g+wgs</sup> , Pseudomonas putida <sup>wgs</sup> , Raoultella planticola <sup>p</sup> , Salmonella enterica <sup>p+wgs</sup> , Serratia marcescens <sup>g+p+wgs</sup> , Shewanella putrefaciens <sup>p+wgs</sup> , Shigella sonnei <sup>wgs</sup> , Vibrio alginolyticus <sup>wgs</sup> , Vibrio cholerae <sup>wgs</sup> , Vibrio parahaemolyticus <sup>wgs</sup> , Vibrio vulnificus <sup>wgs</sup>



"There are **known knowns**.  
These are things we know that  
we know.

"There are **known unknowns**.  
That is to say, there are things  
that we know we don't know.

"But there are also unknown  
unknowns. There are things  
we don't know we don't know."

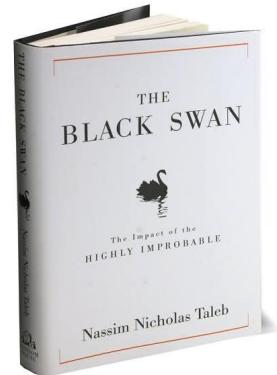


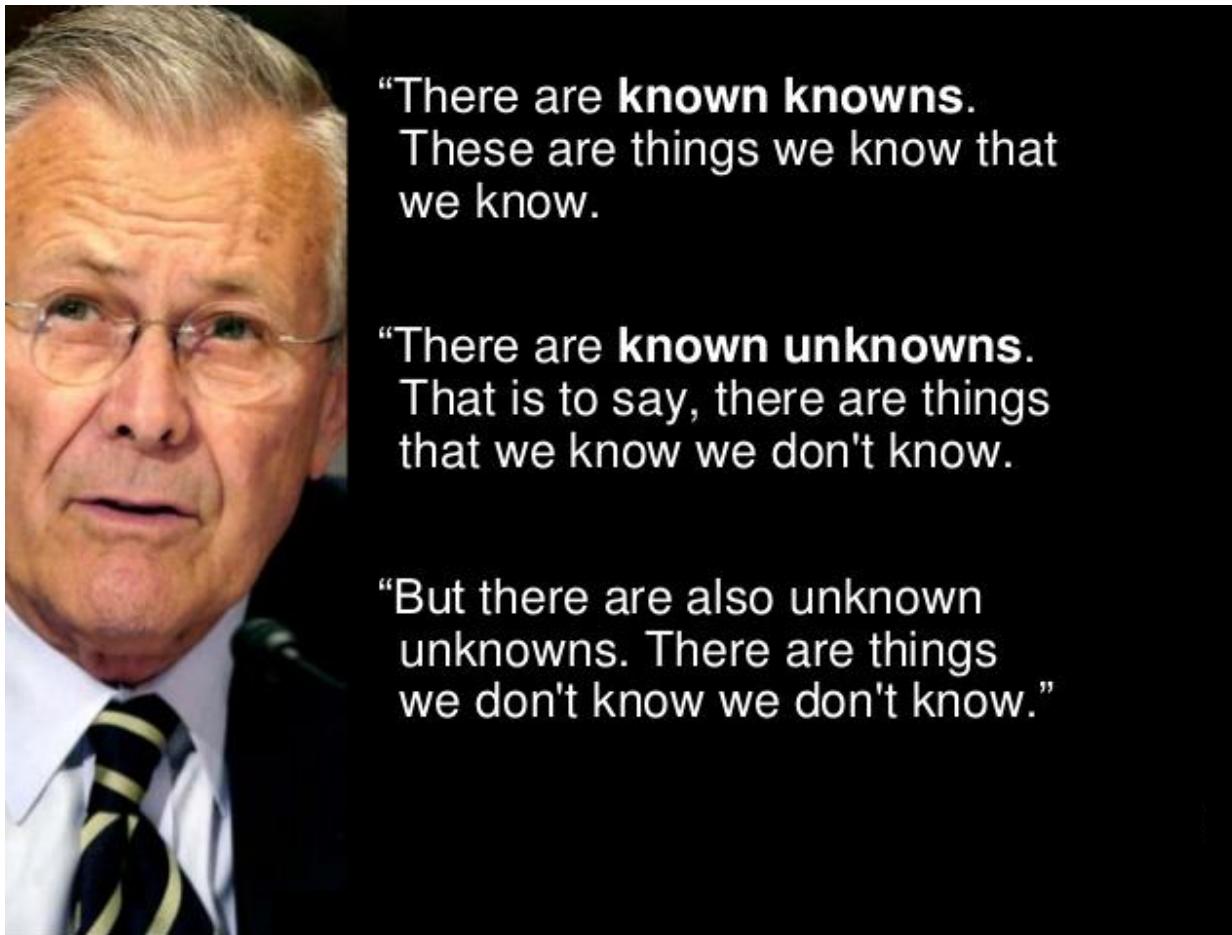
"There are **known knowns**.  
These are things we know that  
we know.

"There are **known unknowns**.  
That is to say, there are things  
that we know we don't know.

"But there are also unknown  
unknowns. There are things  
we don't know we don't know."

## Genes & Pathogens we are tracking





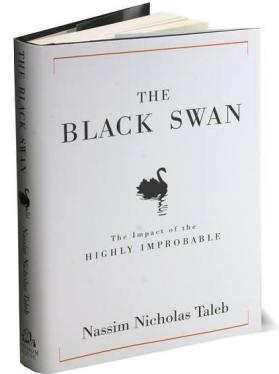
"There are **known knowns**.  
These are things we know that  
we know.

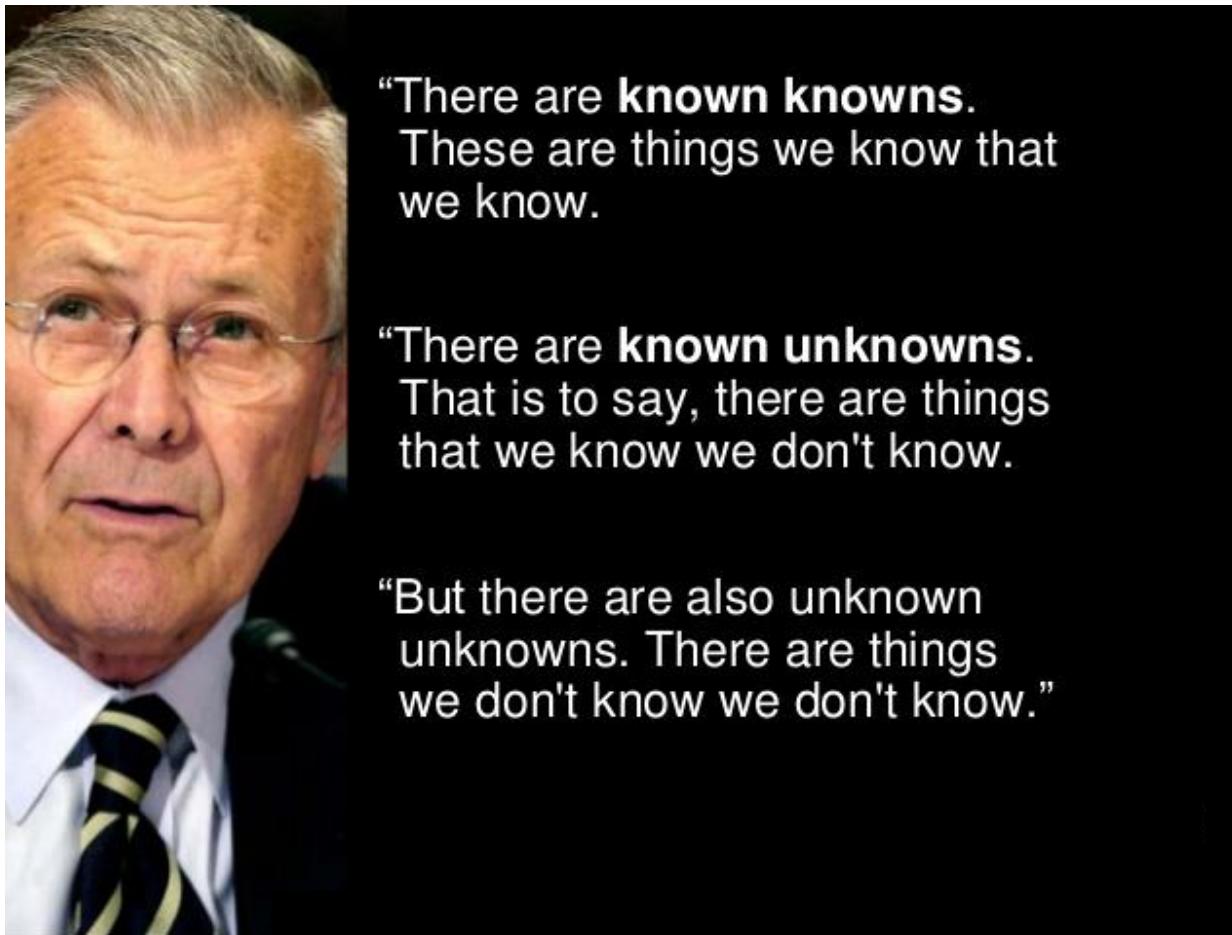
"There are **known unknowns**.  
That is to say, there are things  
that we know we don't know.

"But there are also unknown  
unknowns. There are things  
we don't know we don't know."

### Genes & Pathogens we are tracking

**Characterized genes  
we don't routinely  
track; variants of  
known genes; novel  
mutations**





"There are **known knowns**.  
These are things we know that  
we know.

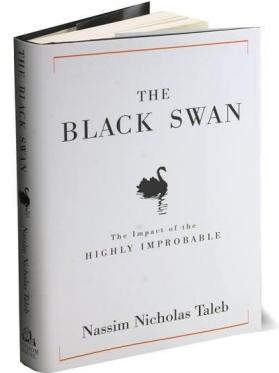
"There are **known unknowns**.  
That is to say, there are things  
that we know we don't know.

"But there are also unknown  
unknowns. There are things  
we don't know we don't know."

### Genes & Pathogens we are tracking

**Characterized genes  
we don't routinely  
track; variants of  
known genes; novel  
mutations**

### Emergent threats





"There are **known knowns**.  
These are things we know that  
we know.

"There are **known unknowns**.  
That is to say, there are things  
that we know we don't know.

"But there are also unknown  
unknowns. There are things  
we don't know we don't know."

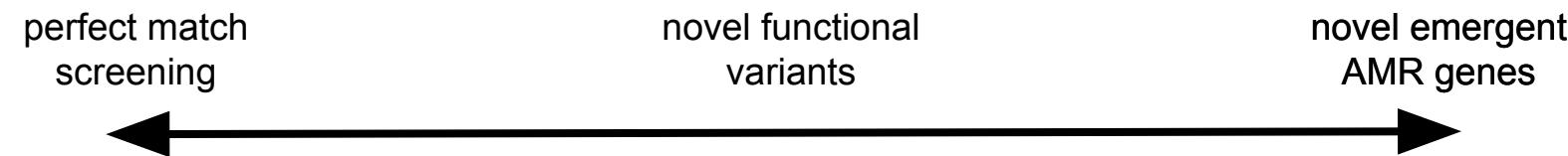


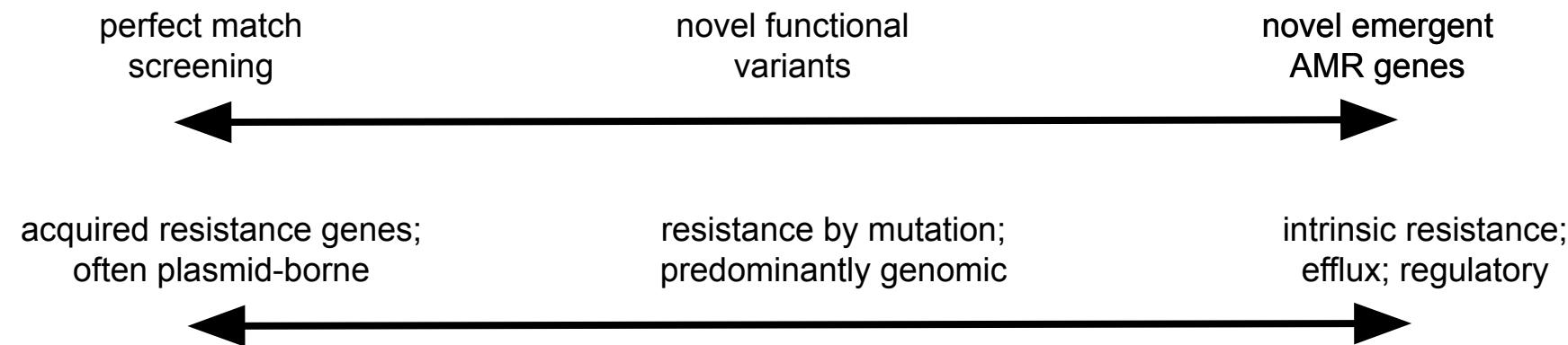
**Genes & Pathogens  
we are tracking**

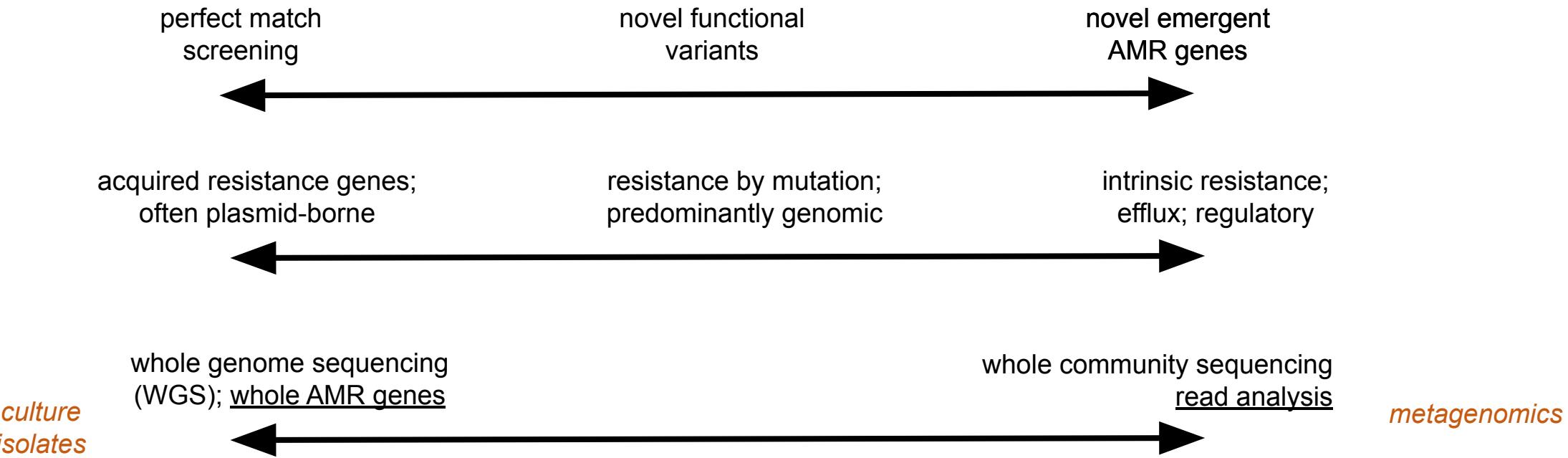
↓ *training gap; info overload*

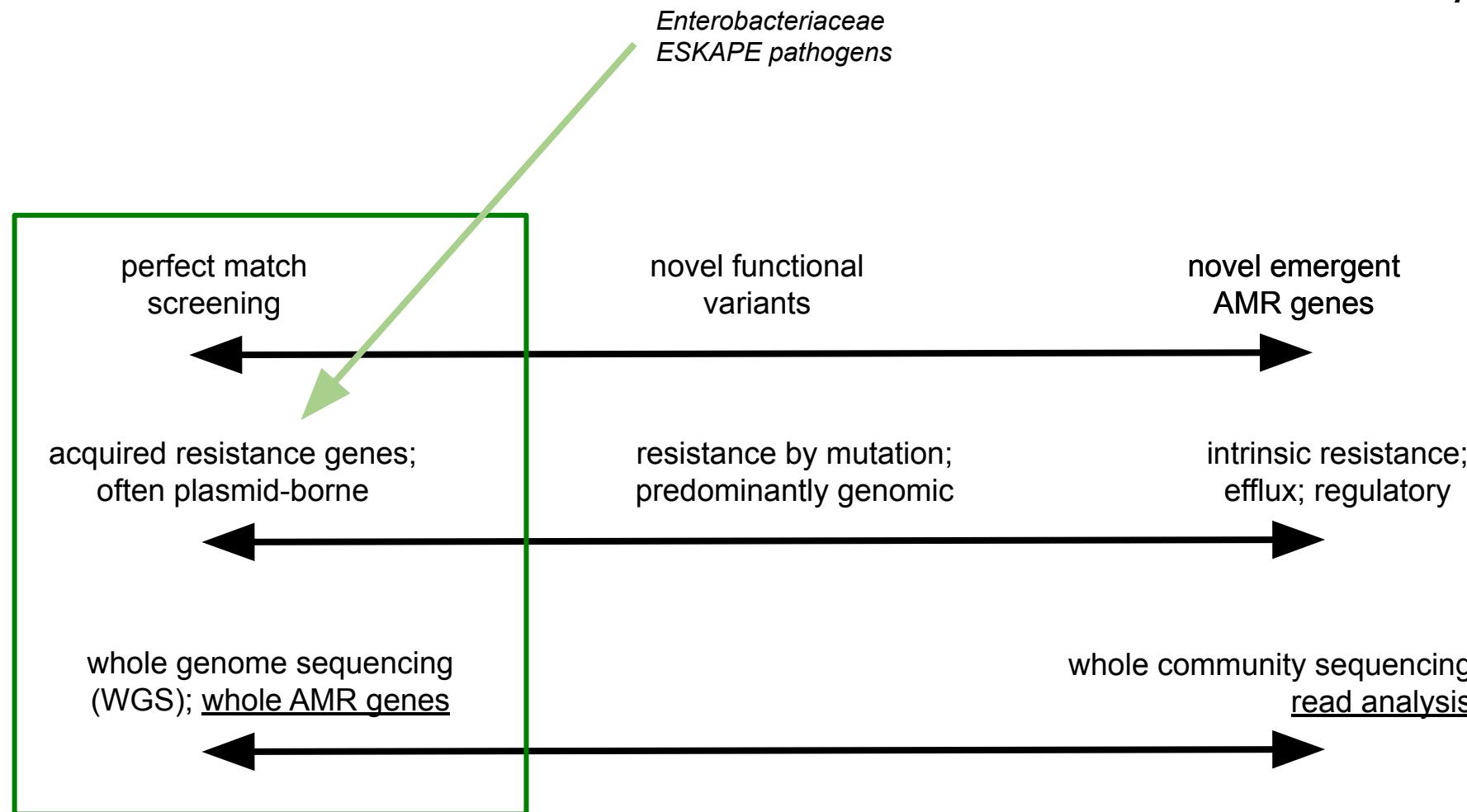
**Characterized genes  
we don't routinely  
track; variants of  
known genes; novel  
mutations**

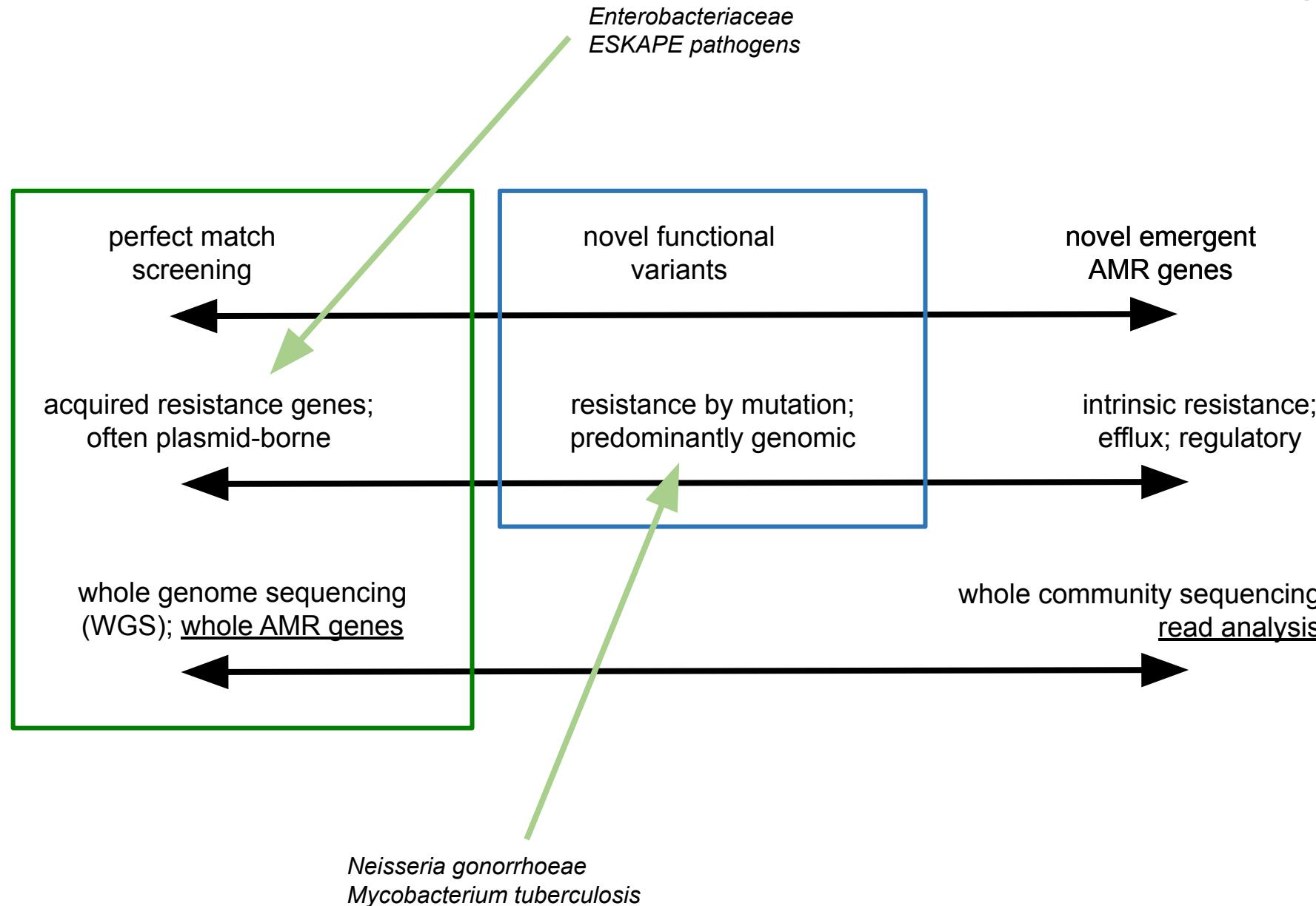
**Emergent threats**

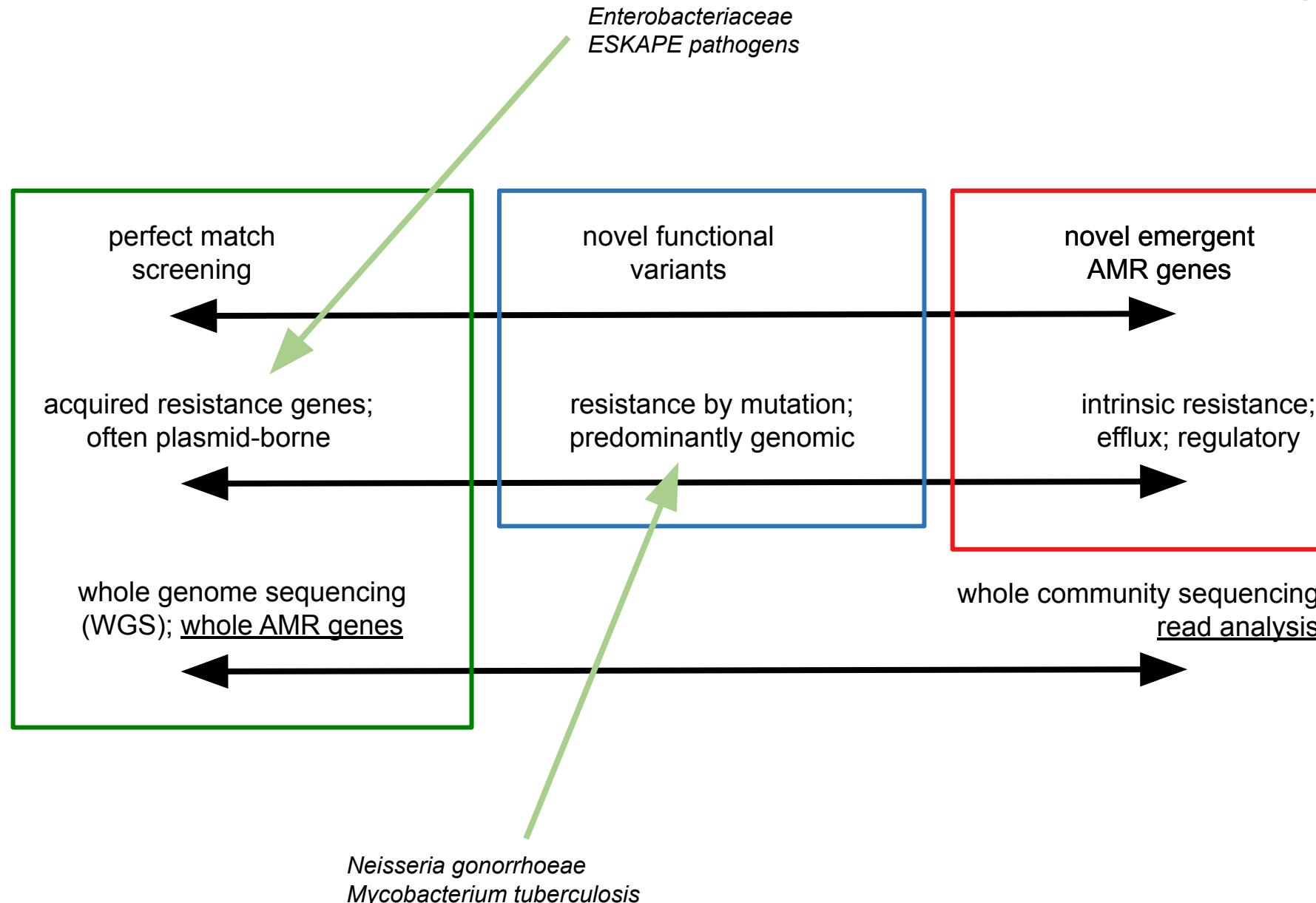


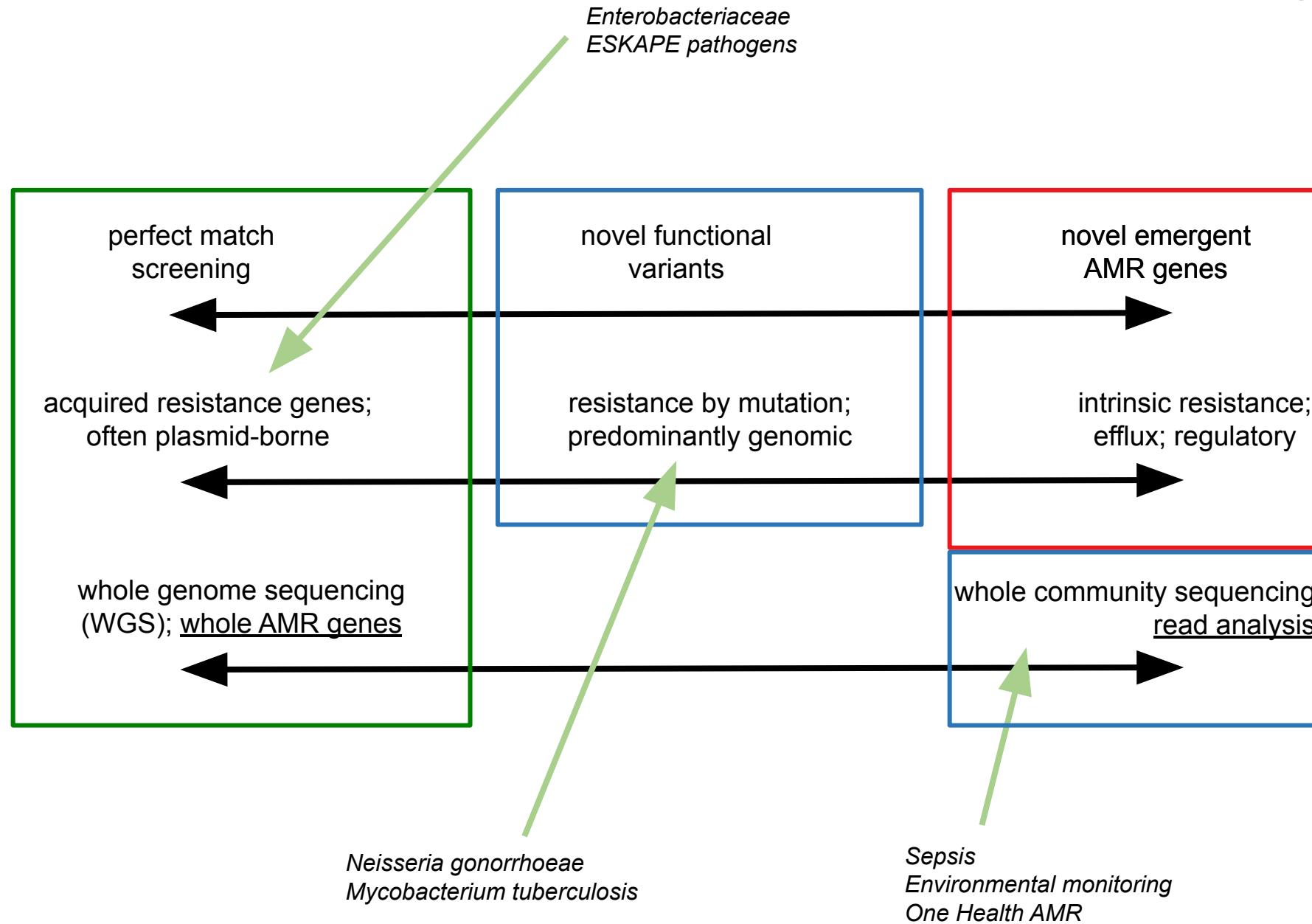












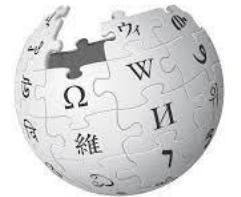
# Biocuration

**Databases on antimicrobial resistance rates and antibiotic consumption** [ edit ]

- CIPARS
- EARS-Net
- ESAC-Net

**Databases on antimicrobial resistance mechanisms** [ edit ]

- |                                      |  |  |
|--------------------------------------|--|--|
| • AMRFinderPlus                      | • The Comprehensive Antibiotic Resistance Database | • PathoPhenoDB                                       |
| • Antimicrobial Drug Database (AMDD) | • FARME  | • PATRIC database                                    |
| • ARDB (no longer maintained)        | • INTEGRALL  | • RAC: Repository of Antibiotic resistance Cassettes |
| • ARGminer                           | • LacED  | • ResFinder  |
| • BacMet                             | • MEGARes  | • TBDReaMDB  |
| • Beta-Lactamase Database (BLAD)     | • MUBII-TB-DB                                      | • u-CARE   |
| • Beta-Lactamase Database (BLDB)     | • Mustard Database                                 | • VFDB   |
| • CBMAR                              | • MvirDB   |  |



**WIKIPEDIA**  
The Free Encyclopedia

### The ‘Big Three’:

- NCBI's Pathogen Detection Reference Gene Catalog
- Resfinder
- Comprehensive Antibiotic Resistance Database

## The Comprehensive Antibiotic Resistance Database

A bioinformatic database of resistance genes, their products and associated phenotypes.

6657 Ontology Terms, 5031 Reference Sequences, 1931 SNPs, 3013 Publications, 5078 AMR Detection Models

Resistome predictions: 377 pathogens, 21079 chromosomes, 2662 genomic islands, 41828 plasmids, 155606 WGS assemblies, 322710 alleles

CARD Bait Capture Platform 1.0.0 | State of the CARD 2021 Presentations & Demonstrations

**Browse**

The CARD is a rigorously curated collection of characterized, peer-reviewed resistance determinants and associated antibiotics, organized by the Antibiotic Resistance Ontology (ARO) and AMR gene detection models.

**Analyze**

The CARD includes tools for analysis of molecular sequences, including BLAST and the Resistance Gene Identifier (RGI) software for prediction of resistome based on homology and SNP models.

**Download**

CARD data and ontologies can be downloaded in a number of formats. RGI software is available as a command-line tool. CARD Bait Capture Platform sequences and protocol available for download.

**Resistomes, Variants, & Prevalence**

Computer-generated resistome predictions for 377 important pathogens. Includes sequence variants beyond those reported in the scientific literature, as well as prevalence statistics for AMR genes among pathogens, genomes, and plasmids.

**CARD:Live**

The CARD:Live project collects pathogen identification, MLST, AMR gene list, date, and geographical region for genome sequences submitted to RGI online, providing a dynamic view of antibiotic resistant isolates being analyzed around the world.

**Timeline**

CARD Developers @arpcard · Oct 20

CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database

pubmed.ncbi.nlm.nih.gov

CARD 2023: expanded curation, support for ...

● 18  ⓘ

*CARD's primary curation paradigm:*

Peer reviewed publication in PubMed

Clear experimental evidence of elevated MIC (except beta-lactamases)

Public sequence record in GenBank



**The Comprehensive Antibiotic Resistance Database**

A bioinformatic database of resistance genes, their products and associated phenotypes.

6657 Ontology Terms, 5031 Reference Sequences, 1931 SNPs, 3013 Publications, 5078 AMR Detection Models

Resistome predictions: 377 pathogens, 21079 chromosomes, 2662 genomic islands, 41828 plasmids, 155606 WGS assemblies, 322710 alleles

CARD Bait Capture Platform 1.0.0 | State of the CARD 2021 Presentations & Demonstrations

**Browse**

The CARD is a rigorously curated collection of characterized, peer-reviewed resistance determinants and associated antibiotics, organized by the Antibiotic Resistance Ontology (ARO) and AMR gene detection models.

**Analyze**

The CARD includes tools for analysis of molecular sequences, including BLAST and the Resistance Gene Identifier (RGI) software for prediction of resistome based on homology and SNP models.

**Download**

CARD data and ontologies can be downloaded in a number of formats. RGI software is available as a command-line tool. CARD Bait Capture Platform sequences and protocol available for download.

**Resistomes, Variants, & Prevalence**

Computer-generated resistome predictions for 377 important pathogens. Includes sequence variants beyond those reported in the scientific literature, as well as prevalence statistics for AMR genes among pathogens, genomes, and plasmids.

**CARD:Live**

The CARD:Live project collects pathogen identification, MLST, AMR gene list, date, and geographical region for genome sequences submitted to RGI online, providing a dynamic view of antibiotic resistant isolates being analyzed around the world.

**Timeline**

CARD Developers @arpcard · Oct 20

CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database

pubmed.ncbi.nlm.nih.gov

CARD 2023: expanded curation, support for ...

● 18 ●

*CARD's primary curation paradigm:*

Peer reviewed publication in PubMed

Clear experimental evidence of elevated MIC (except beta-lactamases)

Public sequence record in GenBank



## ANT(6) [AMR Gene Family]

[Download Sequences](#)

Accession	ARO:3000225
Synonym(s)	<i>aadE</i>
Definition	Nucleotidylylation of streptomycin at the hydroxyl group at position 6
Drug Class	<a href="#">aminoglycoside antibiotic</a>
Resistance Mechanism	<a href="#">antibiotic inactivation</a>
Classification	<a href="#">9 ontology terms</a>   <a href="#">Hide</a> + <a href="#">process or component of antibiotic biology or chemistry</a> + <a href="#">mechanism of antibiotic resistance</a> + <a href="#">antibiotic molecule</a> + <a href="#">determinant of antibiotic resistance</a> + <a href="#">aminoglycoside antibiotic</a> [Drug Class] + <a href="#">antibiotic inactivation</a> [Resistance Mechanism] + <a href="#">nucleotidylation of antibiotic conferring resistance</a> + <a href="#">antibiotic inactivation enzyme</a> + <a href="#">determinant of aminoglycoside resistance</a>
Parent Term(s)	<a href="#">2 ontology terms</a>   <a href="#">Hide</a> + <a href="#">confers_resistance_to aminoglycoside antibiotic</a> [Drug Class] + <a href="#">aminoglycoside nucleotidyltransferase (ANT)</a>
Sub-Term(s)	<a href="#">4 ontology terms</a>   <a href="#">Hide</a> + <a href="#">ANT(6)-Ia</a> + <a href="#">aadK</a> + <a href="#">aad(6)</a> + <a href="#">ANT(6)-Ib</a>

Permanent accession, nomenclature & synonyms, definition



Brian Alcock, Lead CARD Curator

### Antibiotic Resistance Ontology knowledgebase

The screenshot displays two web pages side-by-side. On the left is the "GEMONIC EPIDEMIOLOGY ONTOLOGY" page, which features a banner for "Market Day", navigation links like "HOME", "PUB", "CONSORTIUM", "CATALOGUE", "BLOG", and "GET INVOLVED". Below the banner is a section titled "RECENT NEWS" with a link to "Get involved with our continued development of an open source knowledgebase for antibiotic resistance and resistance investigation". On the right is the "IRIDA" page, which has a banner for "The Integrated Rapid Infectious Disease Analysis (IRIDA) project is a Canadian-led initiative to build an open source, end-to-end platform for public health genomics". It includes a "HIGHLIGHTS" section with three items: "Data Management", "Analysis Pipeline", and "Visualizations as a Tool".

**ANT(6) [AMR Gene Family]**

[Download Sequences](#)

Accession	ARO:3000225
Synonym(s)	<a href="#">aadE</a>
Definition	Nucleotidylation of streptomycin at the hydroxyl group at position 6
Drug Class	<a href="#">aminoglycoside antibiotic</a>
Resistance Mechanism	<a href="#">antibiotic inactivation</a>
Classification	<b>9 ontology terms</b>   <a href="#">Hide</a> + <a href="#">process or component of antibiotic biology or chemistry</a> + <a href="#">mechanism of antibiotic resistance</a> + <a href="#">antibiotic molecule</a> + <a href="#">determinant of antibiotic resistance</a> + <a href="#">aminoglycoside antibiotic</a> [Drug Class] + <a href="#">antibiotic inactivation</a> [Resistance Mechanism] + <a href="#">nucleotidylation of antibiotic conferring resistance</a> + <a href="#">antibiotic inactivation enzyme</a> + <a href="#">determinant of aminoglycoside resistance</a>
Parent Term(s)	<b>2 ontology terms</b>   <a href="#">Hide</a> + <a href="#">confers_resistance_to aminoglycoside antibiotic</a> [Drug Class] + <a href="#">aminoglycoside nucleotidyltransferase (ANT)</a>
Sub-Term(s)	<b>4 ontology terms</b>   <a href="#">Hide</a> + <a href="#">ANT(6)-Ia</a> + <a href="#">aadK</a> + <a href="#">aad(6)</a> + <a href="#">ANT(6)-Ib</a>

**Classification:**

**AMR Gene Family (458)**  
**Resistance Mechanism (8)**  
**Drug Class (64)**

**Highly Curated:**

**AMR Gene Family -> Drug Class (458)**

**Less Curated:**

**AMR Gene/Mutation -> Antibiotic (3000+)**

[Protein](#)    [DNA](#)

```
>gb|CBH51824.1|+|ANT(6)-Ib [Campylobacter fetus subsp. fetus]
MKMRTEKQIYDTILNFAKADDRIRVVTLEGSRTNNINIIPDDFQDYDITFFVTDMQSFINSD
EVLNVGERLIMQKPEDMELFPKEEKGYS
YLMLFWDGVKIDLTLPLEVLDEYFTWDKLVKLLLLDKDNRVTNIPVPTDEDYYIEHPTARSF
DDCCNEFWNTVTYVVKGCRKEILFAID
HLNNIVRMELLRMISWKVGIEQGQYSFLGKNYKFLERYISPTELWKKILATYNMGSYTEMWKS
LELCMGIFRMSKEVAQCLNYLYPDYDK
NISNYVIRQKEKYQR
```

## ANT(6)-Ia

[Download Sequences](#)

Accession	ARO:3002626
Synonym(s)	<i>ant6, aadE</i>
Definition	ANT(6)-Ia is an aminoglycoside nucleotidyltransferase gene encoded by plasmids and chromosomes in <i>Staphylococcus epidermidis</i> , <i>E. faecium</i> , <i>Streptococcus suis</i> , <i>S. aureus</i> , <i>E. faecalis</i> and <i>Streptococcus mitis</i>
AMR Gene Family	<a href="#">ANT(6)</a>
Drug Class	<a href="#">aminoglycoside antibiotic</a>
Resistance Mechanism	<a href="#">antibiotic inactivation</a>
Resistomes with Perfect Matches	<i>Clostridioides difficile</i> <sup>g+wgs</sup> , <i>Clostridium perfringens</i> <sup>wgs</sup> , <i>Enterococcus casseliflavus</i> <sup>p</sup> , <i>Enterococcus faecalis</i> <sup>p+wgs</sup> , <i>Enterococcus faecium</i> <sup>g+p+wgs+qi</sup> , <i>Enterococcus hirae</i> <sup>g+wgs+qi</sup> , <i>Erysipelothrix rhusiopathiae</i> <sup>g+qi</sup> , <i>Herbinix luporum</i> <sup>g±qi</sup> , <i>Jeotgalibaca arthritidis</i> <sup>g</sup> , <i>Lactobacillus animalis</i> <sup>wgs</sup> , <i>Listeria innocua</i> <sup>p+wgs</sup> , <i>Staphylococcus aureus</i> <sup>g+p+wgs+qi</sup> , <i>Staphylococcus haemolyticus</i> <sup>g+wgs</sup> , <i>Staphylococcus saprophyticus</i> <sup>p+wgs</sup> , <i>Streptococcus agalactiae</i> <sup>g+wgs</sup> , <i>Streptococcus pasteurianus</i> <sup>wgs</sup> , <i>Streptococcus pneumoniae</i> <sup>wgs</sup> , <i>Streptococcus pyogenes</i> <sup>wgs</sup> , <i>Streptococcus suis</i> <sup>g+wgs+qi</sup>
Resistomes with Sequence Variants	<i>Bacteroides ovatus</i> <sup>wgs</sup> , <i>Campylobacter jejuni</i> <sup>g+p+wgs+qi</sup> , <i>Clostridioides difficile</i> <sup>g+wgs</sup> , <i>Clostridium perfringens</i> <sup>wgs</sup> , <i>Enterobacter hormaechei</i> <sup>wgs</sup> , <i>Enterococcus avium</i> <sup>p</sup> , <i>Enterococcus casseliflavus</i> <sup>p</sup> , <i>Enterococcus faecalis</i> <sup>g+p+wgs+qi</sup> , <i>Enterococcus faecium</i> <sup>g+p+wgs+qi</sup> , <i>Enterococcus hirae</i> <sup>g+p+wgs+qi</sup> , <i>Erysipelothrix rhusiopathiae</i> <sup>g+qi</sup> , <i>Escherichia coli</i> <sup>wgs</sup> , <i>Herbinix luporum</i> <sup>g±qi</sup> , <i>Jeotgalibaca arthritidis</i> <sup>g</sup> , <i>Klebsiella pneumoniae</i> <sup>wgs</sup> , <i>Lactobacillus animalis</i> <sup>wgs</sup> , <i>Listeria innocua</i> <sup>p+wgs</sup> , <i>Staphylococcus aureus</i> <sup>g+p+wgs+qi</sup> , <i>Staphylococcus haemolyticus</i> <sup>g+wgs</sup> , <i>Staphylococcus pseudintermedius</i> <sup>g+wgs</sup> , <i>Staphylococcus saprophyticus</i> <sup>p+wgs</sup> , <i>Streptococcus agalactiae</i> <sup>g+wgs</sup> , <i>Streptococcus pasteurianus</i> <sup>wgs</sup> , <i>Streptococcus pneumoniae</i> <sup>wgs</sup> , <i>Streptococcus pyogenes</i> <sup>wgs</sup> , <i>Streptococcus suis</i> <sup>g+wgs+qi</sup>
Classification	<a href="#">9 ontology terms</a>   <a href="#">Show</a>
Parent Term(s)	<a href="#">2 ontology terms</a>   <a href="#">Show</a>
Publications	Pinto-Alphandary H, et al. 1990. Antimicrob Agents Chemother 34(6): 1294-1296. Emergence of aminoglycoside resistance genes <i>aadA</i> and <i>aadE</i> in the genus <i>Campylobacter</i> . ( <a href="#">PMID 2168151</a> ) Gill SR, et al. 2005. J Bacteriol 187(7): 2426-2438. Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant <i>Staphylococcus aureus</i> strain and a biofilm-producing methicillin-resistant <i>Staphylococcus epidermidis</i> strain. ( <a href="#">PMID 15774886</a> )

*in silico* prediction of AMR genes,  
resistomes, and sequence variants

377 pathogens, 21079 chromosomes, 2662 genomic islands,  
41828 plasmids, 155606 WGS assemblies, 322710 alleles

The Comprehensive Antibiotic Resistance Database  
A bioinformatic database of resistance genes, their products and associated phenotypes.  
6657 Ontology Terms, 5031 Reference Sequences, 1931 SNPs, 3013 Publications, 5078 AMR Detection Models  
Resistome predictions: 377 pathogens, 21079 chromosomes, 2662 genomic islands, 41828 plasmids, 155606 WGS assemblies, 322710 alleles  
CARD Bait Capture Platform 1.0.0 | State of the CARD 2023 Presentations & Demonstrations

**BROWSE**  
The CARD is a rigorously curated collection of characterized, peer-reviewed resistance determinants and associated antibiotics, organized by the Antibiotic Resistance Ontology (ARO) and AMR gene detection models.

**Analyze**  
The CARD includes tools for analysis of molecular sequences, including BLAST and the Resistance Gene Identifier (RGI) software for prediction of resistome based on homology and SNP models.

**Resistomes, Variants, & Prevalence**  
Computer-generated resistome predictions for 377 important pathogens. Includes sequence variants beyond those reported in the scientific literature, as well as prevalence viewing of AMR genes among pathogens, genomes, and plasmids.

**CARD:Live**  
The CARD:Live project collects pathogen identification, MLST, AMR gene list, date, and geographical region for genome sequences submitted to RGI online, providing a dynamic view of antibiotic resistant isolates being analyzed around the world.

**Download**  
CARD data and ontologies can be downloaded in a number of formats. RGI software is available as a command-line tool. CARD Bait Capture Platform sequences and protocol available for download.

**Timeline**

- CARD Developers @ercard - Oct 20 CARD 2022: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database
- PubMed.ncbi.nlm.nih.gov CARD 2023: expanded curation, support for ...

- High quality reference database on the molecular basis of AMR
- Expert curation, guided by CARD\*Shark PubMED text mining algorithms
- Breadth of data & AMR mechanisms covered
- Advanced analytics & AMR gene discovery
- Tools for data harmonization, standardized names for machine learning
- Antibiotic Resistance Ontology
- Dedicated curators – updates every 1-3 months

# Bioinformatics

## Infectious Disease Combined Rounds



**"A bad penny always comes back: harnessing immunity to combat *Salmonella* superbugs"**

### Dr. Eva Piessens & Dr. Brian Coombes

ASSOCIATE PROFESSOR  
MEDICINE

PROFESSOR AND CHAIR  
BIOCHEMISTRY &  
BIOMEDICAL SCIENCES

In this seminar, Dr. Coombes and Dr. Piessens will be discussing non-typhoidal *Salmonella* and the emergence of highly virulent, multi-drug resistant strains both here and abroad.

They will present a clinical case study and discuss the latest research in trying to overcome drug resistance by harnessing the power of the innate immune system.

**Wednesday, February 5th**  
**8 - 9 a.m. | MDCL 1309**

ID/IIDR Combined Rounds occur on the first Wednesday of every month.

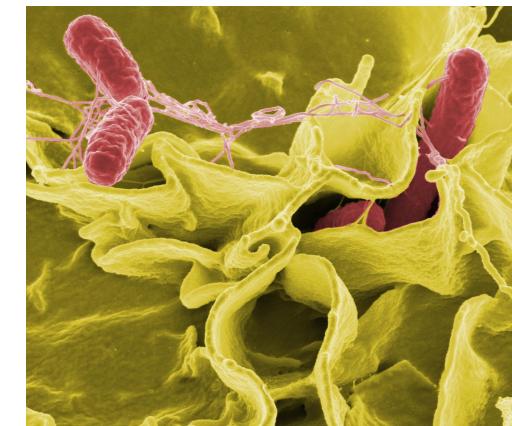
### CLINICAL CASE REPORT

#### Recurrent multidrug-resistant *Salmonella enterica* serovar Typhimurium bacteremia in a returned traveller

Adam S Komorowski MD<sup>1</sup> , Xena X Li MD, FRCPC<sup>1,2</sup>, Eva Piessens MD, MPH, ABIM<sup>2</sup>, Andrew G McArthur PhD<sup>3</sup>, Ameen Patel MD, FRCPC<sup>4</sup>

This case report describes a 68-year-old male with recurrent multidrug-resistant *Salmonella enterica* serovar Typhimurium bacteremia acquired during travel abroad. He experienced a recurrence of bacteremia without a clear source and was successfully treated with 10 weeks of intravenous ertapenem. *Post hoc* genome sequencing revealed an isolate bearing class A, C, and D extended-spectrum  $\beta$ -lactamases (ESBLs). A review of English- and French-language literature since 2000 revealed eight publications that discussed recurrent *S. enterica* serovar Typhimurium bacteremia. Patients with multidrug-resistant *S. enterica* serovar Typhimurium should be monitored frequently for recrudescence, even in the absence of risk factors.

**KEY WORDS:** bacteremia, multidrug-resistant *Salmonella*, *Salmonella*, *Salmonella enterica*, *Salmonella* infections, *Salmonella Typhimurium*



Infectious Disease  
Combined Rounds



**"A bad penny always comes back: harnessing immunity to combat *Salmonella* superbugs"**

**Dr. Eva Piessens & Dr. Brian Coombes**

ASSOCIATE PROFESSOR  
MEDICINE

PROFESSOR AND CHAIR  
BIOCHEMISTRY &  
BIOMOLECULAR SCIENCES

In this seminar, Dr. Coombes and Dr. Piessens will be discussing non-typhoidal *Salmonella* and the emergence of highly virulent, multi-drug resistant strains both here and abroad.

They will present a clinical case study and discuss the latest research in trying to overcome drug resistance by harnessing the power of the innate immune system.

**Wednesday, February 5th**  
8 - 9 a.m. | MDCL 1309

ID/IIDR Combined Rounds occur on the first Wednesday of every month.

>Salmonella genome sequence  
 CTGTCCCTTGGCTTTTTACTATGGTCTCGAAATGAACGTTGTATGACTTGCCATCC  
 GCTTGACGCTCTTCATAGTGATTGGCTGTGTAATCGCGGATGCTCGCACGCCAAAAGCCG  
 ATGGCAGGCATGTTGTTCCAATTGAAAGACTTCCCAGGCACCCGGGAACATTTGAAG  
 ATTTGATGGGAGACCCTTCCGATTCCCAGCCGACGGTAGCGTTCATGATAAAAAAT  
 TCAGCAATGGAGTGGATCGATGGCAAATTGCATCAGGGTTGTGAGCATGCGGACGAGG  
 ACAAAACCAGCCAGCTTACCATTAACCTTGATGAAAAAGGGATGGCGCTCTGTTCAACC  
 CAATATTCTCAAAGTAGTCGTAGCCAAACAAGCCGTCGTACATCTACTTCGAAAGGCTCA  
 AACTCAGTGAAGTCGTACGTATACAGCTCAACAGTTGCTTGAGAGCGTGCTCTGTTCA  
 AAAGGAAACTCGTTCTATCGAGAATGCTTGATTCACTGAGAAAACCTCCTTTCTTGAG  
 ATATAGGAGTATCGTACTTTATTTAGGTTGGGAACATATGGTGGACACTTCTTTA  
 TAGAACCTCACACGATTGATCATTACTTGAGGTGGTCCAGTCGGGCTATACTAGAGT  
 CAAACGAAAGAAAATCAATACGAAGCAAGACAGCAGCCATACGGTGATCAGCCGGTGGC  
 TGAATTGTTATTCTTATCTCAATAACAGCCTCGAGAATGTAACAACACTGGGAGACGAACGA  
 CAAATTAATCAACGGAAACATCACATAATAGGATGGCGATGTAAAGTGAAGGTACTTGGT  
 CATTTCCTTTCTAAAGATGACATGGAGGAAATATCCGTATTAAGATCTATGTATTGGG

## RGI Resistance Gene Identifier

RGI can be used to predict resistomes from protein or nucleotide data based on homology and SNP models. Analyses can be performed via this web portal (20 Mb limit), via the command line, or via use of a [Galaxy wrapper](#). The command line version can be obtained from the [Download section of the CARD website](#). You can additionally install RGI from Conda or run RGI from Docker.

This web portal supports analysis of genomes, genome assemblies, metagenomic contigs, or proteomes. The command line tool additionally supports analysis of metagenomic reads and k-mer prediction of pathogen-of-origin for AMR genes.

**Web portal - RGI 5.1.0, CARD 3.0.7:** Open Reading Frame (ORF) prediction using [Prodigal](#), homolog detection using [DIAMOND](#), and Strict significance based on CARD curated bitscore cut-offs. Options included for percent identity filtering, low quality/coverage assemblies, merged metagenomic reads, small plasmids or assembly contigs (<20,000 bp).

Online RGI results cached for 7 days. As the CARD curation evolves, the results of the RGI evolve. RGI targets, reference sequences, and significance cut-offs are under constant curation. Full documentation for the RGI can be found at [GitHub](#).

Use RGI:

**Enter a GenBank accession(s):**

Enter accessions separated by commas

Nucleotide sequences will undergo ORF calling to generate predicted protein sequences. Examples:

**Select Data Type:**

DNA sequence

Protein sequence



## Infectious Disease Combined Rounds

## **"A bad penny always comes back: harnessing immunity to combat *Salmonella* superbugs"**

**Dr. Eva Piessens & Dr. Brian Coombes**

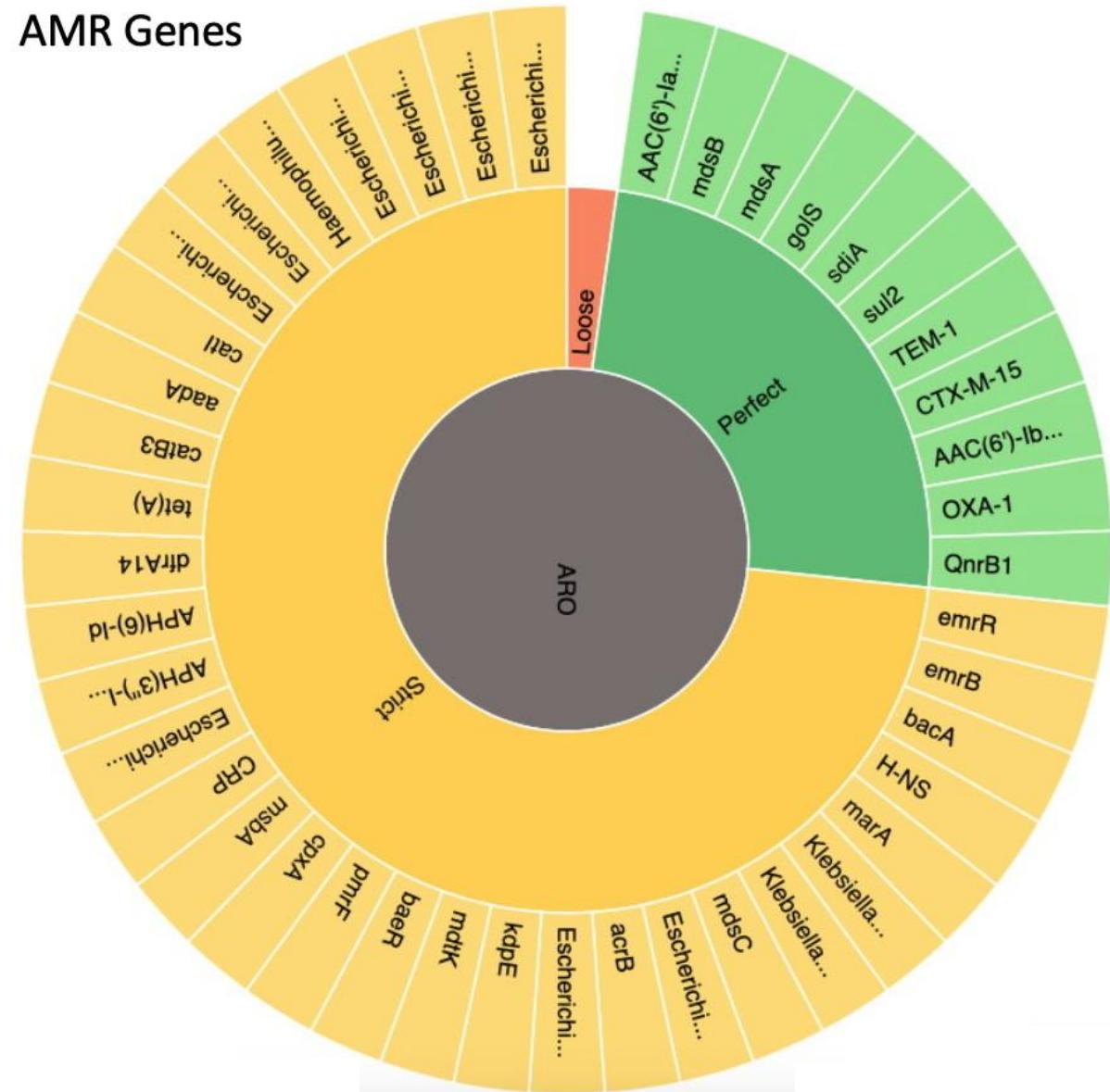
**ASSOCIATE PROFESSOR**  
MEDICINE

**PROFESSOR AND CHAIR**  
BIOCHEMISTRY &  
BIOMEDICAL SCIENCES

In this seminar, Dr. Coombes and Dr. Piessens will be discussing non-typhoidal *Salmonella* and the emergence of highly virulent, multi-drug resistant strains both here and abroad. They will present a clinical case study and discuss the latest research in trying to overcome drug resistance by harnessing the power of the innate immune system.

Wednesday, February 5th  
8 - 9 a.m. | MDCL 1309

ID/IIDR Combined Rounds occur on the first Wednesday of every month.





"A bad penny always comes back: harnessing immunity to combat *Salmonella* superbugs"

Dr. Eva Piessens & Dr. Brian Coombes

ASSOCIATE PROFESSOR  
MEDICINE

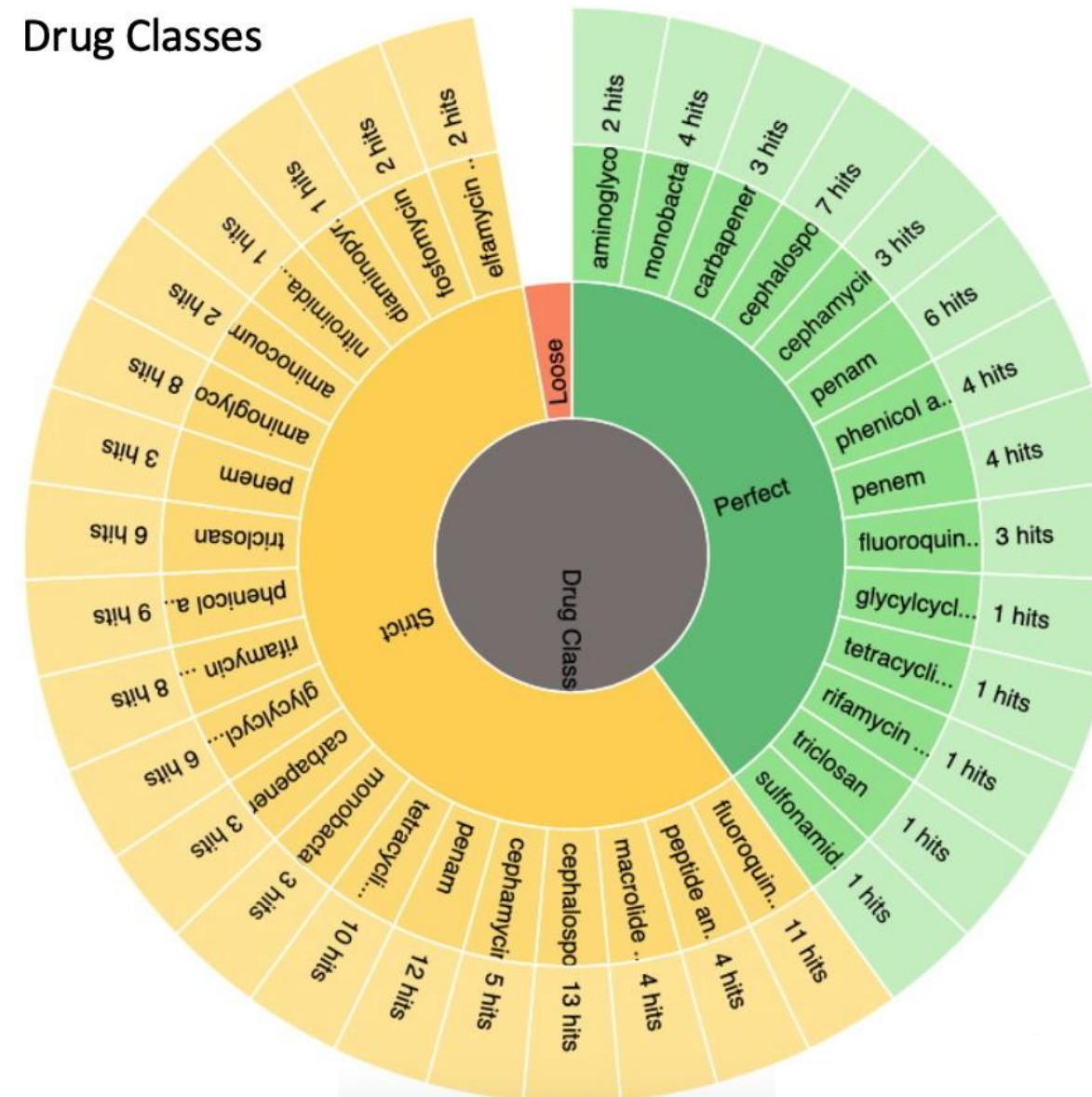
PROFESSOR AND CHAIR  
BIOCHEMISTRY &  
BIOMEDICAL SCIENCES

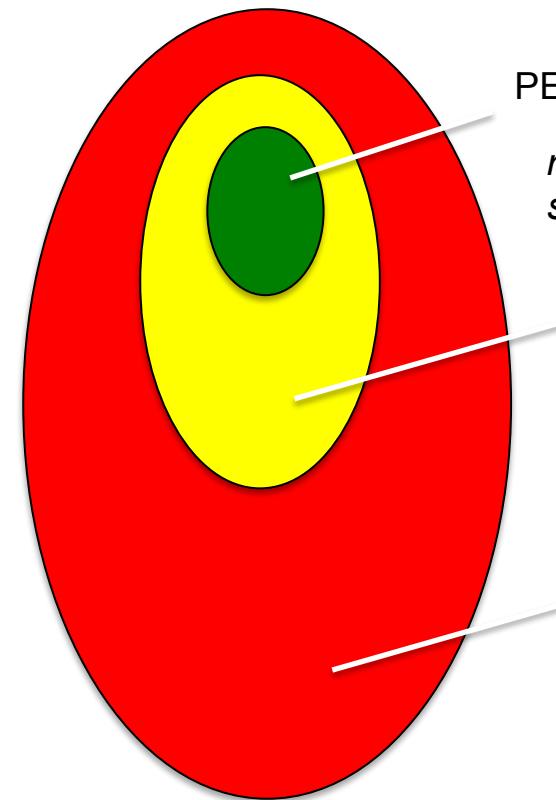
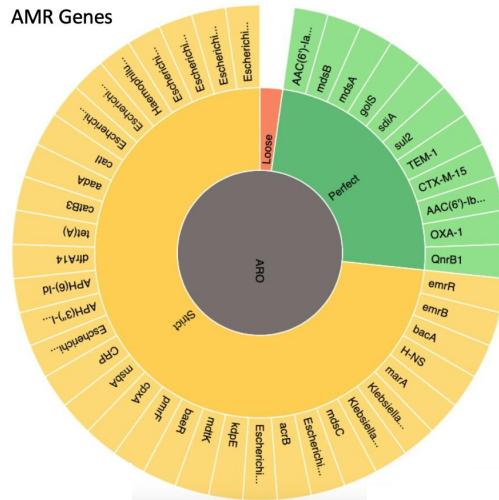
In this seminar, Dr. Coombes and Dr. Piessens will be discussing non-typhoidal *Salmonella* and the emergence of highly virulent, multi-drug resistant strains both here and abroad. They will present a clinical case study and discuss the latest research in trying to overcome drug resistance by harnessing the power of the innate immune system.

Wednesday, February 5th  
8 - 9 a.m. | MDCL 1309

ID/IIDR Combined Rounds occur on the first Wednesday of every month.

## Drug Classes





## PERFECT

*matches reference sequence*

## STRICT

*similarity within “known unknown”  
model; functional  
variant*

LOOSE

*similarity outside  
of model; needs  
validation*

## ANT(6)-Ia

protein homolog model

**Model Type:** protein homolog model

**Model Definition:** Protein Homolog Models (PHM) detect protein sequences based on their similarity to a curated reference sequence, using curated BLASTP bitscore cut-offs. Protein Homolog Models apply to all genes that confer resistance through their presence in an organism, such as the presence of a beta-lactamase gene on a plasmid. PHMs include a reference sequence and a bitscore cut-off for detection using BLASTP. A Perfect RGI match is 100% identical to the reference protein sequence along its entire length, a Strict RGI match is not identical but the bit-score of the matched sequence is greater than the curated BLASTP bit-score cutoff, Loose RGI matches have a bit-score less than the curated BLASTP bit-score cut-off.

**Bit-score Cut-off (blastP): 500**

Protein    DNA

```
>gb|AHE40557.1| - |ANT(6)-Ia [Exiguobacterium sp. S3-2]
MRSEKEMMDLVLSLAEQDERIRIVTLEGSRANINIPKDEFQDYDITYFVSDIEPFISNDDWLQFGNIIMMQKPEDMELFPPEEKGFSYL
MLFDYDYNKIDLTLLPLEELDNLKGDKLIKVLIDKDCRIKRDIVPTDIDYHVRKPSAREYDDCCNEFWNVTPYVIKGLCRKEILFAIDHL
NQILRFELLRMMSWKVGKIKTEFSLSVGKNKYINKYIDEDLWNRLLSTYRMDSYENIWKSLFICHQLFREVSKEVAELLGFDPYPEYGKNI
TRYTEDMYKKYVENDYF
```

## The BLAST Search Algorithm

query word ( $W = 3$ )

Query: GSVEDTTGSQSIAALLNKCKTPGGQRQLVNQWIKQPLMDKNRIEERLNVEAFVDAELRQTLQEDL

neighborhood words	score
PQG	18
PEG	15
PRG	14
PKG	14
PNG	13
PDG	13
PHG	13
<b>PMG</b>	13
PSG	13
PQA	12
PQN	12
etc...	

neighborhood score threshold  
( $T = 13$ )

Query: 325 SLAALLNKCKTPGGQRQLVNQWIKQPLMDKNRIEERLNVEA 365  
 +LA++L+ TP G R++ +U+ P+ D + ER + A  
 Sbjct: 290 TLASVLDCTVTE**MGSRMLKRWLHMPVRDTRVLLERQQTIGA** 330

High-scoring Segment Pair (HSP)

<https://www.ncbi.nlm.nih.gov/books/NBK1762/>

# gyrB mutations

protein variant model

**Model Type:** protein variant model

**Model Definition:** Protein Variant Models (PVM) perform a similar search as Protein Homolog Models (PHM), i.e. detect protein sequences based on their similarity to a curated reference sequence, but secondarily screen query sequences for curated sets of mutations to differentiate them from antibiotic susceptible wild-type alleles. PVMs are designed to detect AMR acquired via mutation of house-keeping genes or antibiotic targets, e.g. a mutated gyrase resistant to aminocoumarin antibiotics. PVMs include a protein reference sequence (often from antibiotic susceptible wild-type alleles), a curated bit-score cut-off, and mapped resistance variants. Mapped resistance variants may include any or all of single point mutations, insertions, or deletions curated from the scientific literature. A Strict RGI match has a BLASTP bit-score above the curated BLASTP cutoff value and contains at least one curated mutation from amongst the mapped resistance variants, while a Loose RGI match has a bit-score less than the curated BLASTP bit-score cut-off but still contains at least one curated mutation from amongst the mapped resistance variants.

**Bit-score Cut-off (blastP): 1200**

**Legend:**

- discovered in clinical, agricultural, or environmental isolates
- discovered via laboratory selection experiments
- ReSeqTB <https://platform.reseqtb.org>

**Published Variants:**

<a href="#">PMID: 25003707</a>	45612,D91G,V199I+ 45613,D481E   45612,D91N+45613,D481E,D484K   45612,D91N,V199I+45613,D481E,D484K   45612,N87K+45613,F438S
	45612,N87K,V199A+45613,D481E,R484K   45612,N87K,V199I+45613,D484K
<a href="#">PMID: 22221614</a>	E463K
<a href="#">PMID: 30774400</a>	45612,A129T+45613,S479G   45612,D91N+45613,R484K   45612,D91N,A129T+45613,S479G   45612,N87Y+45613,R484K

[Protein](#)   [DNA](#)

```
>gb|WP_001182024.1|+|Helicobacter pylori gyrB conferring resistance to fluoroquinolones [Helicobacter pylori]
MQNYQSHSIKVLEGVRKPGMYIGDTNVGGLHHMYYEVVDNAVDESMAGFCDTINIT
LTDEGSCIVEDNGRGPVDIHPTEKIPACTVVLTILHAGGKFDNDTYKVSGGLHGVGSV
VNALSKRLIMTIKKEGQIYRQEFEKGPIPTSELEIIGTKSAKESGTTIEFFPDESVMEVV
EFQAGILQKRFKEMAYLNDGLKISFKEEKTLQETYFYEDGLKQFVKDSAKKELLTPIS
FKSMDEERTSIEVALAYADDYNENTLSFVNNIKTSEGGTHEAGFKMGLSKAILQYIGNN
IKTKESRPISEDIKEGLIAVVSLSKMSPLFEGQTKSKLGSSYARALVSKLVYDKIHQFLE
ENPNEAKIIANKALLAAKAREASKKARELTRKKDNLNSVGTLPGKLADCQSKDPLESEIFL
VEGDSAGGSAKOGRDRVFOAILPLKGKILNVEKSHLSKILKSEEIKNMITAFCGGIQESF
DIERLRYHKIIIIMTDADVDGSHIQTLLMTFFYRYLRPLIEQGHVYIAQAPLYKYKKGKTE
TYLKDSVALDHFLIEHGINSVDIEGIGKNDLMNLLKVARHYRYALLELEKRYNILLEILRF
```

mutations  
conferring  
resistance

protein homolog model

e.g. extended spectrum beta-lactamase

protein variant model \*

e.g. fluoroquinolone resistant gyrB

rRNA gene variant model \*

e.g. hygromycin B resistant 23S rRNA

protein overexpression model \*

e.g. upregulated MDR in *Pseudomonas*

protein knockout model

e.g. imipenem resistance in *A. baumannii*



Karyn Mukiri, PhD Student

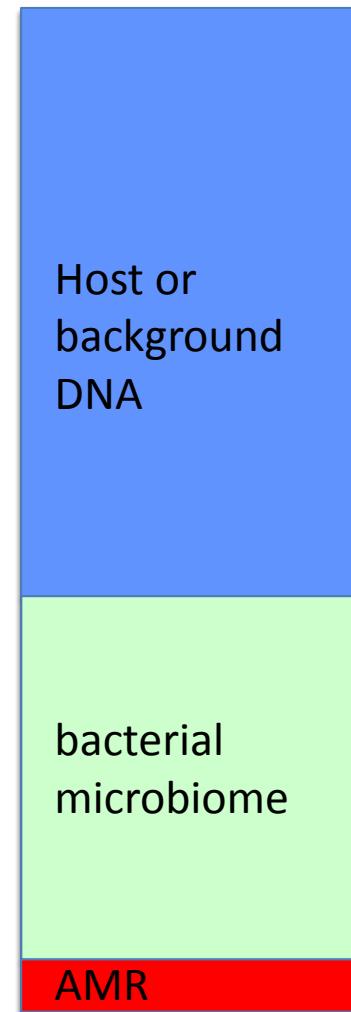
nonfunctional insertion model

gene cluster meta-model

e.g. vancomycin resistance

efflux pump system meta-model

e.g. MexAB-OprM efflux pump system



AMR genes are the needle in the haystack

It takes a lot of (expensive) DNA sequencing

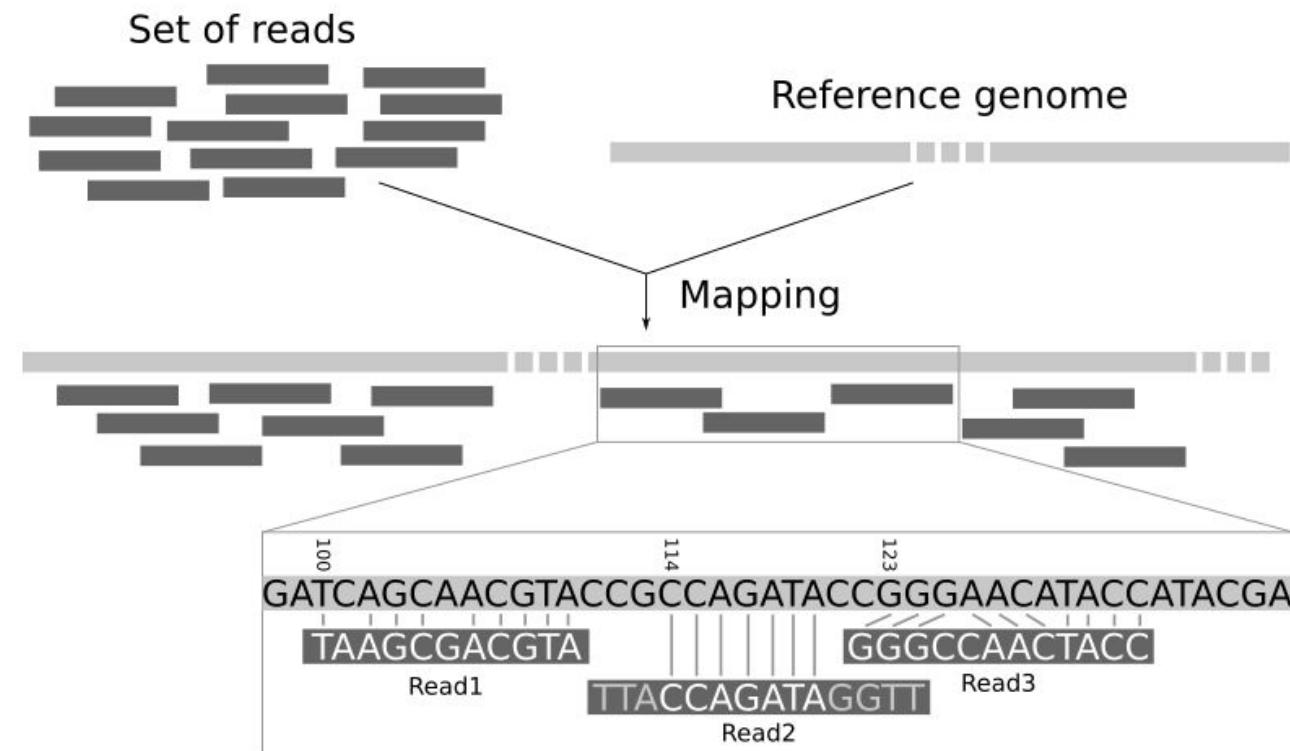
The data is very fragmentary

Analysis sensitive to the diversity of reference data

Even if we find an AMR gene, which bacteria did it come from?

Which gene did that 250 bp sequencing read come from?

The Burrows–Wheeler transform, a.k.a. 'read mapping'



*Common software:*  
bwa  
bowtie2

<https://galaxyproject.github.io>

# The AMR Allele Network Problem

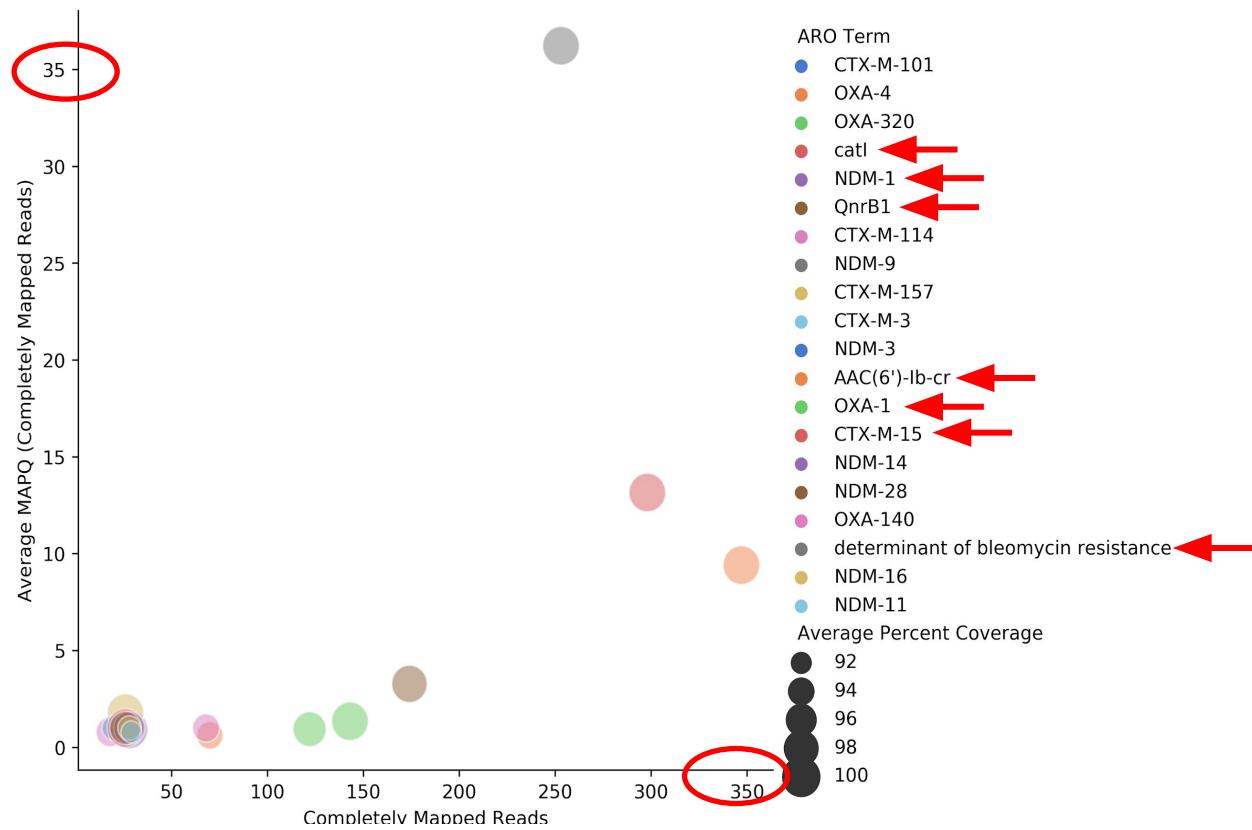
**TEM beta-lactamase [AMR Gene Family]**

[Download Sequences](#)

Accession	ARO:3000014
Definition	TEM-1 is the most commonly-encountered beta-lactamase in gram-positive bacteria. It is responsible for the ampicillin and penicillin resistance that is seen often found in <i>E. coli</i> and <i>K. pneumoniae</i> , they are also found in other bacteria. TEM-1 is responsible for the ESBL phenotype cluster around the active site of the beta-lactamases. The active site to beta-lactam substrates also typically enhances the resistance of TEM-1. Although inhibitor-resistant TEM-1 is not yet described, there are at least 19 distinct inhibitor-resistant TEM beta-lactamases. In <i>K. pneumoniae</i> , <i>Klebsiella oxytoca</i> , <i>P. mirabilis</i> , and <i>Citrobacter freundii</i> , TEM-1 is resistant to ampicillin and sulbactam, thereby showing clinical resistance to the beta-lactam antibiotics. In <i>E. coli</i> , TEM-1 is resistant to ampicillin/sulbactam, they normally remain susceptible to inhibition by clavulanic acid and carbenicillin-clavulanate, although resistance has been described.
Drug Class	<a href="#">penam</a> , <a href="#">monobactam</a> , <a href="#">cephalosporin</a> , <a href="#">penem</a>
Resistance Mechanism	<a href="#">antibiotic inactivation</a>
Classification	<a href="#">12 ontology terms</a>   <a href="#">Show</a>
Parent Term(s)	<a href="#">6 ontology terms</a>   <a href="#">Show</a>
Sub-Term(s)	<p><b>232 ontology terms</b>   <a href="#">Hide</a></p> <ul style="list-style-type: none"> <li>+ <a href="#">TEM-1</a></li> <li>+ <a href="#">TEM-2</a></li> <li>+ <a href="#">TEM-3</a></li> <li>+ <a href="#">TEM-4</a></li> <li>+ <a href="#">TEM-5</a></li> <li>+ <a href="#">TEM-6</a></li> <li>+ <a href="#">TEM-7</a></li> <li>+ <a href="#">TEM-8</a></li> <li>+ <a href="#">TEM-9</a></li> <li>+ <a href="#">TEM-10</a></li> <li>+ <a href="#">TEM-11</a></li> <li>+ <a href="#">TEM-12</a></li> <li>+ <a href="#">TEM-13</a></li> <li>+ <a href="#">TEM-15</a></li> <li>+ <a href="#">TEM-16</a></li> <li>+ <a href="#">TEM-17</a></li> <li>+ <a href="#">TEM-18</a></li> <li>+ <a href="#">TEM-19</a></li> <li>+ <a href="#">TEM-20</a></li> <li>+ <a href="#">TEM-21</a></li> <li>+ <a href="#">TEM-22</a></li> <li>+ <a href="#">TEM-24</a></li> <li>+ <a href="#">TEM-25</a></li> <li>+ <a href="#">TEM-26</a></li> <li>+ <a href="#">TEM-27</a></li> <li>+ <a href="#">TEM-28</a></li> <li>+ <a href="#">TEM-29</a></li> <li>+ <a href="#">TEM-30</a></li> <li>+ <a href="#">TEM-31</a></li> <li>+ <a href="#">TEM-32</a></li> <li>+ <a href="#">TEM-33</a></li> <li>+ <a href="#">TEM-34</a></li> </ul>

## The AMR Allele Network Problem

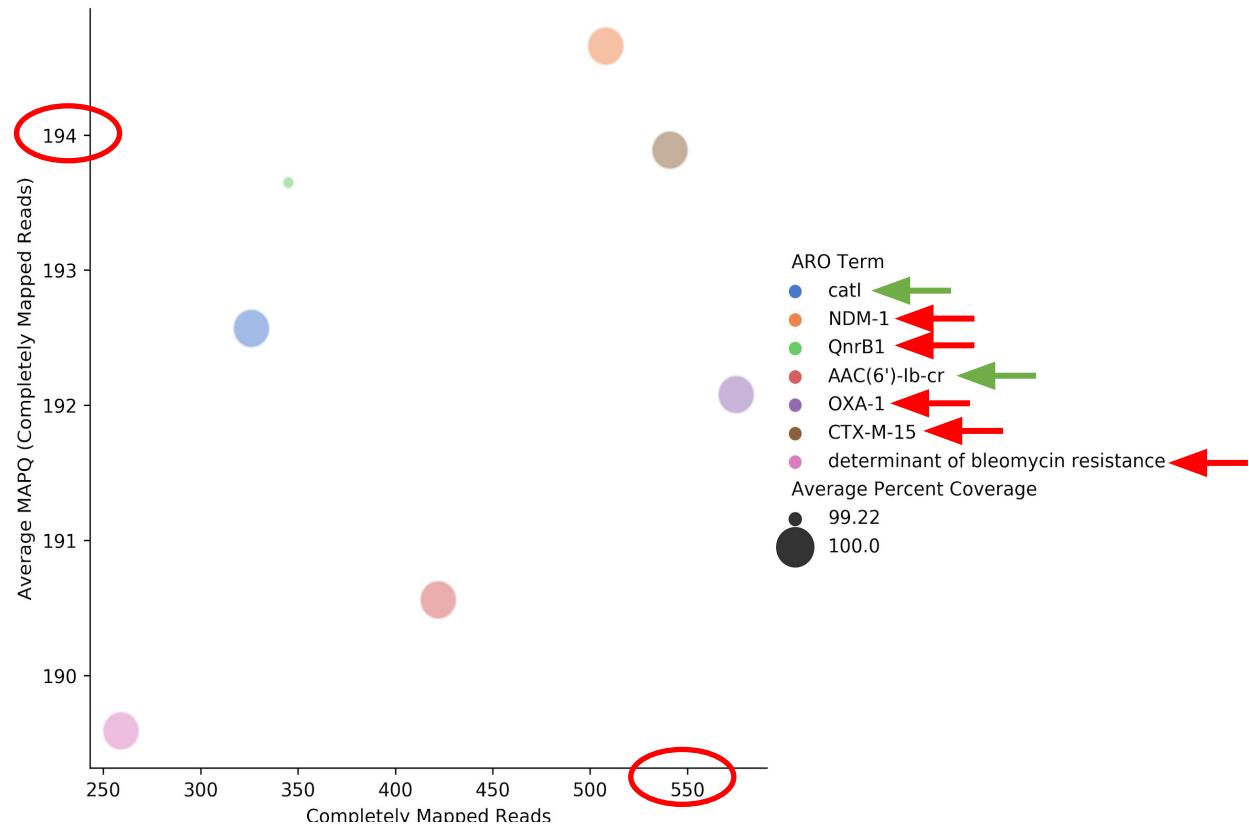
Simulated genome (10x coverage) – Bowtie2



Amos Raphenya, Lead CARD Developer

## The AMR Allele Network Problem

Simulated genome (10x coverage) – KMA



Amos Raphenya, Lead CARD Developer

Thank you Resfinder team!

## The AMR Allele Network Problem #2

The screenshot shows the homepage of the Comprehensive Antibiotic Resistance Database (CARD). The top navigation bar includes links for Browse, Analyze, Download, and About, along with a search bar. The main content area features a title "The Comprehensive Antibiotic Resistance Database" and a subtitle "A bioinformatic database of resistance genes, their products and associated phenotypes." Below this, it displays various statistics: 6453 Ontology Terms, 4937 Reference Sequences, 1788 SNPs, 2775 Publications, 4983 AMR Detection Models, Resistome predictions for 263 pathogens, 14795 chromosomes, 2675 genomic islands, 30591 plasmids, 105556 WGS assemblies, and 231629 alleles. It also mentions the CARD Bait Capture Platform 1.0.0 and State of the CARD 2021 Presentations & Demonstrations. The page is divided into several sections: "Browse" (describing the curated collection of resistance determinants and antibiotics), "Analyze" (mentioning BLAST and RGI software), "Download" (describing data formats and availability), "Resistomes, Variants, & Prevalence" (describing computer-generated predictions for important pathogens), "CARD:Live" (describing the project for pathogen identification and analysis), and a "Timeline" section at the bottom right.

‘Canonical CARD’ reflects expert curation of the literature:

- Does not reflect the full diversity of AMR genes, alleles, and mutations in clinical, agricultural, veterinary, and environmental settings

## The AMR Allele Network Problem #2

ANT(6)-Ia	
<a href="#">Download Sequences</a>	
Accession	ARO:3002626
Synonym(s)	<i>ant6</i> , <i>aadE</i>
Definition	ANT(6)-Ia is an aminoglycoside nucleotidyltransferase gene encoded by plasmids and chromosomes in <i>Staphylococcus epidermidis</i> , <i>E. faecium</i> , <i>Streptococcus suis</i> , <i>S. aureus</i> , <i>E. faecalis</i> and <i>Streptococcus mitis</i>
AMR Gene Family	<a href="#">ANT(6)</a>
Drug Class	<a href="#">aminoglycoside antibiotic</a>
Resistance Mechanism	<a href="#">antibiotic inactivation</a>
<b>Molecular Epidemiology</b>	
Resistomes with Perfect Matches	<p><i>Clostridioides difficile</i><sup>q+wgs</sup>, <i>Clostridium perfringens</i><sup>wgs</sup>, <i>Enterococcus casseliflavus</i><sup>p</sup>, <i>Enterococcus faecalis</i><sup>q+p+wgs+qi</sup>, <i>Enterococcus faecium</i><sup>q+p+wgs+qi</sup>, <i>Enterococcus hirae</i><sup>q+wgs+qi</sup>, <i>Erysipelothrix rhusiopathiae</i><sup>q+qi</sup>, <i>Herbinix luporum</i><sup>q+qi</sup>, <i>Jeotgalibaca arthritidis</i><sup>q</sup>, <i>Lactobacillus animalis</i><sup>wgs</sup>, <i>Listeria innocua</i><sup>p+wgs</sup>, <i>Staphylococcus aureus</i><sup>q+p+wgs+qi</sup>, <i>Staphylococcus haemolyticus</i><sup>q+wgs</sup>, <i>Staphylococcus saprophyticus</i><sup>p+wgs</sup>, <i>Streptococcus agalactiae</i><sup>q+wgs</sup>, <i>Streptococcus pasteurianus</i><sup>wgs</sup>, <i>Streptococcus pneumoniae</i><sup>wgs</sup>, <i>Streptococcus pyogenes</i><sup>wgs</sup>, <i>Streptococcus suis</i><sup>q+wgs+qi</sup></p>
Resistomes with Sequence Variants	<p><i>Bacteroides ovatus</i><sup>wgs</sup>, <i>Campylobacter jejuni</i><sup>q+p+wgs+qi</sup>, <i>Clostridioides difficile</i><sup>q+wgs</sup>, <i>Clostridium perfringens</i><sup>wgs</sup>, <i>Enterobacter hormaechei</i><sup>wgs</sup>, <i>Enterococcus avium</i><sup>p</sup>, <i>Enterococcus casseliflavus</i><sup>p</sup>, <i>Enterococcus faecalis</i><sup>q+p+wgs+qi</sup>, <i>Enterococcus faecium</i><sup>q+p+wgs+qi</sup>, <i>Enterococcus hirae</i><sup>q+p+wgs+qi</sup>, <i>Erysipelothrix rhusiopathiae</i><sup>q+qi</sup>, <i>Escherichia coli</i><sup>wgs</sup>, <i>Herbinix luporum</i><sup>q+qi</sup>, <i>Jeotgalibaca arthritidis</i><sup>q</sup>, <i>Klebsiella pneumoniae</i><sup>wgs</sup>, <i>Lactobacillus animalis</i><sup>wgs</sup>, <i>Listeria innocua</i><sup>p+wgs</sup>, <i>Staphylococcus aureus</i><sup>q+p+wgs+qi</sup>, <i>Staphylococcus haemolyticus</i><sup>q+wgs</sup>, <i>Staphylococcus pseudintermedius</i><sup>q+wgs</sup>, <i>Staphylococcus saprophyticus</i><sup>p+wgs</sup>, <i>Streptococcus agalactiae</i><sup>q+wgs</sup>, <i>Streptococcus pasteurianus</i><sup>wgs</sup>, <i>Streptococcus pneumoniae</i><sup>wgs</sup>, <i>Streptococcus pyogenes</i><sup>wgs</sup>, <i>Streptococcus suis</i><sup>q+wgs+qi</sup></p>
Classification	<a href="#">9 ontology terms</a>   <a href="#">Show</a>
Parent Term(s)	<a href="#">2 ontology terms</a>   <a href="#">Show</a>
Publications	Pinto-Alphandary H, et al. 1990. Antimicrob Agents Chemother 34(6): 1294-1296. Emergence of aminoglycoside resistance genes <i>aadA</i> and <i>aadE</i> in the genus <i>Campylobacter</i> . ( <a href="#">PMID 2168151</a> ) Gill SR, et al. 2005. J Bacteriol 187(7): 2426-2438. Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant <i>Staphylococcus aureus</i> strain and a biofilm-producing methicillin-resistant <i>Staphylococcus epidermidis</i> strain. ( <a href="#">PMID 15774886</a> )

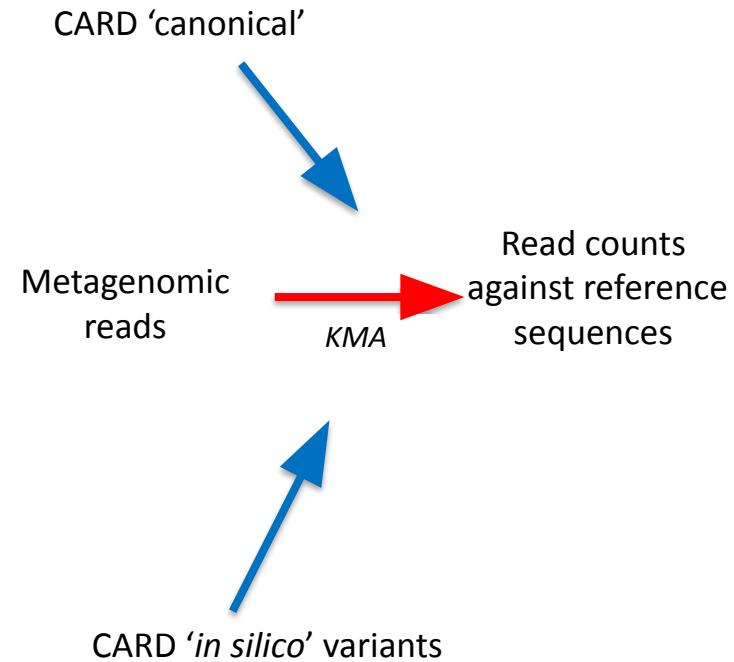
377 pathogens, 21079 chromosomes, 2662 genomic islands,  
 41828 plasmids, 155606 WGS assemblies, **322710 alleles**

RGI Perfect &  
 Strict Hits!

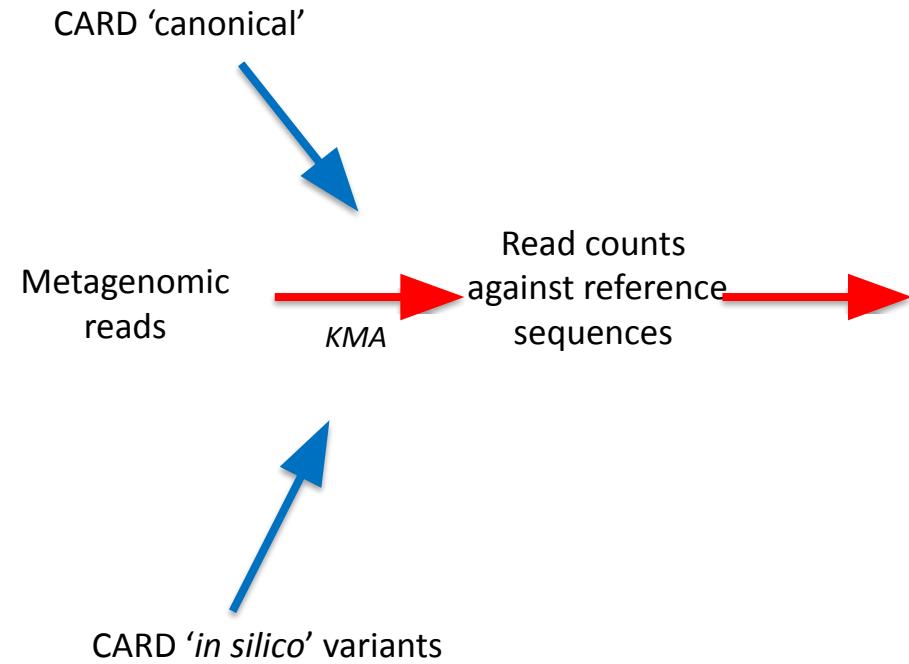
## A metagenomics workflow

Metagenomic  
reads

## A metagenomics workflow

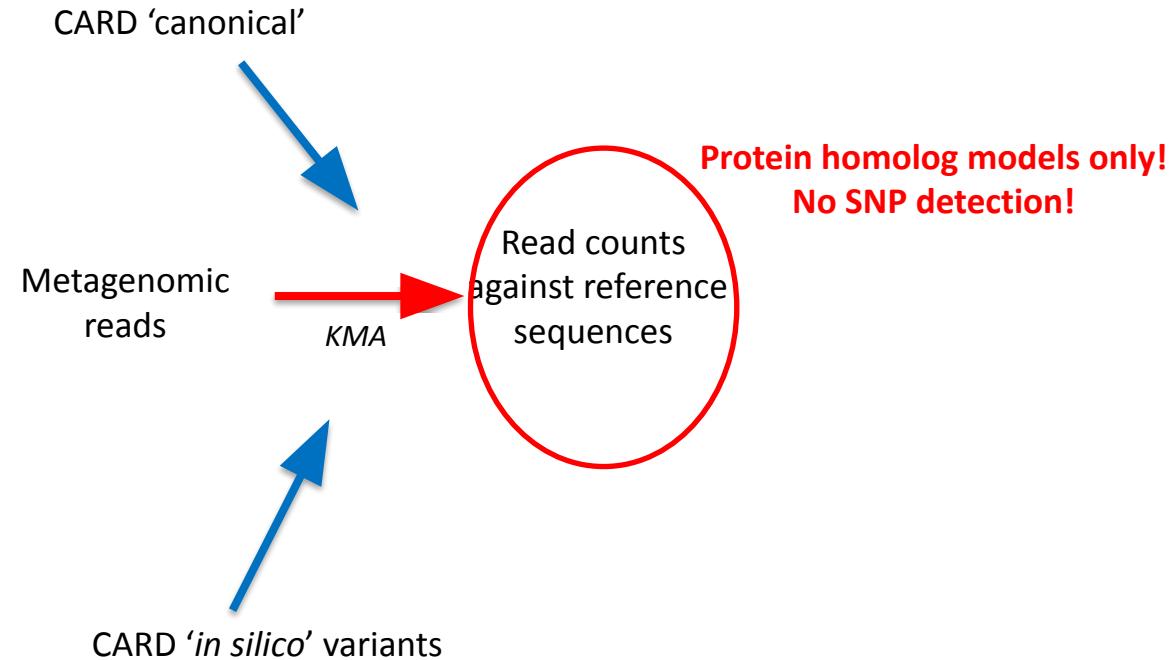


## A metagenomics workflow

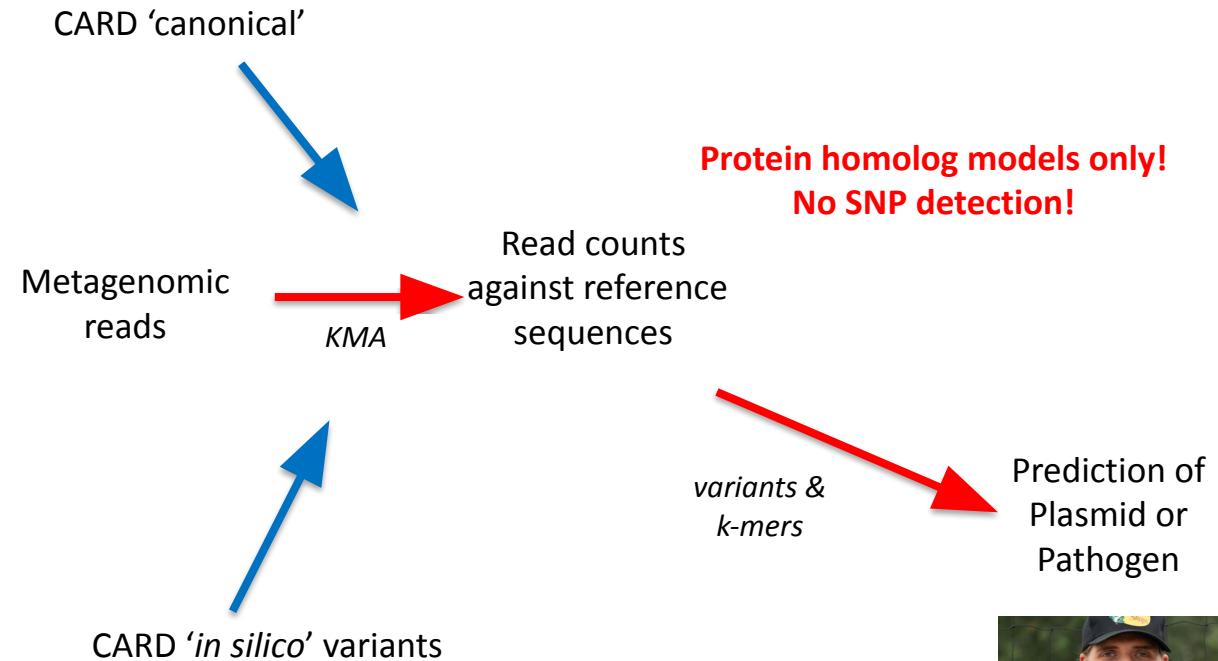


ARO Term	All Mapped Reads
tet(Q)	42684
tet(X)	7393
acrD	1881
APH(6)-Id	1418
sul2	961
tet(W)	939
aad(6)	99
tet(X4)	2

## A metagenomics workflow

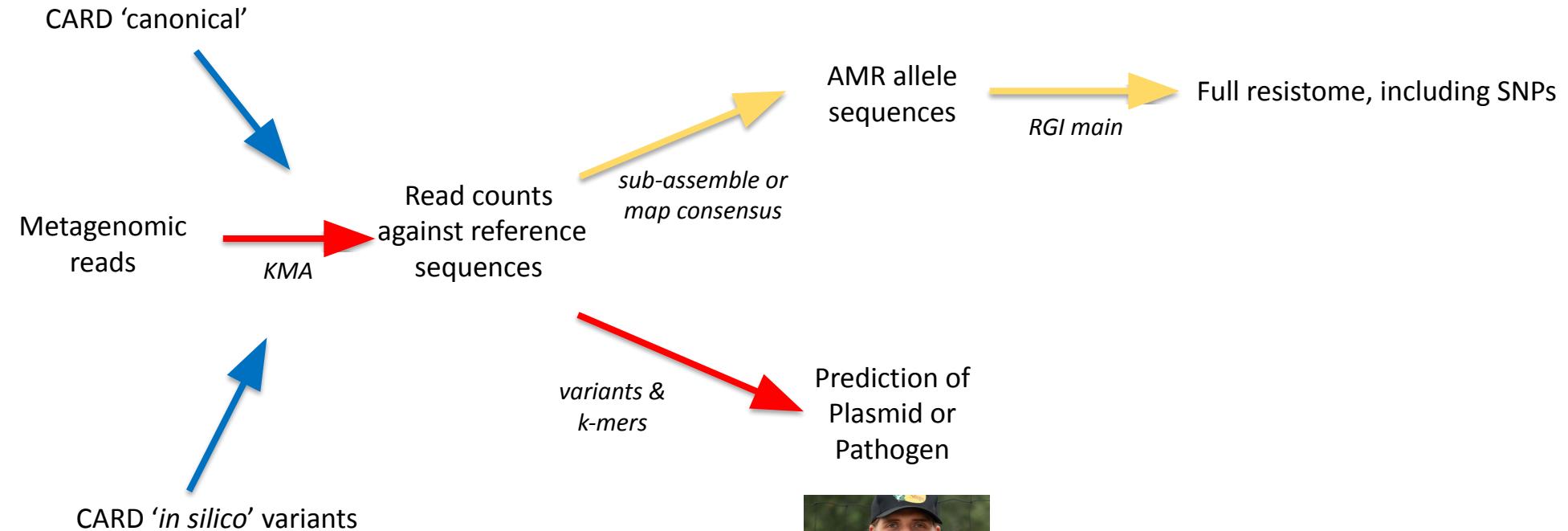


## A metagenomics workflow



Mateusz Włodarski  
MSc Student

## A metagenomics workflow



Mateusz Włodarski  
MSc Student

# CARD Bait Capture Platform

A robust and reliable targeted bait capture method for metagenomic detection of antibiotic resistance determinants in complex samples, including hybridization bait FASTA sequences and laboratory protocol.



Allison Gutor, Ph.D.



Brian Alcock, Lead CARD Curator

## Capturing the Resistome: a Targeted Capture Method To Reveal Antibiotic Resistance Determinants in Metagenomes

Allison K. Gutor,<sup>a,b,c</sup> Amogelang R. Raphenya,<sup>a,b,c</sup> Jennifer Klunk,<sup>c,d</sup> Melanie Kuch,<sup>c,d</sup> Brian Alcock,<sup>a,b,c</sup> Michael G. Surette,<sup>a,b,c</sup>  
Andrew G. McArthur,<sup>a,b,c</sup> Hendrik N. Poinar,<sup>a,b,c,d</sup> Gerard D. Wright<sup>a,b,c</sup>

<sup>a</sup>David Braley Centre for Antibiotic Discovery, McMaster University, Hamilton, Ontario, Canada

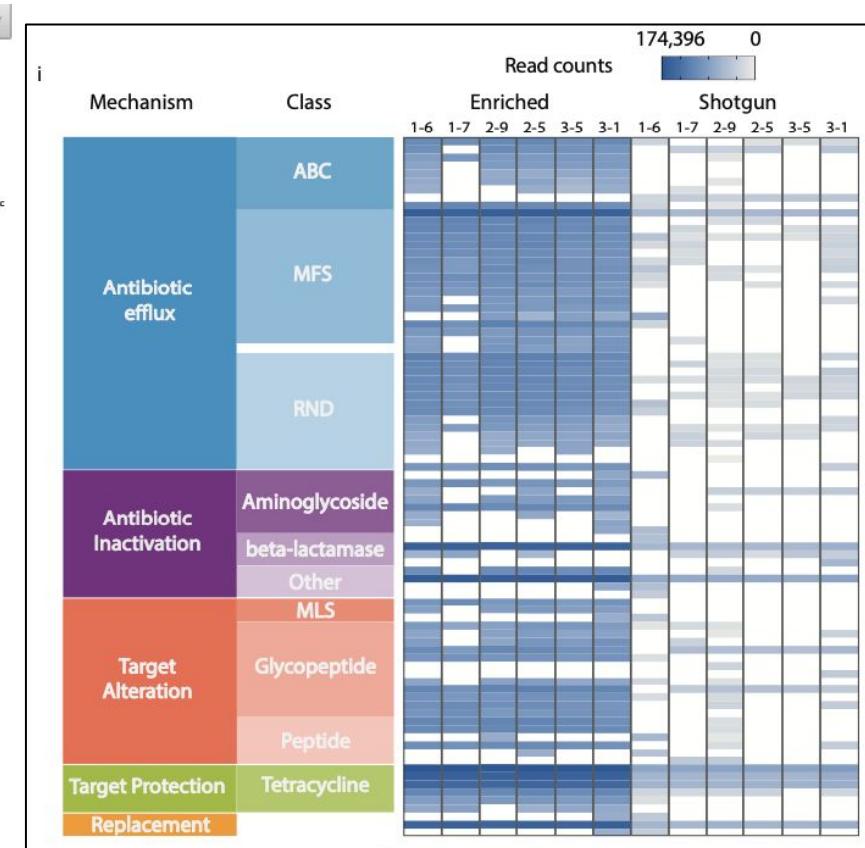
<sup>b</sup>Michael G. DeGroote Institute for Infectious Disease Research, McMaster University, Hamilton, Ontario, Canada

<sup>c</sup>Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario, Canada

<sup>d</sup>McMaster Ancient DNA Centre, Department of Anthropology and Biochemistry, McMaster University, Hamilton, Ontario, Canada

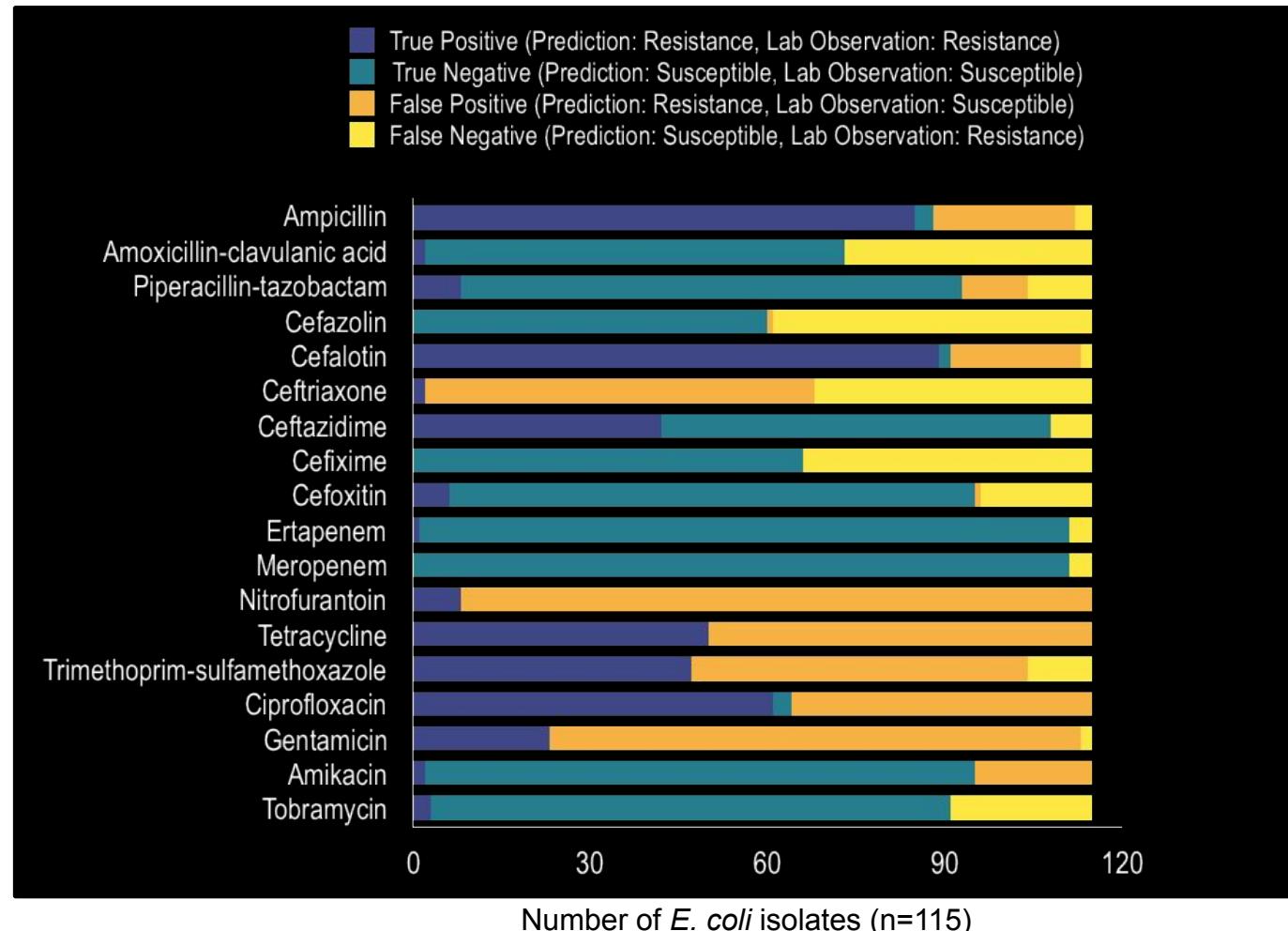
**ABSTRACT** Identification of the nucleotide sequences encoding antibiotic resistance elements and determination of their association with antibiotic resistance are critical to improve surveillance and monitor trends in antibiotic resistance. Current methods to study antibiotic resistance in various environments rely on extensive deep sequencing or laborious culturing of fastidious organisms, both of which are heavily time-consuming operations. An accurate and sensitive method to identify both rare and common resistance elements in complex metagenomic samples is needed. Referencing the sequences in the Comprehensive Antibiotic Resistance Database, we designed a set of 37,826 probes to specifically target over 2,000 nucleotide sequences associated with antibiotic resistance in clinically relevant bacteria. Testing of this probe set on DNA libraries generated from multidrug-resistant bacteria to selectively capture resistance genes reproducibly produced higher numbers of reads on target at a greater length of coverage than shotgun sequencing. We also identified additional resistance gene sequences from human gut microbiome samples that sequencing alone was not able to detect. Our method to capture the resistome enables a sensitive means of gene detection in diverse environments where genes encoding antibiotic resistance represent less than 0.1% of the metagenome.

**KEYWORDS** antibiotic resistance, resistome, sequencing, targeted capture



# Analytics

# Rules-based algorithm leads to poor phenotype prediction

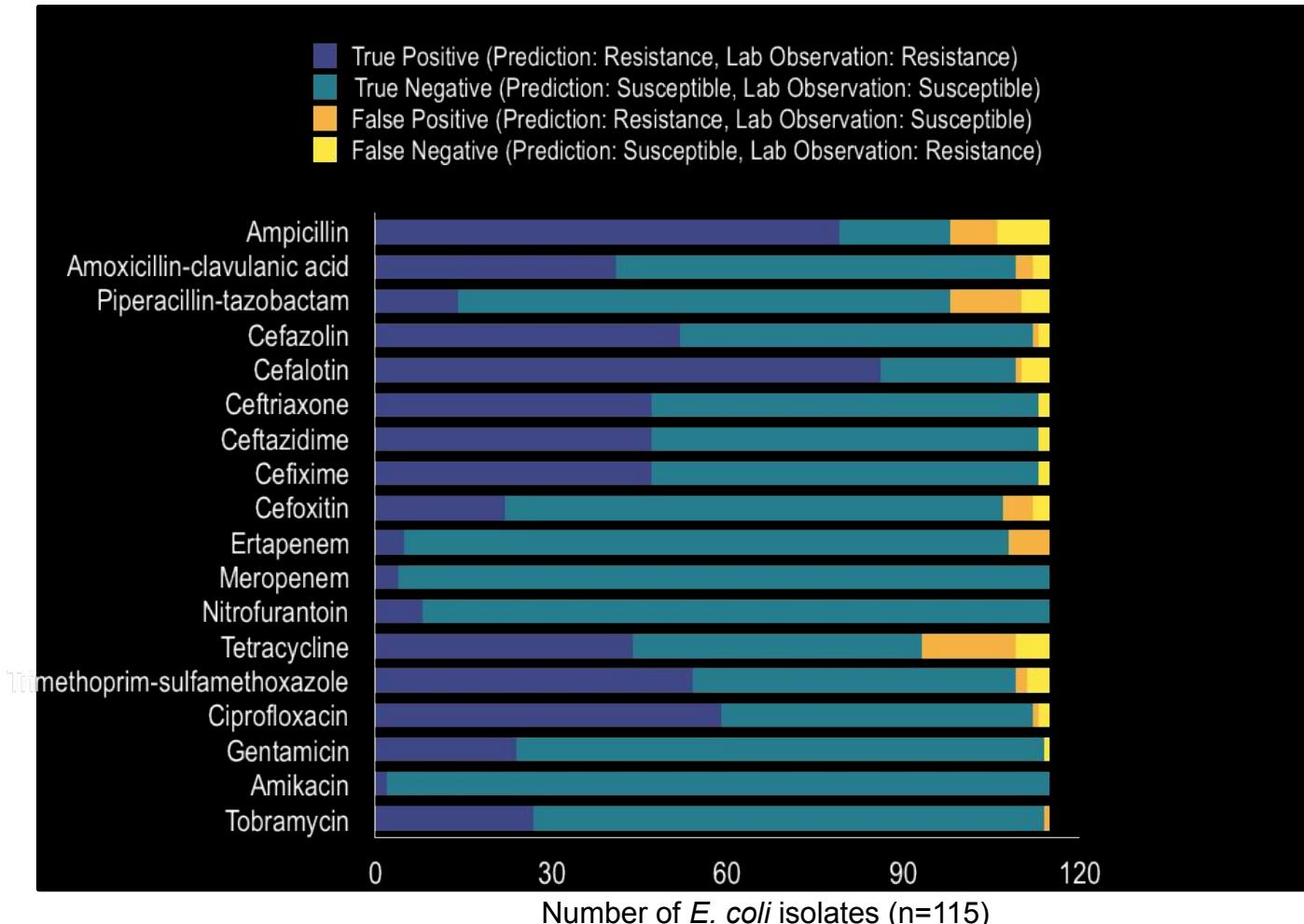


Presence or absence of a gene is not enough!



Dr. Kara Tsang

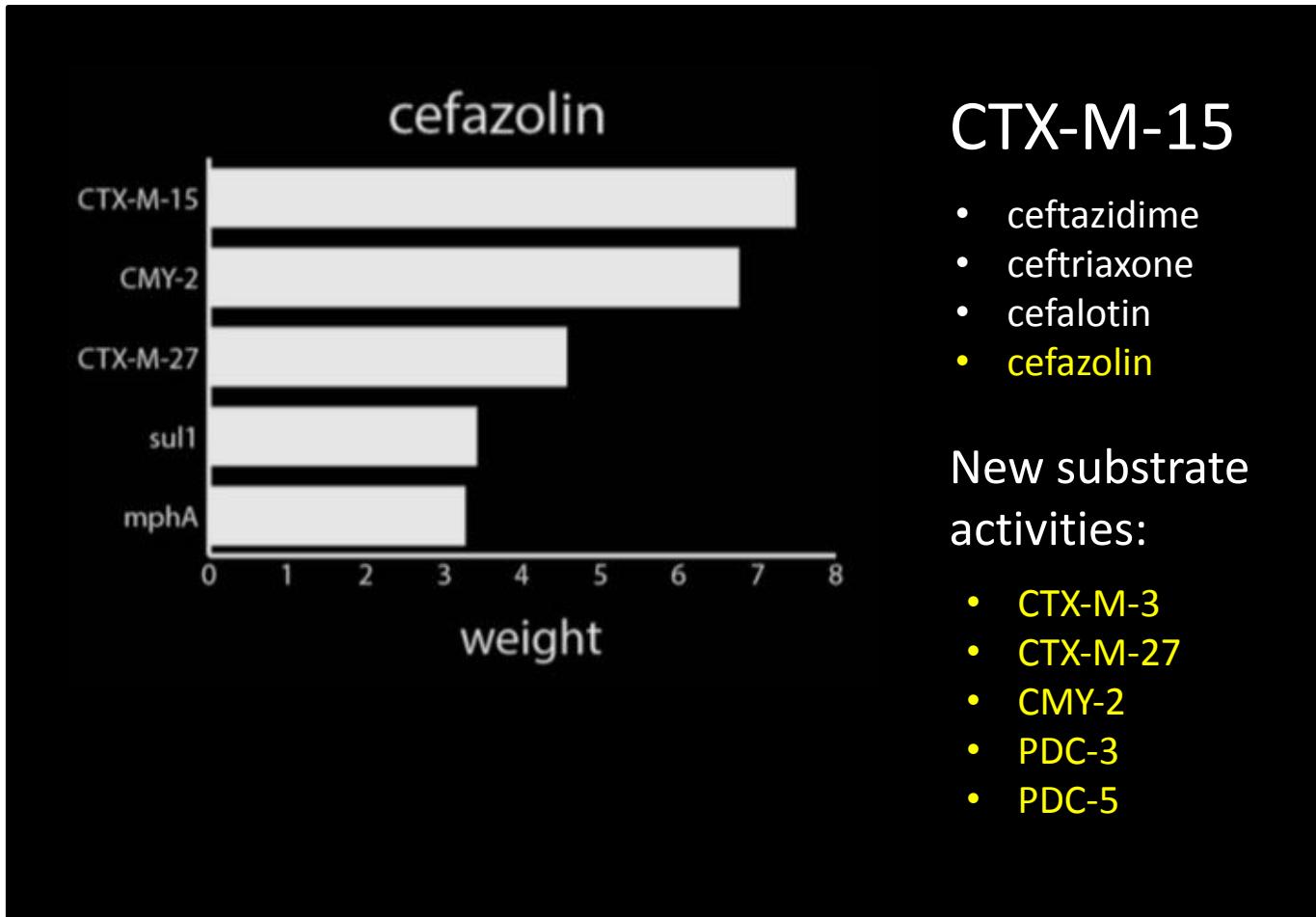
# Machine learning algorithm leads to improved phenotype prediction accuracy



Dr. Kara Tsang

# Weights of predictive importance identify novel substrate activity

Antibiotic Resistance Platform

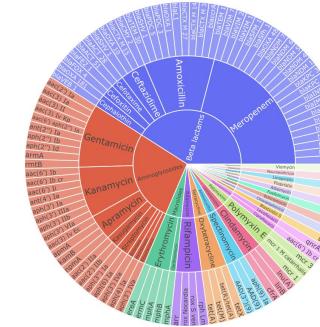


## CTX-M-15

- ceftazidime
- ceftriaxone
- cefalotin
- **cefazolin**

## New substrate activities:

- CTX-M-3
- CTX-M-27
- CMY-2
- PDC-3
- PDC-5



The Antibiotic Resistance Platform (ARP) is a cell-based array of mechanistically distinct individual resistance elements in an identical genetic background. The ARP consists of resistance elements to 18 classes of antibiotics, represented by the inner wheel. The middle wheel divides the 18 classes into individual antibiotics found within each class. The outer wheel depicts the array of resistance genes within the ARP, grouped based on antibiotic substrate specificity. There are over 100 antibiotic resistance genes currently in the ARP.

<https://www.thewrightlab.com/antibiotic-resistance-platform>

# Machine learning algorithm leads to improved phenotype prediction accuracy?

- 102 *P. aeruginosa* isolates  
Genomes completely sequenced  
Resistance profiles for 18 antibiotics  
**Machine learning less reliable**



*Pseudomonas aeruginosa*

# Last Thoughts

### *Context is Everything*

*Plasmid borne AMR genes are generally high threat  
Does the pathogen form biofilms?  
How to deploy molecular knowledge?*

### *Capacity and Technology are key considerations*

*Breadth and volume of culture capacity?  
Genome assembly and read mapping are computationally intensive  
How to keep up with 5000+ AMR genes?*

### *Confidence can be a challenge*

*How to be sure of the SNP call when there is sequencing error?  
Plasmid AMR genes are easy to find, but resolving plasmids is very hard*

### *How tight is the genotype-phenotype relationship?*

*Some SNPs are potent, others are additive  
How to interpret efflux?  
How good is the underlying knowledge?*

### *Solve the Genome & the Plasmids in the first pass*

*Nanopore long-read sequencing is cheap, fast, and awesome – GPU processors!  
Combined with Illumina you get complete chromosomes & plasmids  
Plasmid analysis software is getting better (e.g. MOB-Suite from PHAC)*

### *By-Pass Culture*

*Bait capture is becoming increasingly effective, thus bypassing culture  
Demand for C. difficile and H. pylori baits in the last two years  
How can you handle a big outbreak if culture is involved?*

### *Metagenomics is a big part of the future*

*Unbiased or baited, either way sequencing costs are dropping  
Algorithms for read alignment to reference are improving (e.g. KMA)*

### *Inverse relationship between speed and scale*

*While sequencing costs are dropping, library preparation costs are not changing  
Library construction is time expensive - only small outbreaks can be done quickly  
Small numbers of samples uses smaller sequencers and is more expensive  
Bar-coding and robotics are the future*

# We are on a Coffee Break & Networking Session

Workshop Sponsors:



Canadian Centre for  
Computational  
Genomics



HPC4Health

