Comprehensive Antibiotic Resistance Database (CARD)

ICARe 2017, Annecy
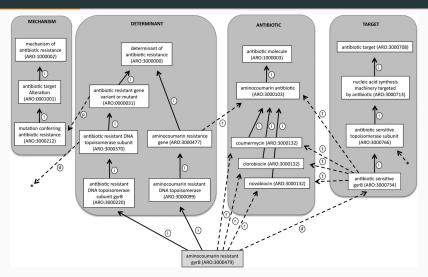
Finlay Maguire

Faculty of Computer Science, Dalhousie University (Nova Scotia, Canada)

- Built around ontological principles
- Automated literature mining with manual curation
- Multiple detection models of AMR determinants
- Specialised analysis tools with per-item detection threshold curation

Antibiotic Resistance Ontology

i=is_a, p=part_of, d=derived_from t=targeted_by, r=confers_resistance

McArthur *et. al.* Antimicrob. Agents Chemother. July 2013 vol. 57 no. 7 3348-3357

2

AAC(3)-Ia

Accession	ARO:3002528
Synonym(s)	*AAC(3)-Ia C aacC1 aacC-A1*
Definition	AAC(3)-Ia is an aminoglycoside acetyltransferase encoded by plasmids, transposons, integrons in S. marcescens, E. coli, Acinetobacter baumannii, Klebsiella pneumoniae, Klebsiella oxytoca, P. aeruginosa, Salmonella typhimurium and Proteus mirabilis
Classification	+ process or component of antibiotic biology or chemistry + mechanism of antibiotic resistance + antibiotic molecule + aminoglycoside antibiotic [Drug Class] + determinant of antibiotic resistance + antibiotic inactivation [Resistance Mechanism] + acylation of antibiotic conferring resistance + determinant of aminoglycoside resistance + antibiotic inactivation enzyme + aminoglycoside acetyltransferase (AAC) + antibiotic mixture
Parent Term(s)	+ gentamicin C (confers_resistance_to_drug) + sisomicin (confers_resistance_to_drug) [Antibiotic] + AAC(3) (is_a) [AMR Gene Family] + gentamicin B (confers_resistance_to_drug) [Antibiotic]
Sub-Term(s)	
Publications	Wohlleben W, et al. 1989. Mol Gen Genet 217(2-3): 202-208. On the evolution of Tn21-like multiresistance transposons: sequence analysis of the gene (aacC1) for gentamicin acetyltransferase-3-I(AAC(3)-I), another member of the Tn21-based expression cassette. (PMID 2549372) Javier Teran F, et al. 1991. J Antimicrob Chemother 28(3): 333-346. Characterization of two aminoglycoside-(3)-N-acetyltransferase genes and assay as epidemiological probes.

CARD*Shark

- 'CARD*Shark' text-mining algorithms to prioritise literature for manual monthly biocuration.
- ARO word-association scoring matrices based on title and abstract to rank new PubMed publications
- High scoring publications prioritised for manual review
- Curation team of experts with dedicated head curator within McArthur Lab
- Determinant inclusion criteria:
  - Must confer resistance to known antimicrobial *in vivo* or *in vitro*
  - Peer-reviewed scientific publication
  - Available in NCBI's GenBank repository

Detection Models

- Available:
  - Protein Homology
  - Protein Variant
  - Ribosomal Variant
- In process of release:
  - Efflux Pump
  - Gene Cluster
- In development:
  - Protein Knockout
  - Nonfunction Insertion
  - Protein Over-expression
  - Protein Domain

Resistance Gene Identifier (RGI)

- Analyse protein sequences or genomic contigs
- Homology detection using DIAMOND (BLASTP/X accelerator)
- ARO allows summarising of results in terms of class/resistances/mechanism etc.
- Curated bitscore cut-off thresholds:
  - Perfect (100% match to curated sequences for clinical surveillance)
  - Strict (manually curated cut-off for identification of previously unknown variants)
  - Loose (discovery mode to help detect emergent or distant homologues)

6

CARDPredicted

Prevalence of **NDM-1** among the sequenced genomes, plasmids, and whole-genome shotgun assemblies available at NCBI for clinically important pathogens (see methodological details and complete list of analyzed pathogens). Values reflect percentage of genomes, plasmids, or whole-genome shotgun assemblies that have at least one hit to the AMR detection model. Default view includes percentages calculated based on Perfect plus Strict RGI hits. Select the checkbox to view percentages based on only Perfect matches to AMR reference sequences curated in CARD (note: this excludes resistance via mutation as references in protein variant models are often wild-type, sensitive sequences).

protein homolog model prevalence table: [View raw data]

Species	NCBI Chromosome	NCBI Plasmid	NCBI WGS
Acinetobacter baumannii	5%	0.62%	0.99%
Enterobacter aerogenes	0%	0%	0.79%
Enterobacter cloacae	0%	1.69%	2.4%
Klebsiella pneumoniae	0%	6.15%	3.39%
Pseudomonas aeruginosa	1.68%	0%	0.31%

☐ Show Perfect Only

CARDPredicted

Phenotype

Prevalence of AMR genes and variants organized by Antibiotic Resistance Ontology phenotypic classification. Values reflect percentage of completely sequenced genomes, completely sequenced plasmids, or whole-genome shotgun assemblies that have at least one AMR determinant associated with the phenotypic class. Default view includes percentages calculated based on Perfect plus Strict RGI hits. Select the checkbox to view percentages based on only Perfect matches to AMR reference sequences curated in CARD (note: this excludes resistance via mutation as references in protein variant models are often wild-type, sensitive sequences).

Species	aminocoumarin antibiotic	aminoglycoside antibiotic	beta-lactam antibiotic	diaminopyrimidine antibiotic	fluoroquinolone antibiotic	fosfomycin	fusidic acid	gl
Acinetobacter baumannii	0%	92.42%	94.57%	3.04%	0.04%	0.08%	0%	
Campylobacter coli	0%	12.02%	52.6%	0%	0.4%	0%	0%	
Campylobacter jejuni	0%	5.46%	51.9%	0%	26.27%	0%	0%	
Enterobacter aerogenes	0%	15.11%	23.02%	10.79%	20.14%	94.24%	0%	
Enterobacter cloacae	0%	47.1%	92.66%	37.37%	43.52%	87.71%	0%	
Enterococcus faecium	0%	89.21%	0.11%	49.67%	0%	0.44%	0%	
Haemophilus influenzae	0%	1.88%	16.88%	0%	0%	0%	0%	

Contribute

- McArthur Lab always keen for assistance of new curators/corrections
- CARD curators or developers responsive and available directly at card@mcmaster.ca or via Twitter at @arpcard
- Commercial users can contact and organise licensing agreements for usage