The Comprehensive Antibiotic Resistance Database

Eddie Azinge and Mary Balducci

Department of Biology &

Department of Computer Science

Loyola Marymount University
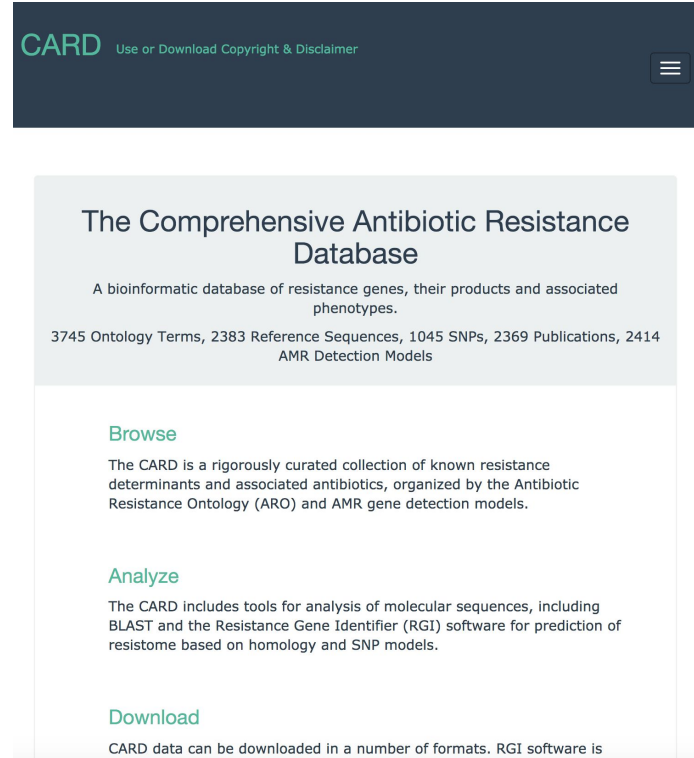
October 3-5, 2017

Outline

- **General Information about CARD**
- Scientific Quality of CARD
- General Use to the Scientific Community
- Summary & Acknowledgments

What is the CARD?

- "The Comprehensive Antibiotic Resistance Database ("CARD") provides data, models, and algorithms relating to the molecular basis of antimicrobial resistance." [1]
- It stores sequences, models, general information, and even provides algorithms for use in analysis of these specimen.

[1] https://card.mcmaster.ca/home

How the CARD is maintained

- The CARD was created by Drs. Gerry Wright and Andrew G. McArthur of McMaster University's Department of Biochemistry & Biomedical Sciences with the help of a global team of collaborators.
- Their research has been supported by funds from various companies:
  - The Canadian Foundation for Innovation
  - The Canadian Institutes of Health Research
  - Natural Sciences and Engineering Research Council of Canada
  - Medical Research Council (UK)
  - Ontario Research Fund
  - Various independent donors

Outline

- General Information about CARD
- **Scientific Quality of CARD**
- General Use to the Scientific Community
- Summary & Acknowledgments

The CARD's utility

- The content contains a detailed enough profile of antimicrobial resistance molecules to be useful to those studying the area given the community involvement.
  - While the database contains a limited amount of information, currently around 3000 entries (13 MB worth of data) in json and various other file formats, the main use of CARD is provided through the easy to use tools for analysis that they include on their website.
- The data in CARD, however, is particularly focused on the field of Antibiotic Research, and does not provide much utility outside of it.

How the CARD maintains its relevance

- The data can be used to perform a multitude of analytics on the various species included in the database, especially using their tools for performing BLAST searches and visualizing RGI information.
- The data is updated on a monthly basis, most recently on September 7th.

Outline

- General Information about CARD
- Scientific Quality of CARD
- **General Use to the Scientific Community**
- Summary & Acknowledgments

The CARD contains links to another Database, and is Simple to Browse

- Throughout the database, links can be found to pages in the NCBI taxonomy database and NCBI Gene

- Browsing the data is convenient. There are two subsections under the "Browse" page:
  - Antibiotic Resistance Ontology: contains data on antibiotic resistance and antibiotics
  - Model Ontology: organizes data into multiple models of resistance

Browsing Under the Antibiotic Resistance Ontology

- Data is organized into terms and subterms.
- Clicking on the link to this ontology brings up a page with the definition of "antibiotic resistance"
  - From there, there are links to more specific terms. For example, "antibiotic molecule", "antibiotic biosynthesis", etc…)
  - The information starts out as broad, but becomes more and more specific the deeper you go into the database
- All terms are related by specific "relationship types":
  - For example the term "antibiotic molecule" is linked to under the term "process or component of antibiotic biology or chemistry" by the relationship term "is_a"
    - So, antibiotic molecule **is a** process or component of antibiotic biology or chemistry
  - Other terms include: "part_of", "derives_from", "confers_resistance_to"

Browsing Under the Model Ontology

- Data is organized into models for how specific antibiotic resistances come about.
- For example, the "protein homolog" is a model of sequences which do not have a mutation that causes resistance. The "protein variant" is a model for sequences which can vary and cause resistance that way.
  - These are both subterms under the "detection model or meta-model page.
- Clicking on the "protein homolog" model brings you to a page with it's definition, from which there are links to more specific terms related by the relationship types used in the Antibiotic Resistance Ontology ("is_a", "part_of")

Downloading the data

- The data is organized for download into these categories:
  - CARD data: separated into each month (every update), all of the data in the database
  - CARD data changes
  - CARD prevalence data: prevalence of each resistance sequence
  - Software
- Data can be downloaded in at least one of these standard file formats:
  - JSON, TAB, OBO, FASTA, OWL, XML, CSV, PYTHON, TXT, GZ

The CARD is Overall User-Friendly

- The website is user friendly, as long as you know what you're looking for. A lot of technical terms are used throughout which makes it hard to understand the information or find new information if you're not experienced in the field.
- The organization makes sense, but it can slightly confusing when using for the first time
  - Data is organized based on its relationship to other data stored in the database
- There's no tutorial or help section
  - There is information on contacting the owners and there is a FAQ section
- There are no search options to narrow a search, instead all results are shown when you type into the search bar
  - From each result there are lists of synonyms, and terms related to that result (other genes, drugs, etc)

Sample Query for an Antibiotic

- I searched for "penicillin" as a test (this is a type of antibiotic).
- The results made sense, I got a list of types of penicillin as well as resistance sequences associated with penicillin
  - Example: Penicillin V is a derivative of penicillin and streptococcus pneumoniae PBP2b is a gene that confers resistance to penicillin

detection parameter

Download Sequences

Accession	MO:0000038
Synonym(s)	
Definition	Model parameters are used within detection models or meta-models. Collectively, parameters describe numerical or other measurable factors that define models, providing the framework for development of analytical tools. For example, many detection models rely on a curated expectation value or bitscore cut-off value to differentiate between functional homologs of known sequences and more divergent homologs with different function.
Classification	
Parent Term(s)	+ detection model or parameter (is_a)
Sub-Term(s)	+ BLASTP e-value (is_a) + taxa restriction (is_a) + BLASTP bit-score (is_a) + BLASTN bit-score (is_a) + single resistance variant (is_a) + multiple resistance variants (is_a) + nonsense mutation (is_a) + co-dependent single resistance variant (is_a) + frameshift mutation (is_a) + co-dependent nonsense SNP (is_a) + co-dependent insertion/deletion (is_a) + gene order (is_a) + gene type (is_a) + efflux pump components (is_a) + domain order (is_a) + domain direction or strand (is_a) + mutation in regulatory element (is_a) + insertion mutation (is_a) + deletion mutation (is_a)

Accessing the CARD

- No membership or license required to view or download the information
- Free for non-commercial, research, or academic use
- Must get written permission and pay a fee to use the information commercially.

Outline

- General Information about CARD
- Scientific Quality of CARD
- General Use to the Scientific Community
- **Summary & Acknowledgments**

Summary

- For someone who is naive in the field of antibiotic resistance, I would not recommend using CARD as a place to start.
- Some effort is required to find a lot of useful information
  - Each entry is very specific, meaning that you have to visit many entries to gather a large amount of information
  - However, within each page there are plenty of links to all related pages, so navigating through the data is tedious, but not difficult
- CARD is organized very well, but still contains very technical terminology that someone not in the field may have trouble understanding

Acknowledgments

Dr. Dahlquist and Dr. Dionisio

The designers and developers of the Comprehensive Antibiotic Resistance Database

References

- Comprehensive Antibiotic Resistance Database. (2017). Retrieved September 29, 2017, from https://card.mcmaster.ca/
- Journals.ASM.org (2017). The Comprehensive Antibiotic Resistance Database. Retrieved October 1, 2017, from http://aac.asm.org/content/57/7/3348.full.pdf
- LMU BioDB 2017. (2017). Week 5. Retrieved September 30, 2017, from https://xmlpipedb.cs.lmu.edu/biodb/fall2017/index.php/Week_5