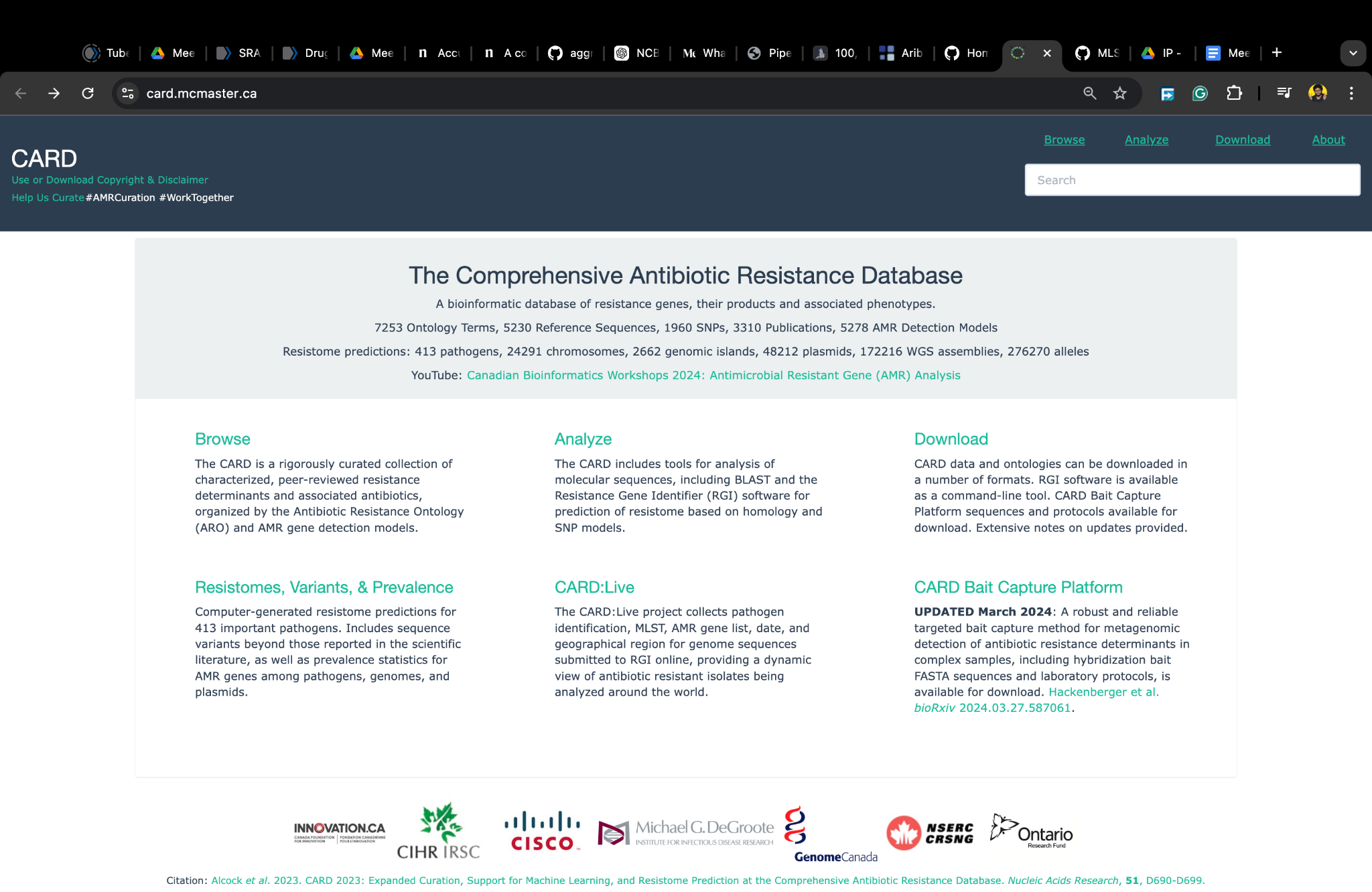
**CARD**



ARIBA

# **ARIBA: Antimicrobial Resistance Identification By Assembly**

ARIBA is a tool that identifies antibiotic resistance genes by running local assemblies. It can also be used for [MLST calling](https://github.com/sanger-pathogens/ariba/wiki/MLST-calling-with-ARIBA).

The input is a FASTA file of reference sequences (can be a mix of genes and noncoding sequences) and paired sequencing reads. ARIBA reports which of the reference sequences were found, plus detailed information on the quality of the assemblies and any variants between the sequencing reads and the reference sequences.

<https://github.com/sanger-pathogens/ariba/wiki>

### **What is Local Assembly?**

In the context of ARIBA, **local assembly** means reconstructing a specific genomic region (such as a gene or a group of genes) from the sequencing reads that map to that region. Instead of assembling the entire genome, ARIBA isolates the parts of the sequencing data that match the reference sequences provided (e.g., resistance genes), and then "assembles" the reads that map to those regions into a continuous sequence.

This allows ARIBA to:

* **Identify specific antibiotic-resistance genes** present in the sequenced sample.
* **Compare the assembled sequence** to the reference sequence to find any differences (such as single nucleotide polymorphisms (SNPs) or insertions/deletions (indels)).

### **Steps in Local Assembly Using ARIBA:**

1. **Reference Sequences as Input**: ARIBA requires a set of reference sequences (e.g., known antibiotic resistance genes) in **FASTA format**. These references represent the regions you are interested in analyzing.
2. **Mapping Reads to the Reference**: Paired-end sequencing reads are aligned to these reference sequences. ARIBA looks for sequencing reads from the sample that correspond to the provided reference genes (e.g., genes known to confer antibiotic resistance).
3. **Local Assembly**:
   * After the reads are mapped to the reference genes, ARIBA performs **local assembly** of these reads. This means that it gathers all reads that match the reference gene and assembles them into a continuous sequence.
   * It reconstructs the local sequence based on overlapping reads in the mapped region, similar to genome assembly but confined to specific parts of the genome (like antibiotic resistance genes).
4. **Variant Detection**:
   * After local assembly, ARIBA compares the assembled sequence to the reference sequence and identifies any **variants** (such as SNPs or indels) between the sample and the reference.
   * This step allows ARIBA to determine not only whether a resistance gene is present but also whether there are any mutations that might affect the gene's function.
5. **Output**: ARIBA produces a detailed report on the:
   * **Presence or absence** of the reference genes.
   * **Variants** detected between the assembled sequence and the reference.
   * **Quality of the assembly**, ensuring the accuracy of the variant calls.

### **Why Use Local Assembly?**

* **Focus on Relevant Regions**: Local assembly allows ARIBA to focus only on the regions of interest (like antibiotic resistance genes) rather than processing the entire genome, making it faster and more efficient.
* **Increased Accuracy**: By assembling reads locally, ARIBA can more accurately reconstruct gene sequences and detect variations that are critical for antibiotic resistance, even if those regions are complex or have mutations.
* **Detailed Variant Information**: Local assembly enables the detection of not just the presence of resistance genes but also the specific mutations that might influence drug resistance, providing more detailed and actionable information.

### **Example:**

Imagine you have sequencing data from a tuberculosis sample, and you want to know if it has a gene that confers resistance to rifampicin (an important anti-TB drug). You provide ARIBA with a FASTA file containing the reference sequence of the **rpoB** gene (which is linked to rifampicin resistance) along with your sequencing reads.

ARIBA:

1. Maps the sequencing reads to the **rpoB** reference sequence.
2. Assembles the reads that map to this region into a continuous **rpoB** gene sequence.
3. Compares the assembled sequence with the reference and detects any mutations in the **rpoB** gene that are known to cause rifampicin resistance.
4. Reports whether the gene is present and if it has any significant variants, such as the **S450L** mutation, which is known to confer resistance to rifampicin.

**BV-BRC**

