

Hur kan man använda
resistensdatabaser för WGS-data?

ØS/CG/USJ

WGS workflow for dummies



Data from machine

- Assembly of fastq files to fasta
(processing raw data into a "nice" file)
- Running fasta files with ResFinder
(detection of resistance genes)
- <https://cge.cbs.dtu.dk/services/>

Center for Genomic Epidemiology

Username
Password

Home

Services

Datasets

Overview of Services

Workflows

[Bacterial Analysis Pipeline \(Batch Upload\)](#)

Phenotyping

[ResFinder](#)
[PathogenFinder](#)
[VirulenceFinder](#)
[Restriction-ModificationFinder](#)

Typing

[MLST](#)
[pMLST](#)
[PlasmidFinder](#)
[KmerFinder](#)
[SpeciesFinder](#)
[Reads2Type](#)
[Tapir](#) (OBS! Only works for Firefox)
[TaxonomyFinder](#)
[SerotypeFinder](#)
[spaTyper](#)
[FimTyper](#)
[PAst](#)

Phylogeny



DA

- [pMLST](#)
- [PlasmidFinder](#)
- [KmerFinder](#)
- [SpeciesFinder](#)
- [Reads2Type](#)
- [Tapir \(OBS! Only works for Firefox\)](#)
- [TaxonomyFinder](#)
- [SerotypeFinder](#)
- [spaTyper](#)
- [FimTyper](#)
- [PAst](#)

Phylogeny

- [snpTree](#)
- [CSI Phylogeny](#)
- [NDtree](#)
- [TreeViewer](#)

Genome Assembly

- ~~[Assembler](#)~~
- [SPAdes](#)

Other

- [PanFunPro](#)
- [MyDbFinder](#)
- [SPiFinder](#)
- [HostPhinder](#)
- [DeHumanizer](#)

Support Technical problems

FASTQ format from Wikipedia

A FASTQ file normally uses four lines per sequence.

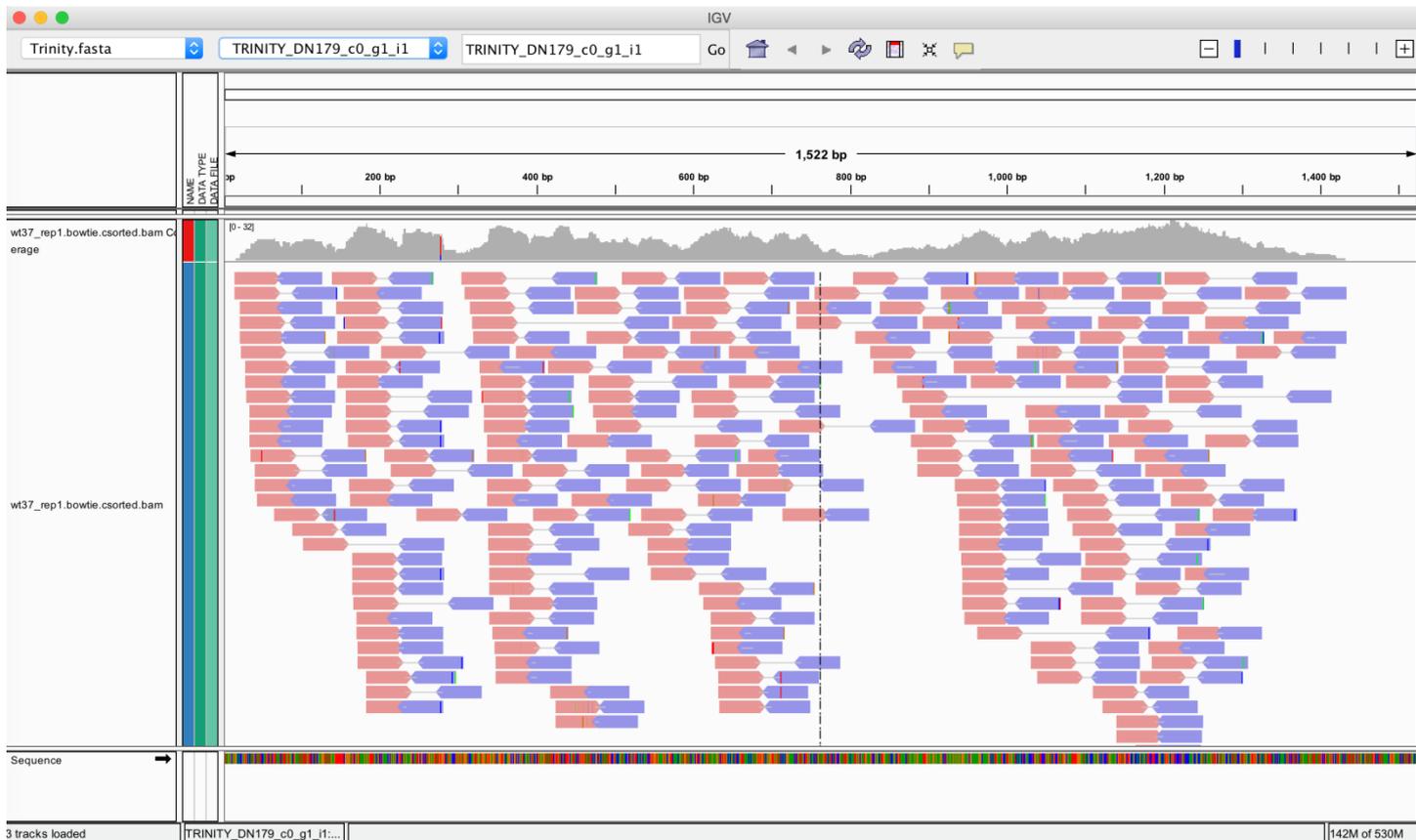
- Line 1 begins with a '@' character and is followed by a sequence identifier and an *optional* description (like a [FASTA](#) title line).
- Line 2 is the raw sequence letters.
- Line 3 begins with a '+' character and is *optionally* followed by the same sequence identifier (and any description) again.
- Line 4 encodes the quality values for the sequence in Line 2, and must contain the same number of symbols as letters in the sequence.

A FASTQ file containing a single sequence might look like this:

```
1.  1@SEQ_ID
2.  GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTTT
3.  +
4.  !"*((((***+))%%%++)(%%%).1***-+*")**55CCF>>>>>CCCCCCC65
```

Genome assembly

”Putting the small pieces together”



Mail - Ulrik Justesen - Outlook CGE Server

Sikker | <https://cge.cbs.dtu.dk/services/SPAdes/>

Outlook Web App SD Outlook Web App OU Gå til forsiden af dr.dk Ulrik Stenz Justesen - TV CGE Services Dashboard - BaseSpa Expert Consult Library TDC hotspot TDC: Mobiltele

Center for Genomic Epidemiology

Username
Password
New Reset Login

Home Services Instructions Output Article abstract

SPAdes 3.9

View the [version history](#) of this server.

Sequencing Platform
Select the sequencing platform used to generate the uploaded reads.
Illumina - paired end reads

Choose careful-mode:
Tries to reduce the number of mismatches and short indels. Also runs MismatchCorrector - a post processing tool, which uses BWA tool. This option is recommended.
Set careful on:

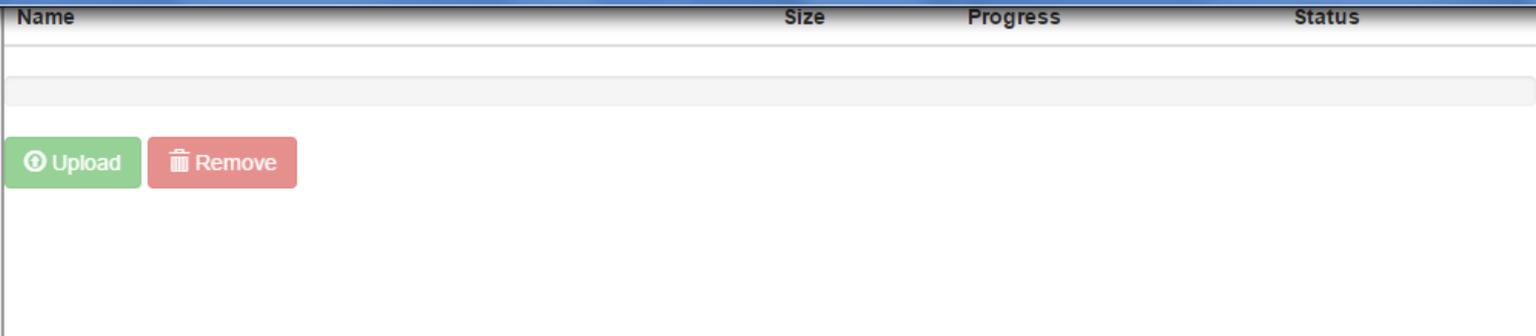
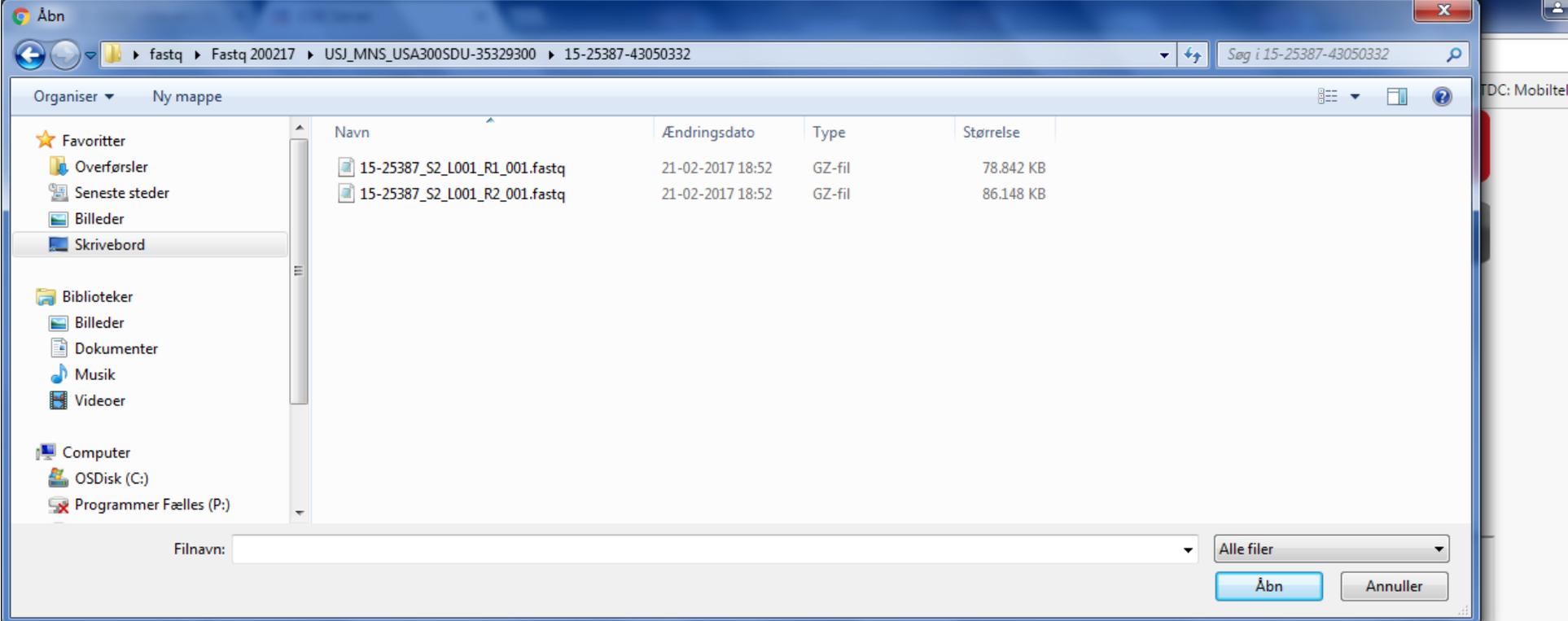
Isolate File

Name	Size	Progress	Status
------	------	----------	--------

Upload Remove

Windows Taskbar: Chrome, File Explorer, Word

Spades: **St. Petersburg** genome assembler



Center for Genomic Epidemiology

Username
Password

- Home
- Services
- Instructions
- Output
- Article abstract

SPAdes 3.9

View the [version history](#) of this server.

Sequencing Platform

Select the sequencing platform used to generate the uploaded reads.

Illumina - paired end reads

Choose careful-mode:

Tries to reduce the number of mismatches and short indels. Also runs MismatchCorrector - a post processing tool, which uses BWA tool. This option is recommended.

Set careful on:

Name	Size	Progress	Status
15-25387_S2_L001_R1_001.fastq.gz	76.99 MB	<div style="width: 50%;"></div>	
15-25387_S2_L001_R2_001.fastq.gz	84.13 MB	<div style="width: 0%;"></div>	

Center for Genomic Epidemiology

Username
Password

Home

Services

Instructions

Output

Article abstract

SPAdes 3.9

View the [version history](#) of this server.

Sequencing Platform

Select the sequencing platform used to generate the uploaded reads.

Illumina - paired end reads

Choose careful-mode:

Tries to reduce the number of mismatches and short indels. Also runs MismatchCorrector - a post processing tool, which uses BWA tool. This option is recommended.

Set careful on:

Isolate File

Name	Size	Progress	Status
15-25387_S2_L001_R1_001.fastq.gz	76.99 MB	<div style="width: 100%;"></div>	✓
15-25387_S2_L001_R2_001.fastq.gz	84.13 MB	<div style="width: 100%;"></div>	

Upload

Remove

Center for Genomic Epidemiology

Your job is being processed

Wait here to watch the progress of your job, or fill in the form below to get an email message upon completion.

To get notified by email:

This page will update itself automatically.



Center for Genomic Epidemiology

- Home
- Services
- Instructions
- Output

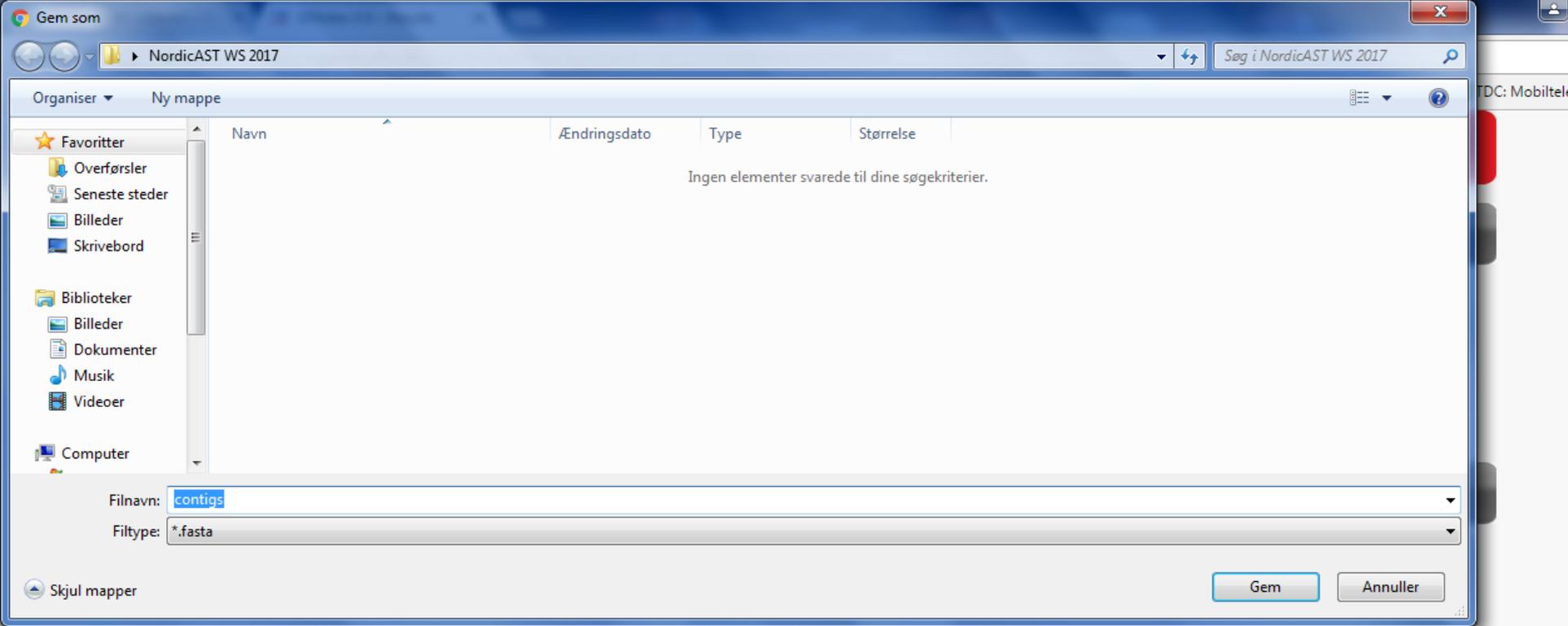
SPAdes-3.9 Server - Results

Input Files: *Funn2_S4_L001_R1_001.fastq Funn2_S4_L001_R2_001.fastq*

Contigs

- Support
- Scientific problems
- Technical problems

Copyright DTU 2011 / All rights reserved
Center for Genomic Epidemiology, DTU, Kemitorvet, Building 204, 2800 Kgs. Lyngby, Denmark
Contact: Vibeke Dybdaht Hammer, Telephone: +45 3588 6420, E-mail: vdha@food.dtu.dk
Funded by: The Danish Council for Strategic Research
Last modified May 22, 2012 11:08:01 GMT



Fasta file with contigs

>NODE_1_length_291111_cov_18.0436

TGGATACCTCACGCCAAACGAAAAATTTAAACAAATTATTAATCAGAATTCTGTTGCATT
TGCAAGTTGAATTCAGCTAATTTTAAGACCTACCTGTCGCTTGATCTTGATCTGACGGTA
GATGATGACCGCCCGATTATACTTATCGGCGGTTCTAATGGCGGTGGAAAGACAACCTCTG ...

>NODE_2_length_209169_cov_16.4241

TGCGGTGCGTACGGGACTCGAACCCGTGACCCCATGCGTGACAGGCATGTATTCTAACCA
ACTGAACTAACGCACCAGAATTTCATTTTTAATCGCTATCTCTCTCGATTGCGGTTGCA
AAGGTAGATATTTATTTTATATCTGCAATAGCTAAGAAAGATTTTTTTTCATCTTTCTTA ...

1571 pages later:

>NODE_230_length_78_cov_30

TTGAGCGGAAAACGAGACTCGAACTCGCGACCCTAACCTTGGCAAGGTTATGCTCTACCA
ACTGAGCTATTTCCGCAA ...

Center for Genomic Epidemiology

Username
Password

Home

Services

Datasets

Overview of Services

Workflows

[Bacterial Analysis Pipeline \(Batch Upload\)](#)

Phenotyping

[ResFinder](#)

[PathogenFinder](#)

[VirulenceFinder](#)

[Restriction-ModificationFinder](#)

Typing

[MLST](#)

[pMLST](#)

[PlasmidFinder](#)

[KmerFinder](#)

[SpeciesFinder](#)

[Reads2Type](#)

[Tapir](#) (OBS! Only works for Firefox)

[TaxonomyFinder](#)

[SerotypeFinder](#)

[spaTyper](#)

[FimTyper](#)

[PAst](#)

Phylogeny



DA

Center for Genomic Epidemiology

Username
Password

Home

Services

Instructions

Output

Overview of genes

Article abstract

ResFinder 2.1

ResFinder identifies acquired antimicrobial resistance genes in total or partial sequenced isolates of bacteria.

Fasta file with test sequence: [Test sequence](#)

NOTE: Currently ResFinder focuses on acquired genes and does therefore not find chromosomal mutations (NAL, FUS, high-level CIP, RIF resistance, etc.)

The database is curated by:
Valeria Bortolaia
(click to contact)

View the [version history](#) of this server.

Select Antimicrobial configuration

Select multiple items, with Ctrl-Click (or Cmd-Click on Mac) - by default all databases are selected

Aminoglycoside
Beta-lactam
Colistin
Fluoroquinolone
Fosfomycin
Fusidic Acid

Select threshold for %ID

90 %

Select minimum length

60 %

Select type of your reads

Assembled Genome/Contigs*

- Select threshold for %ID down to 30%
- Select minimum length down to 20%

Fusidic Acid

Select threshold for %ID
90 %

Select minimum length
60 %

Select type of your reads
Assembled Genome/Contigs*

If you get an "Access forbidden. Error 403": Make sure the start of the web address is https and not just http. Fix it by clicking [here](#).

Isolate File

Name	Size	Progress	Status
------	------	----------	--------

Upload Remove

Confidentiality:
The sequences are kept confidential and will be deleted after 48 hours.

What you see is what you get!

Maybe you can find more in the data, if you have the right skills and tools???



Center for Genomic Epidemiology

Home Services Instructions Output Overview of genes Article abstract

ResFinder-2.1 Server - Results

Aminoglycoside						
Resistance gene	%Identity	HSP length/Query	Contig	Position in contig	Predicted phenotype	Accession number
<i>strB</i>	100.00	837 / 837	NODE_156_length_6288_cov_24.0689_ID_335	2198..3034	Aminoglycoside resistance Alternate name; aph(6)-IId	M96392
<i>strA</i>	100.00	804 / 804	NODE_156_length_6288_cov_24.0689_ID_335	3034..3837	Aminoglycoside resistance Alternate name; aph(3'')-IId	AF3215
<i>aac(3)-IId</i>	99.77	861 / 861	NODE_161_length_5909_cov_3.99949_ID_341	4857..5717	Aminoglycoside resistance	EU0223
<i>aadA2</i>	100.00	792 / 792	NODE_195_length_3200_cov_3.89625_ID_403	923..1714	Aminoglycoside resistance	JQ3649

Beta-lactam						
Resistance gene	%Identity	HSP length/Query	Contig	Position in contig	Predicted phenotype	Accession number
<i>blaOXA-48</i>	100.00	798 / 798	NODE_200_length_2866_cov_30.7859_ID_413	1614..2411	Beta-lactam resistance	AY236073
<i>blaTEM-1B</i>	100.00	861 / 861	NODE_132_length_9090_cov_45.3519_ID_283	193..1053	Beta-lactam resistance Alternate name; RblaTEM-1	JF910132

Colistin

No resistance genes found.

Fluoroquinolone

No resistance genes found.

Fosfomycin

No resistance genes found.

Fusidic Acid

No resistance genes found.

Glycopeptide

No resistance genes found.

Trimethoprim						
Resistance gene	%Identity	HSP length/Query	Contig	Position in contig	Predicted phenotype	Accession number
<i>dfra12</i>	100.00	498 / 498	NODE_195_length_3200_cov_3.89625_ID_403	2122..2619	Trimethoprim resistance	AB571791

extended output

Results as text Results tab separated Hit in genome sequences Resistance gene sequences

Selected %ID threshold: 90.00 %

Selected minimum length: 60 %

Input Files: 14-73932.fasta

CITATIONS

For publication of results, please cite:

- Identification of acquired antimicrobial resistance genes. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. J Antimicrob Chemother. 2012 Jul 10. PMID: [22782487](#) doi: [10.1093/jac/dks261](#)

Support Scientific problems Technical problems

Copyright DTU 2011 / All rights reserved
Center for Genomic Epidemiology, DTU, Kemitorvet, Building 204, 2800 Kgs. Lyngby, Denmark
Contact: Vibeke Dybdahl Hammer, Telephone: +45 3588 6420, E-mail: vdha@food.dtu.dk
Funded by: The Danish Council for Strategic Research
Last modified May 22, 2012 11:08:01 GMT



Telia Wifi x Mail - Ulrik Justesen - Outlook x The Comprehensive Antil... x

McMaster University [CA] | https://card.mcmaster.ca

Outlook Web App SD Outlook Web App OU Gå til forsiden af dr.dk Ulrik Stenz Justesen - TV CGE Services Dashboard - BaseSpa Expert Consult Library TDC hotspot TDC: Mobiltelefoni, br

[Browse](#) [Analyze](#) [Download](#) [About](#)

CARD

[Use or Download Copyright & Disclaimer](#)

The Comprehensive Antibiotic Resistance Database

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

3670 Ontology Terms, 2353 Reference Sequences, 944 SNPs, 2283 Publications, 2356 AMR Detection Models

Browse

The CARD is a rigorously curated collection of known 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 can be downloaded in a number of formats. RGI software is available as a command-line tool. View changes in CARD curation. CARD is updated monthly.

Timeline

- Julian Garneau Retweeted
-  **CARD Developers** @arpcard
CARD 1.1.8 & RGI 3.2.0 #card_release: bitscores, chloramphenicol exporters, ontology changes, Pradinal gene collig

https://card.mcmaster.ca/

Analyze

BLAST

Perform standard BLAST searches against the CARD reference sequences. Results are annotated with extra information from the CARD.

RGI

Use the Resistance Gene Identifier to predict resistome(s) from protein or nucleotide data based on homology and SNP models. Now supporting bulk analyses.

[Visualize external RGI json results](#)



RGI Resistance Gene Identifier

RGI Bulk Analysis: Do you want to use RGI to analyze a large number of genomes? It is now available as a downloadable command-line tool in the [Download section of the CARD website](#).

The CARD includes a curated collection of antimicrobial resistance gene and mutation sequences, bioinformatics models for their detection, and software in the form of the Resistance Gene Identifier (RGI) for their detection in genome or protein sequences. It includes a common framework for the sharing of antimicrobial resistance data in the form of the novel Antibiotic Resistance Ontology (ARO), allowing the RGI to catalog resistome predictions by both drug class and resistance mechanism.

The RGI provides a preliminary annotation of your DNA or protein sequence(s), based upon the data available in CARD. The RGI currently only analyzes protein sequences - if genome sequences or assembly contigs are submitted, the RGI first predicts open reading frames using the [Prodigal](#) software (ignoring those less than 30 bp) and analyzes the predicted protein sequences. The RGI currently supports two detection model types - *protein homolog models* (use of BLAST sequence similarity cut-offs to detection functional homologs of AMR genes) and *protein variant models* (for accurate differentiation between susceptible intrinsic genes and intrinsic genes that have acquired mutations conferring AMR, based on curated SNP matrices).



Use the RGI:

Enter a GenBank accession for a nucleotide or protein sequence:

You can enter multiple accessions separated by commas

Nucleotide sequences will undergo ORF calling to generate predicted protein sequences. Short or partial gene sequences are unlikely to work. Examples: JN420336.1, AY123251.1, HQ451074.1, AL123456

Upload sequence(s):

Vælg filer Der er ikke valgt nogen fil

Upload a **plain text file** containing DNA or protein sequence(s) in FASTA format (20 Mb limit). The file can contain more than one FASTA formatted sequence, such as assembly contigs or multiple proteins. Each file will be treated as a single sample.

Select data type for uploaded file(s):

DEFAULT - DNA sequence

Select Criteria:

DEFAULT - Perfect and Strict hits only

Home All Results

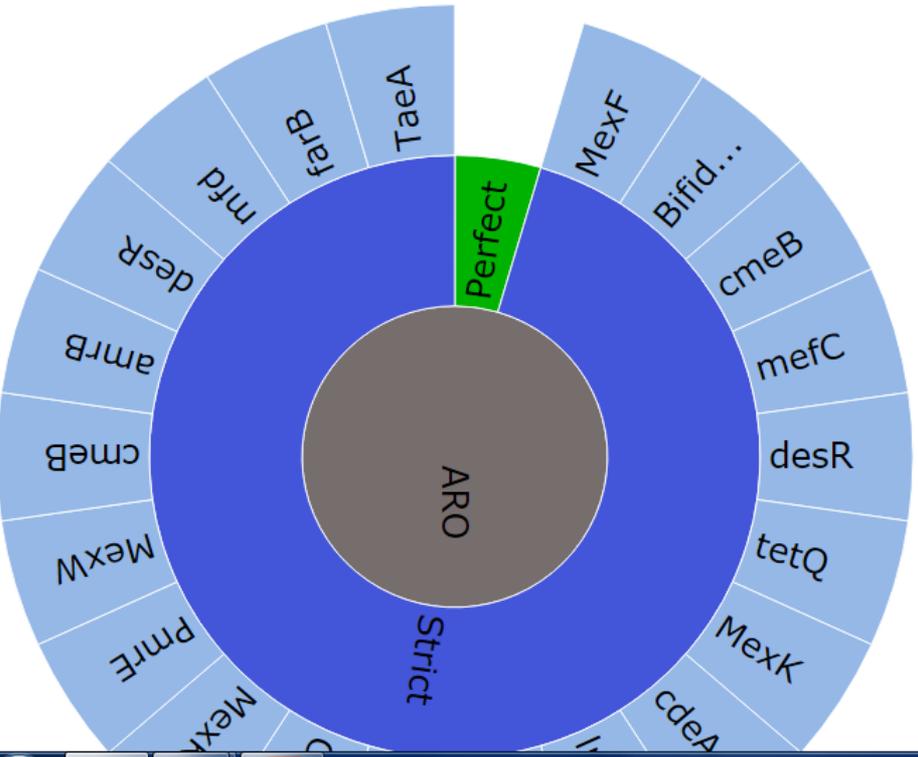
Resistance Gene Identifier (RGI) Results

Loose hits of **e-10** or better can be visualized in the "View by Functional Categories" display. Loose hits are not shown in the "View by Gene" display.

Filename	Date (UTC)	Criteria	Visualize	All Results (JSON)	Tab-delimited Summary	FASTA Format	Heat Map PDF
DCMOUH BF-18.fa	May 22, 2017 10:48:32	Perfect, Strict	View by Gene View by Functional Categories	Download	Download	Download	Download

Home All Results ecd2a0f3e7d628de87fc3d186518f318-PS x

RGI Results (Click on the RGI wheel to browse results)



Tools

Download