Center for Genomic Epidemiology
MyDbFinder
Tutorial

JOANA MOURÃO

EURL-AR

Objectives

Objective 1: Understand the purpose of MyDbFinder in bacterial genomics

Objective 2: Learn how to create a custom database of relevant MRSA genes

Objective 3: Perform a step-by-step analysis using MyDbFinder

Objective 4: Analyze the results and understand their significance

Introduction to MyDbFinder

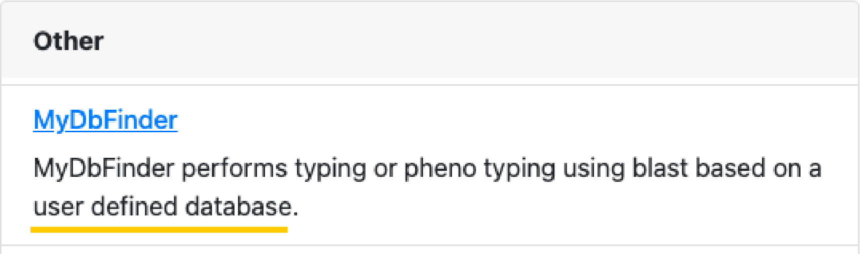
What is MyDbFinder?

- A tool that allows for the identification of genes from a **custom database** in bacterial whole-genome sequences

Why is it useful?

- Essential for tracking antibiotic resistance, virulence genes, and other relevant genes important for understanding genomic epidemiology

MyDbFinder Location



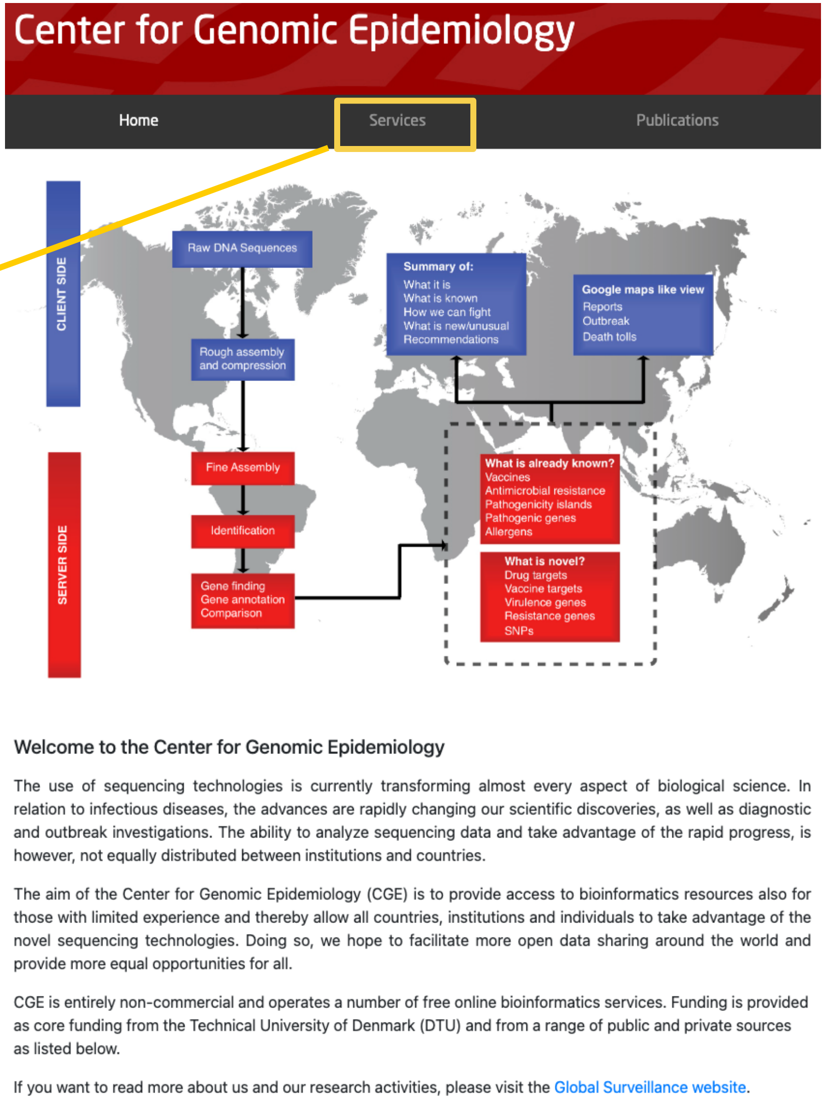
Other

[MyDbFinder](#)

MyDbFinder performs typing or pheno typing using blast based on a user defined database.

- **Create** and **use** custom gene databases ➡️ specific research goals
- Operates through a **simple web interface** ➡️ accessible to researchers without extensive bioinformatics skills

https://www.genomicepidemiology.org/



Welcome to the Center for Genomic Epidemiology

The use of sequencing technologies is currently transforming almost every aspect of biological science. In relation to infectious diseases, the advances are rapidly changing our scientific discoveries, as well as diagnostic and outbreak investigations. The ability to analyze sequencing data and take advantage of the rapid progress, is however, not equally distributed between institutions and countries.

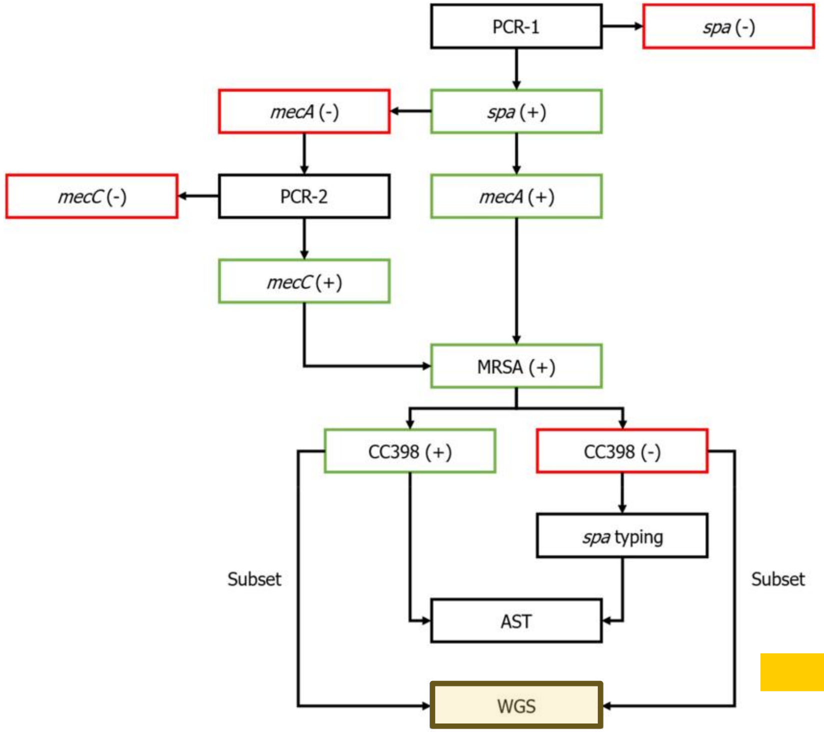
The aim of the Center for Genomic Epidemiology (CGE) is to provide access to bioinformatics resources also for those with limited experience and thereby allow all countries, institutions and individuals to take advantage of the novel sequencing technologies. Doing so, we hope to facilitate more open data sharing around the world and provide more equal opportunities for all.

CGE is entirely non-commercial and operates a number of free online bioinformatics services. Funding is provided as core funding from the Technical University of Denmark (DTU) and from a range of public and private sources as listed below.

If you want to read more about us and our research activities, please visit the Global Surveillance website.

Create a Good Database

Development of a **custom database** with relevant genes in preparation for the EU-wide baseline survey on the prevalence of methicillin-resistant *Staphylococcus aureus* (MRSA) in pigs



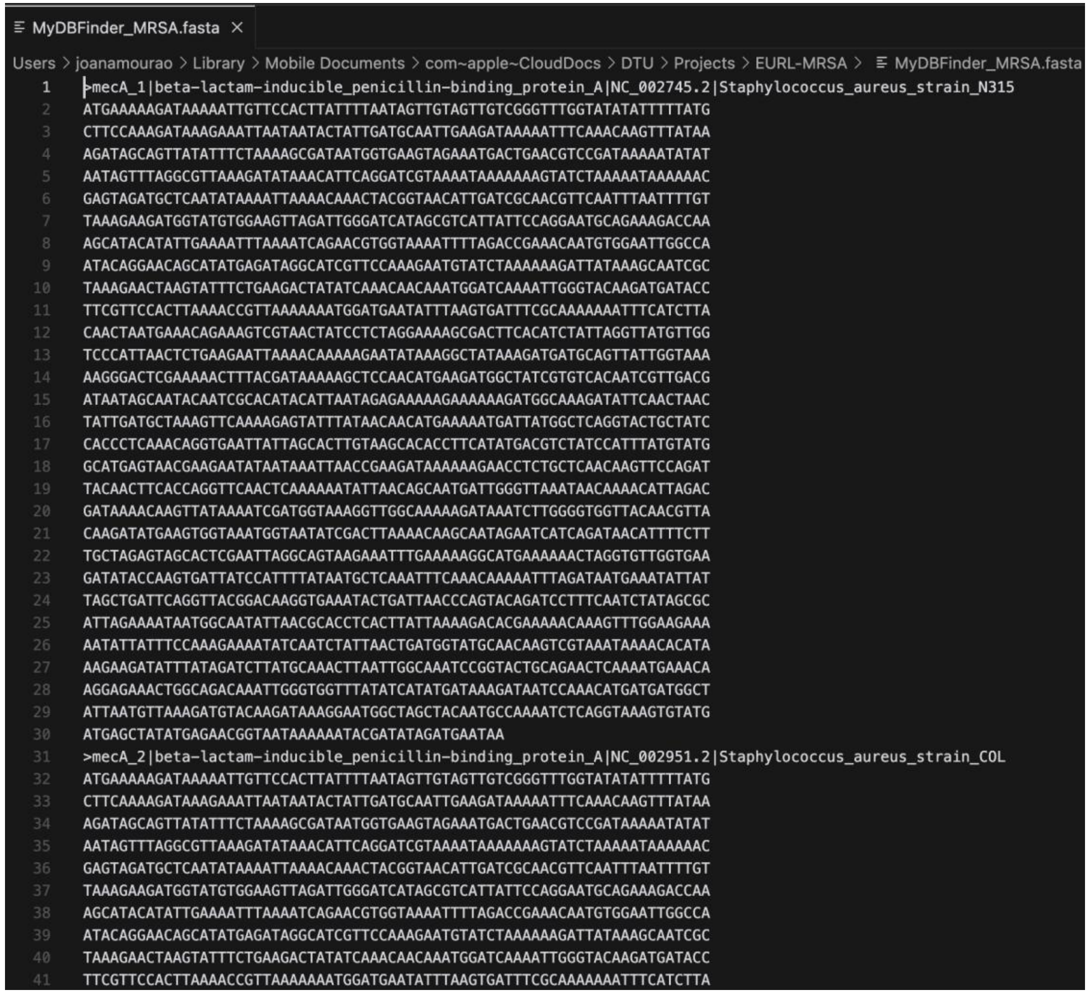
The database includes genes used to classify *S. aureus* as MRSA or MSSA (*mecA*/*mecC*), genes associated with human or swine hosts (PVL, *scn*, and CC398), and other relevant genes (e.g., part of the IEC-1)

Create a Good Database

A. Collect relevant **gene sequences** for your database

B. Format the gene sequences for MyDbFinder - ensure the sequences are in a **DNA FASTA format** and named clearly for ease of identification

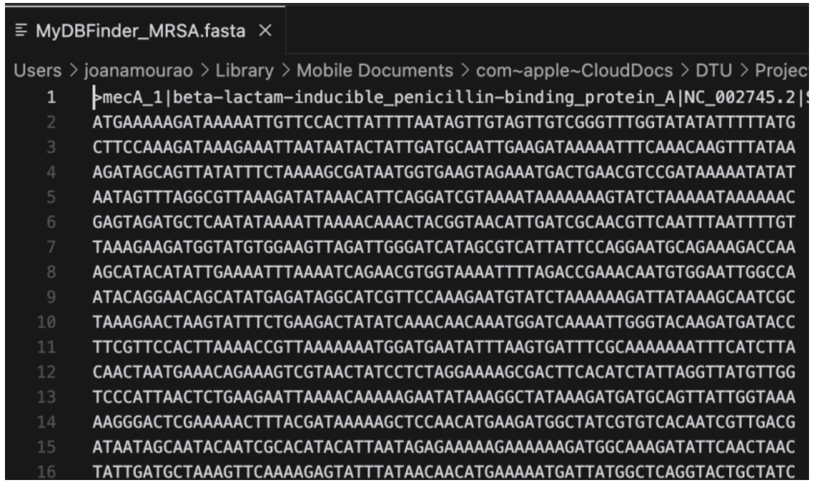
- Gene **name**/variant
- Gene **description**
- **Accession** number
- **Isolate** genus/specie and **name**



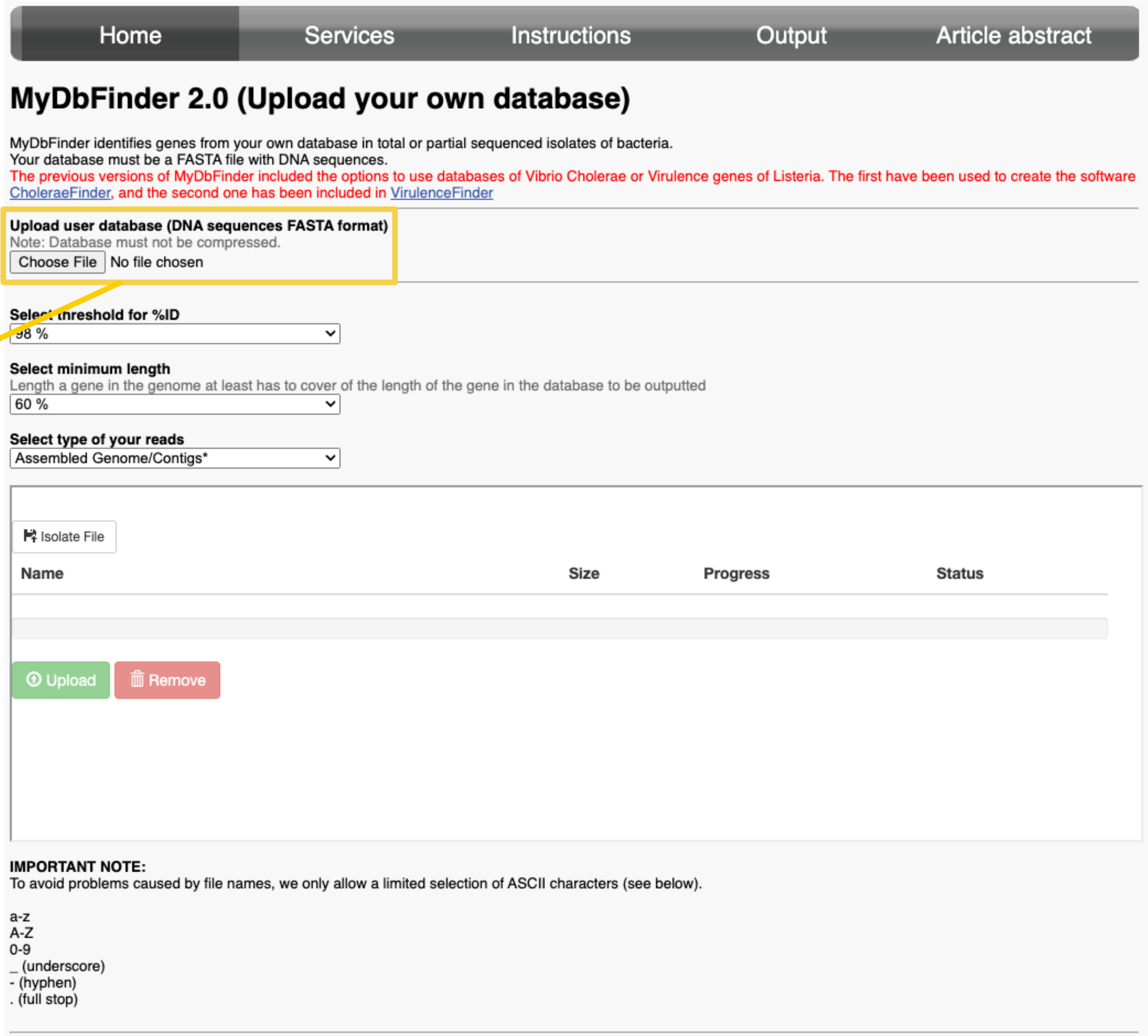
```
>mecA_1|beta-lactam-inducible_penicillin-binding_protein_A|NC_002745.2|Staphylococcus_aureus_strain_N315
ATGAAAAAGATAAAAATTGTTCCACTTATTTTAATAGTTGTAGTTGTCGGGTTTGGTATATATTTTTATG
CTTCCAAAGATAAAGAAATTAATAATACTATTGATGCAATTGAAGATAAAAATTTCAAACAAGTTTATAA
AGATAGCAGTTATATTTCTAAAAGCGATAATGGTGAAGTAGAAATGACTGAACGTCCGATAAAAATATAT
AATAGTTTAGGCGTTAAAGATATAAACATTCAGGATCGTAAAATAAAAAAAGTATCTAAAAATAAAAAAC
GAGTAGATGCTCAATATAAAATTAAAACAAACTACGGTAACATTGATCGCAACGTTCAATTTAATTTTGT
TAAAGAAGATGGTATGTGGAAGTTAGATTGGGATCATAGCGTCATTATTCCAGGAATGCAGAAAGACCAA
AGCATACATATTGAAAATTTAAAATCAGAACGTGGTAAAATTTTAGACCGAAACAATGTGGAATTGGCCA
ATACAGGAACAGCATATGAGATAGGCATCGTTCCAAAGAATGTATCTAAAAAAGATTATAAAGCAATCGC
TAAAGAACTAAGTATTTCTGAAGACTATATCAAACAACAAATGGATCAAAATTGGGTACAAGATGATACC
TTCGTTCCACTTAAAACCGTTAAAAAAATGGATGAATATTTAAGTGATTTCGCAAAAAAATTTCATCTTA
CAACTAATGAAACAGAAAGTCGTAACTATCCTCTAGGAAAAGCGACTTCACATCTATTAGGTTATGTTGG
TCCCATTAACTCTGAAGAATTAAAACAAAAAGAATATAAAGGCTATAAAGATGATGCAGTTATTGGTAAA
AAGGGACTCGAAAAACTTTACGATAAAAAGCTCCAACATGAAGATGGCTATCGTGTCACAATCGTTGACG
ATAATAGCAATACAATCGCACATACATTAATAGAGAAAAAGAAAAAAGATGGCAAAGATATTCAACTAAC
TATTGATGCTAAAGTTCAAAAGAGTATTTATAACAACATGAAAAATGATTATGGCTCAGGTACTGCTATC
CACCCTCAAACAGGTGAATTATTAGCACTTGTAAGCACACCTTCATATGACGTCTATCCATTTATGTATG
GCATGAGTAACGAAGAATATAATAAATTAACCGAAGATAAAAAAGAACCTCTGCTCAACAAGTTCCAGAT
TACAACTTCACCAGGTTCAACTCAAAAAATATTAACAGCAATGATTGGGTTAAATAACAAAACATTAGAC
GATAAAACAAGTTATAAAATCGATGGTAAAGGTTGGCAAAAAGATAAATCTTGGGGTGGTTACAACGTTA
CAAGATATGAAGTGGTAAATGGTAATATCGACTTAAAACAAGCAATAGAATCATCAGATAACATTTTCTT
TGCTAGAGTAGCACTCGAATTAGGCAGTAAGAAATTTGAAAAAGGCATGAAAAAACTAGGTGTTGGTGAA
GATATACCAAGTGATTATCCATTTTATAATGCTCAAATTTCAAACAAAAATTTAGATAATGAAATATTAT
TAGCTGATTCAGGTTACGGACAAGGTGAAATACTGATTAACCCAGTACAGATCCTTTCAATCTATAGCGC
ATTAGAAAATAATGGCAATATTAACGCACCTCACTTATTAAAAGACACGAAAAACAAAGTTTGGAAGAAA
AATATTATTTCCAAAGAAAATATCAATCTATTAACTGATGGTATGCAACAAGTCGTAAATAAAACACATA
AAGAAGATATTTATAGATCTTATGCAAACTTAATTGGCAAATCCGGTACTGCAGAACTCAAAATGAAACA
AGGAGAAACTGGCAGACAAATTGGGTGGTTTATATCATATGATAAAGATAATCCAAACATGATGATGGCT
ATTAATGTTAAAGATGTACAAGATAAAGGAATGGCTAGCTACAATGCCAAAATCTCAGGTAAAGTGTATG
ATGAGCTATATGAGAACGGTAATAAAAAATACGATATAGATGAATAA
>mecA_2|beta-lactam-inducible_penicillin-binding_protein_A|NC_002951.2|Staphylococcus_aureus_strain_COL
ATGAAAAAGATAAAAATTGTTCCACTTATTTTAATAGTTGTAGTTGTCGGGTTTGGTATATATTTTTATG
CTTCAAAAGATAAAGAAATTAATAATACTATTGATGCAATTGAAGATAAAAATTTCAAACAAGTTTATAA
AGATAGCAGTTATATTTCTAAAAGCGATAATGGTGAAGTAGAAATGACTGAACGTCCGATAAAAATATAT
AATAGTTTAGGCGTTAAAGATATAAACATTCAGGATCGTAAAATAAAAAAAGTATCTAAAAATAAAAAAC
GAGTAGATGCTCAATATAAAATTAAAACAAACTACGGTAACATTGATCGCAACGTTCAATTTAATTTTGT
TAAAGAAGATGGTATGTGGAAGTTAGATTGGGATCATAGCGTCATTATTCCAGGAATGCAGAAAGACCAA
AGCATACATATTGAAAATTTAAAATCAGAACGTGGTAAAATTTTAGACCGAAACAATGTGGAATTGGCCA
ATACAGGAACAGCATATGAGATAGGCATCGTTCCAAAGAATGTATCTAAAAAAGATTATAAAGCAATCGC
TAAAGAACTAAGTATTTCTGAAGACTATATCAAACAACAAATGGATCAAAATTGGGTACAAGATGATACC
TTCGTTCCACTTAAAACCGTTAAAAAAATGGATGAATATTTAAGTGATTTCGCAAAAAAATTTCATCTTA
```

Running MyDbFinder

1. Upload the custom **MRSA gene database** you've previously created



```
>mecA_1|beta-lactam-inducible_penicillin-binding_protein_A|NC_002745.2|
ATGAAAAAGATAAAAATTGTTCCACTTATTTTAATAGTTGTAGTTGTCGGGTTTGGTATATATTTTTATG
CTTCCAAAGATAAAGAAATTAATAATACTATTGATGCAATTGAAGATAAAAATTTCAAACAAGTTTATAA
AGATAGCAGTTATATTTCTAAAAGCGATAATGGTGAAGTAGAAATGACTGAACGTCCGATAAAAATATAT
AATAGTTTAGGCGTTAAAGATATAAACATTCAGGATCGTAAAATAAAAAAAGTATCTAAAAATAAAAAAC
GAGTAGATGCTCAATATAAAATTAAAACAAACTACGGTAACATTGATCGCAACGTTCAATTTAATTTTGT
TAAAGAAGATGGTATGTGGAAGTTAGATTGGGATCATAGCGTCATTATTCCAGGAATGCAGAAAGACCAA
AGCATACATATTGAAAATTTAAAATCAGAACGTGGTAAAATTTTAGACCGAAACAATGTGGAATTGGCCA
ATACAGGAACAGCATATGAGATAGGCATCGTTCCAAAGAATGTATCTAAAAAAGATTATAAAGCAATCGC
TAAAGAACTAAGTATTTCTGAAGACTATATCAAACAACAAATGGATCAAAATTGGGTACAAGATGATACC
TTCGTTCCACTTAAAACCGTTAAAAAAATGGATGAATATTTAAGTGATTTCGCAAAAAAATTTCATCTTA
CAACTAATGAAACAGAAAGTCGTAACTATCCTCTAGGAAAAGCGACTTCACATCTATTAGGTTATGTTGG
TCCCATTAACTCTGAAGAATTAAAACAAAAAGAATATAAAGGCTATAAAGATGATGCAGTTATTGGTAAA
AAGGGACTCGAAAAACTTTACGATAAAAAGCTCCAACATGAAGATGGCTATCGTGTCACAATCGTTGACG
ATAATAGCAATACAATCGCACATACATTAATAGAGAAAAAGAAAAAAGATGGCAAAGATATTCAACTAAC
TATTGATGCTAAAGTTCAAAAGAGTATTTATAACAACATGAAAAATGATTATGGCTCAGGTACTGCTATC
```



MyDbFinder 2.0 (Upload your own database)

MyDbFinder identifies genes from your own database in total or partial sequenced isolates of bacteria.
Your database must be a FASTA file with DNA sequences.
The previous versions of MyDbFinder included the options to use databases of Vibrio Cholerae or Virulence genes of Listeria. The first have been used to create the software CholeraeFinder, and the second one has been included in VirulenceFinder

Upload user database (DNA sequences FASTA format)
Note: Database must not be compressed.
Choose File No file chosen

Select threshold for %ID
98 %

Select minimum length
Length a gene in the genome at least has to cover of the length of the gene in the database to be outputted
60 %

Select type of your reads
Assembled Genome/Contigs*

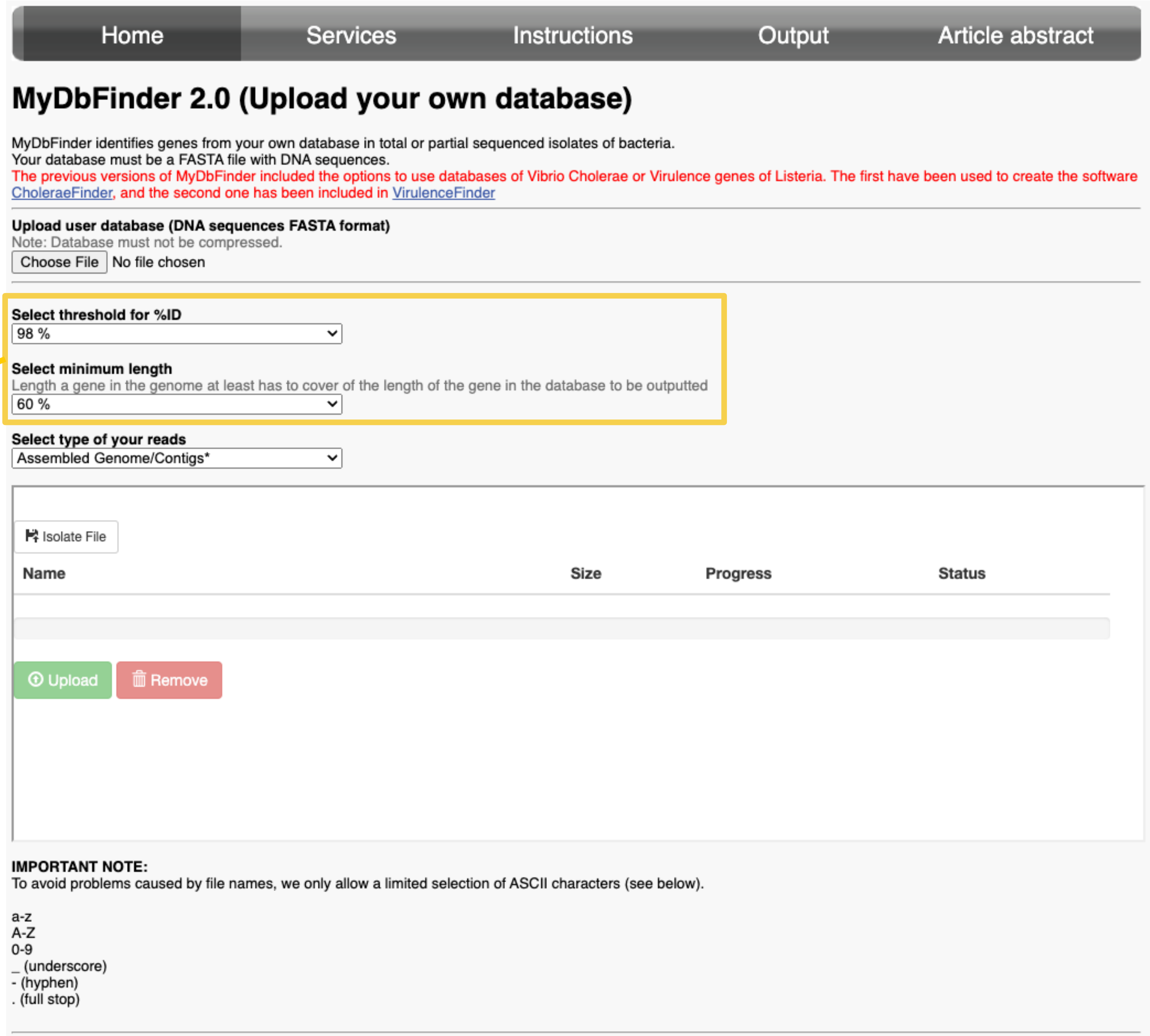
IMPORTANT NOTE:
To avoid problems caused by file names, we only allow a limited selection of ASCII characters (see below).

a-z
A-Z
0-9
_ (underscore)
- (hyphen)
. (full stop)

Running MyDbFinder

2. Customize the search **parameters** (e.g., **identity**, **coverage**)

- **Sequence Identity:** ≥90-95% ➡️ lower identity may yield non-specific hits
- **Coverage:** set the minimum length of 60-80% to ensure that most of the gene's sequence is covered



MyDbFinder 2.0 (Upload your own database)

MyDbFinder identifies genes from your own database in total or partial sequenced isolates of bacteria.
Your database must be a FASTA file with DNA sequences.
The previous versions of MyDbFinder included the options to use databases of Vibrio Cholerae or Virulence genes of Listeria. The first have been used to create the software CholeraeFinder, and the second one has been included in VirulenceFinder

Upload user database (DNA sequences FASTA format)
Note: Database must not be compressed.
No file chosen

Select threshold for %ID
98 %

Select minimum length
Length a gene in the genome at least has to cover of the length of the gene in the database to be outputted
60 %

Select type of your reads
Assembled Genome/Contigs*

Name	Size	Progress	Status

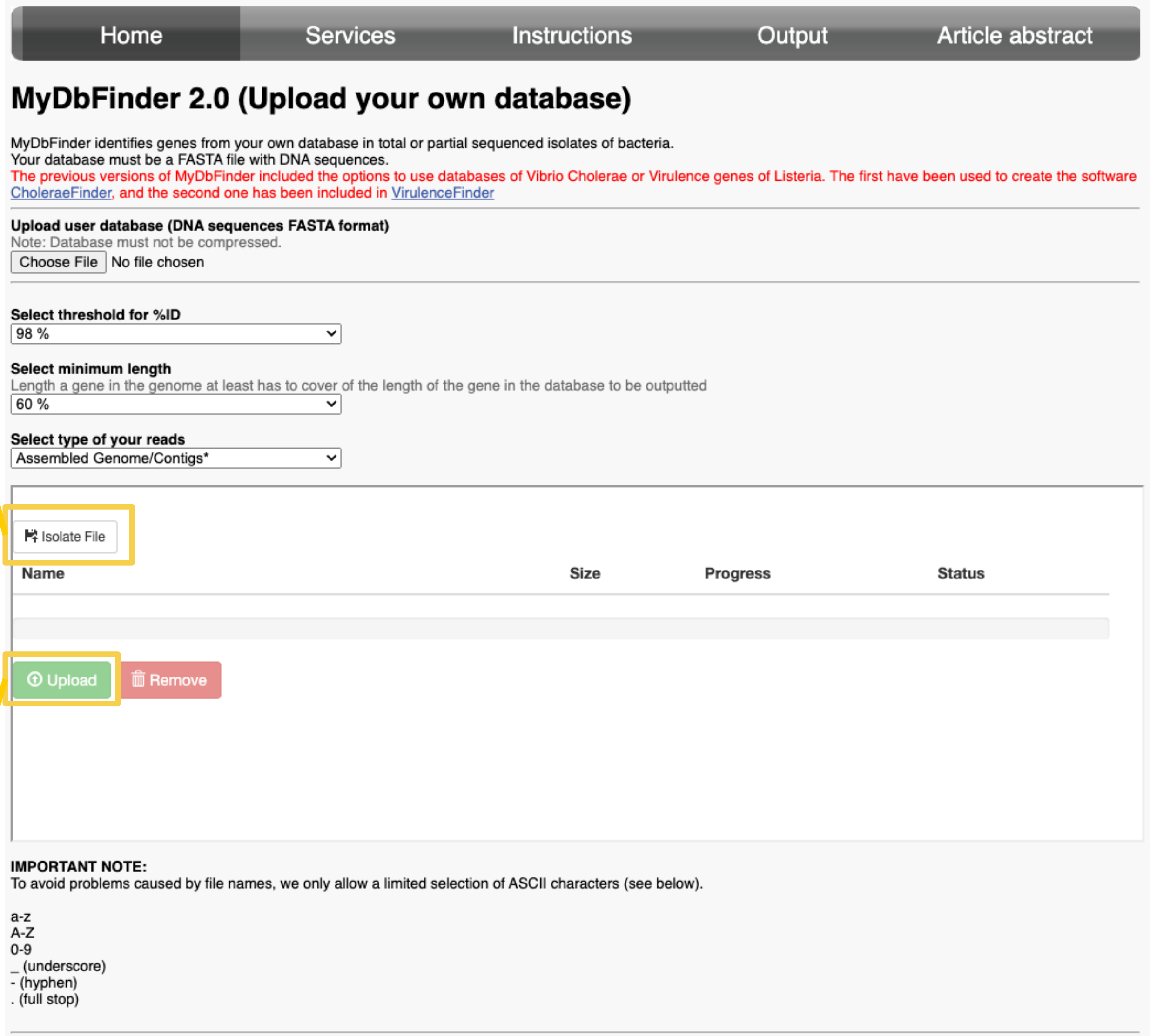
IMPORTANT NOTE:
To avoid problems caused by file names, we only allow a limited selection of ASCII characters (see below).

a-z
A-Z
0-9
_ (underscore)
- (hyphen)
. (full stop)

Running MyDbFinder

3. Upload the genome(s) you want to analyze (e.g., raw data, assembled genomes in FASTA format)

Click **”Upload"** and wait for the analysis to complete



MyDbFinder 2.0 (Upload your own database)

MyDbFinder identifies genes from your own database in total or partial sequenced isolates of bacteria.
Your database must be a FASTA file with DNA sequences.
The previous versions of MyDbFinder included the options to use databases of Vibrio Cholerae or Virulence genes of Listeria. The first have been used to create the software CholeraeFinder, and the second one has been included in VirulenceFinder

Upload user database (DNA sequences FASTA format)
Note: Database must not be compressed.
Choose File No file chosen

Select threshold for %ID
98 %

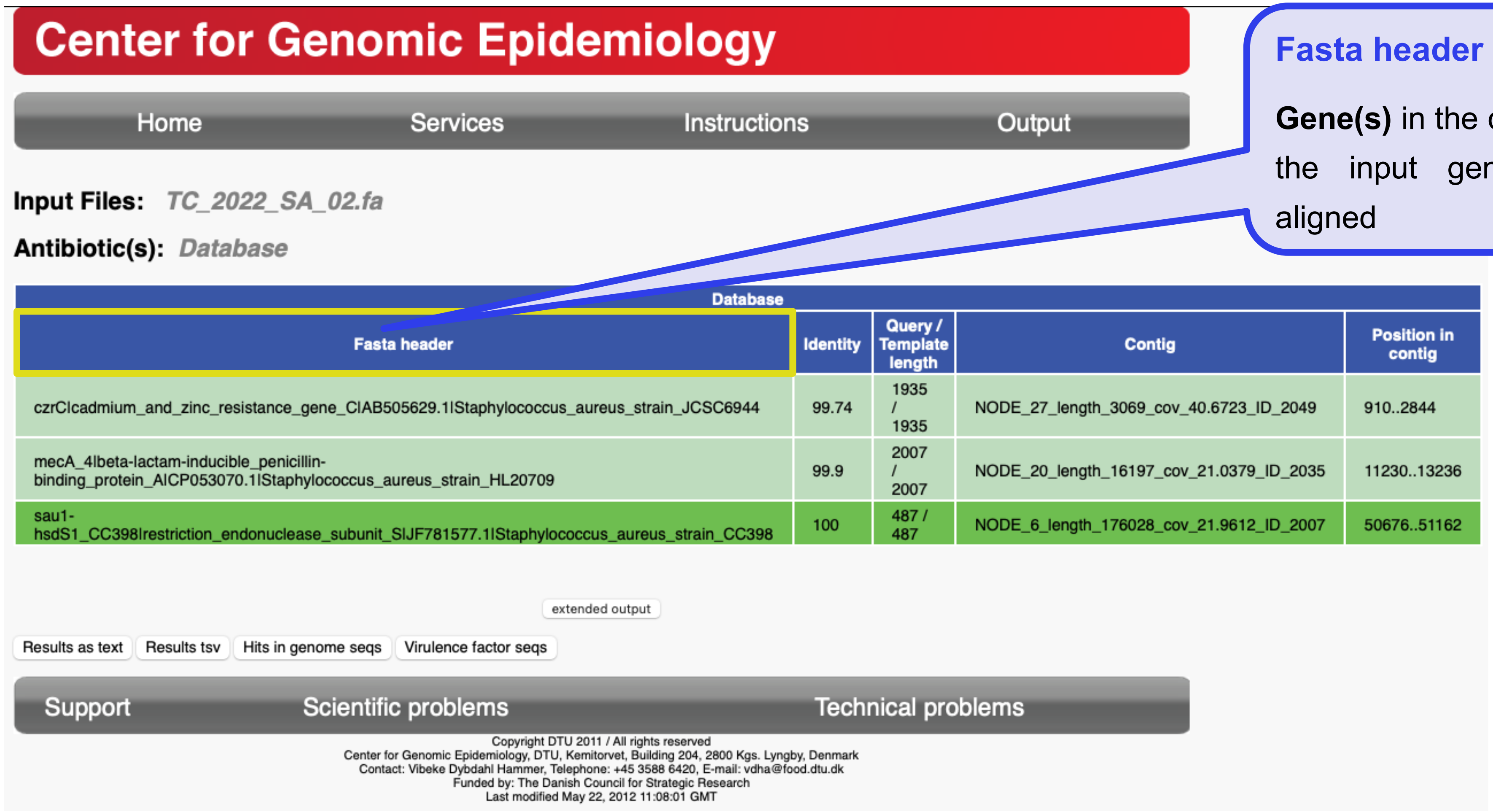
Select minimum length
Length a gene in the genome at least has to cover of the length of the gene in the database to be outputted
60 %

Select type of your reads
Assembled Genome/Contigs*

IMPORTANT NOTE:
To avoid problems caused by file names, we only allow a limited selection of ASCII characters (see below).

a-z
A-Z
0-9
_ (underscore)
- (hyphen)
. (full stop)

Analysis of Results



Input Files: *TC_2022_SA_02.fa*

Antibiotic(s): *Database*

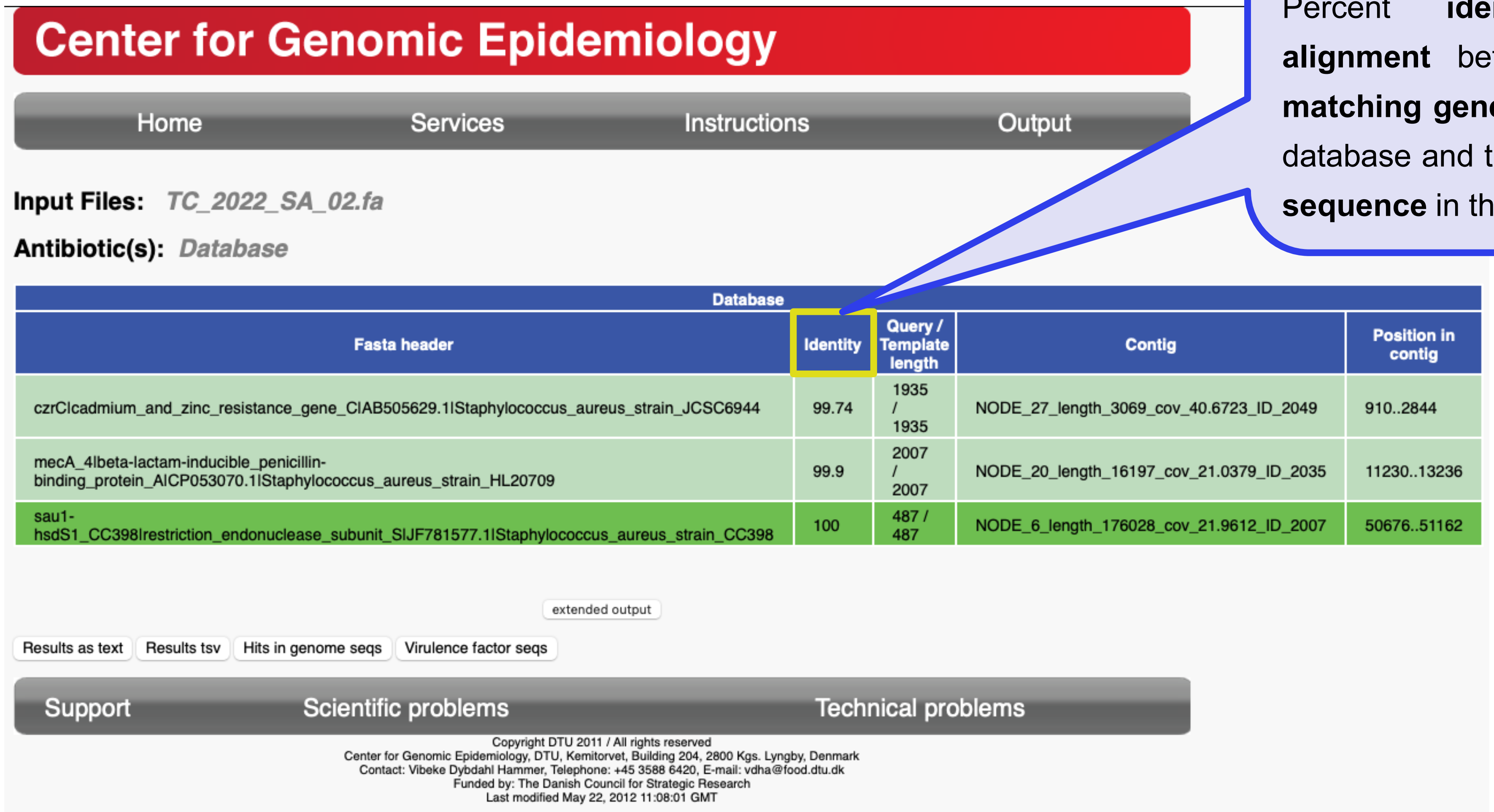
Database				
Fasta header	**Identity**	**Query / Template length**	**Contig**	**Position in contig**
czrC|cadmium_and_zinc_resistance_gene_C|AB505629.1|Staphylococcus_aureus_strain_JCSC6944	99.74	1935 / 1935	NODE_27_length_3069_cov_40.6723_ID_2049	910..2844
mecA_4|beta-lactam-inducible_penicillin-binding_protein_A|CP053070.1|Staphylococcus_aureus_strain_HL20709	99.9	2007 / 2007	NODE_20_length_16197_cov_21.0379_ID_2035	11230..13236
sau1-hsdS1_CC398|restriction_endonuclease_subunit_S|JF781577.1|Staphylococcus_aureus_strain_CC398	100	487 / 487	NODE_6_length_176028_cov_21.9612_ID_2007	50676..51162

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

Fasta header

Gene(s) in the database to which the input genome has been aligned

Analysis of Results



Center for Genomic Epidemiology

Input Files: *TC_2022_SA_02.fa*

Antibiotic(s): *Database*

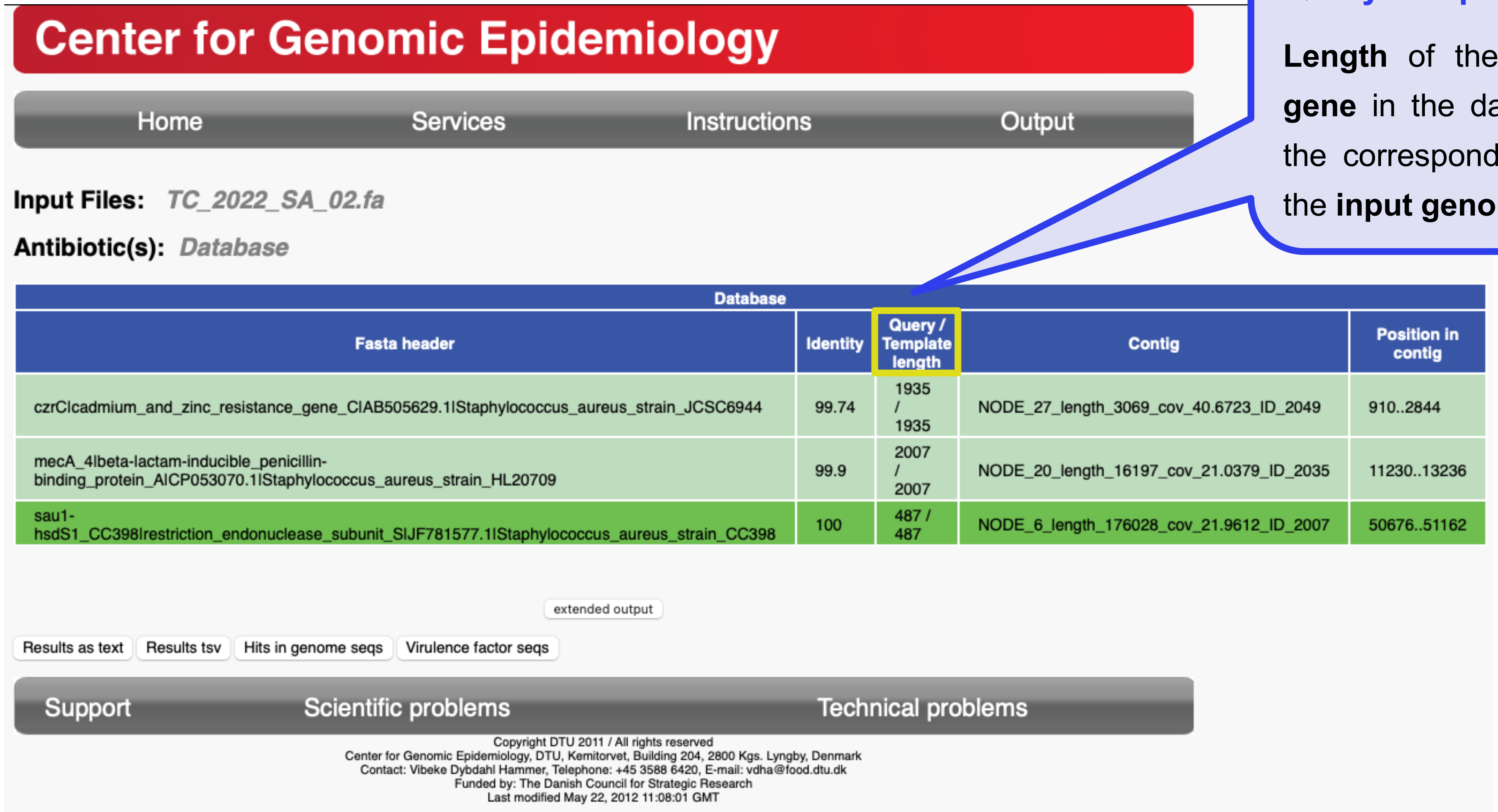
Database				
Fasta header	**Identity**	**Query / Template length**	**Contig**	**Position in contig**
czrC|cadmium_and_zinc_resistance_gene_C|AB505629.1|Staphylococcus_aureus_strain_JCSC6944	99.74	1935 / 1935	NODE_27_length_3069_cov_40.6723_ID_2049	910..2844
mecA_4|beta-lactam-inducible_penicillin-binding_protein_A|CP053070.1|Staphylococcus_aureus_strain_HL20709	99.9	2007 / 2007	NODE_20_length_16197_cov_21.0379_ID_2035	11230..13236
sau1-hsdS1_CC398|restriction_endonuclease_subunit_S|JF781577.1|Staphylococcus_aureus_strain_CC398	100	487 / 487	NODE_6_length_176028_cov_21.9612_ID_2007	50676..51162

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

% Identity

Percent **identity** in the **alignment** between the **best matching gene** in the uploaded database and the corresponding **sequence** in the **input genome**

Analysis of Results



Center for Genomic Epidemiology

Input Files: *TC_2022_SA_02.fa*

Antibiotic(s): *Database*

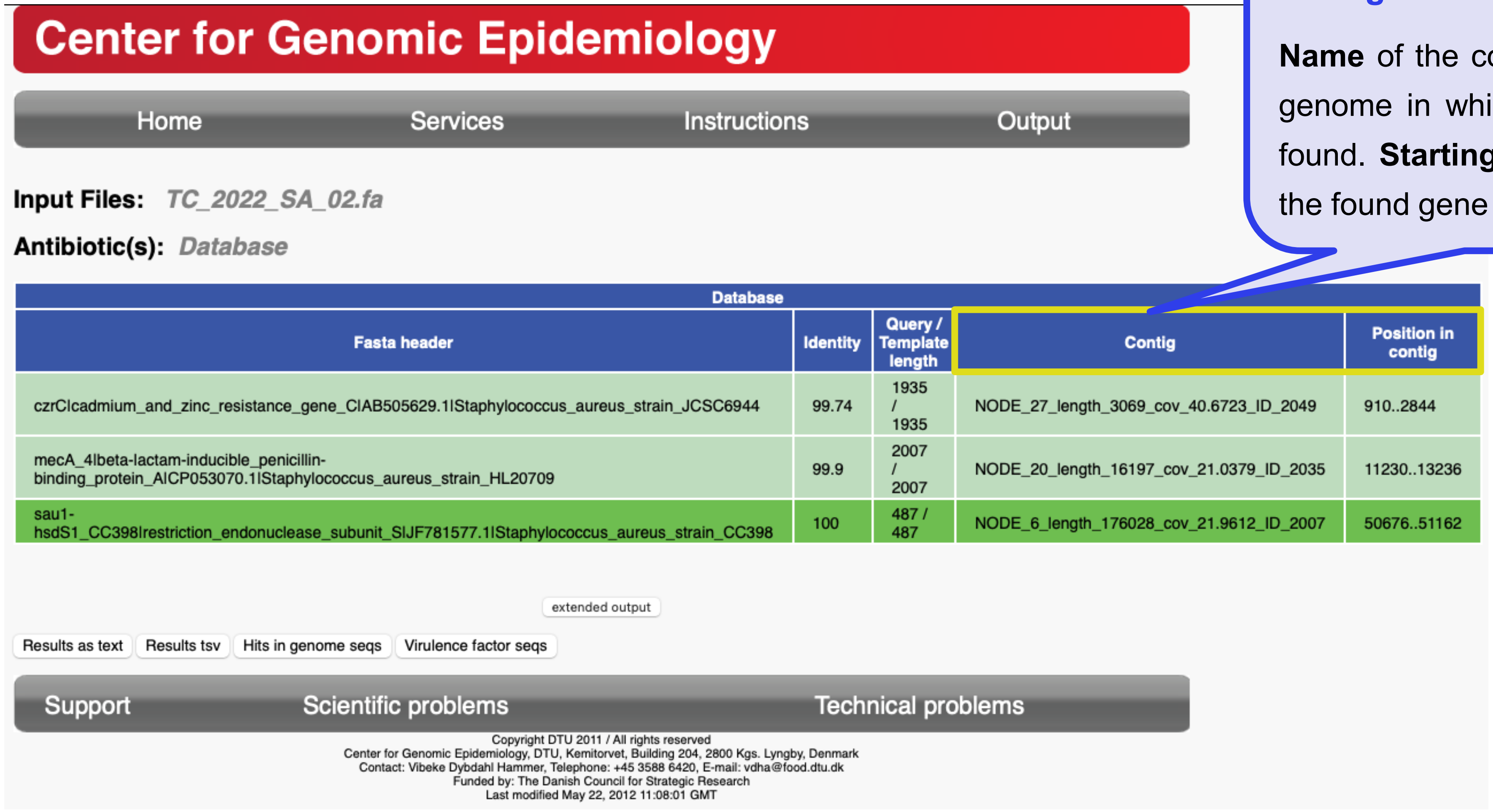
Database				
Fasta header	**Identity**	**Query / Template length**	**Contig**	**Position in contig**
czrC|cadmium_and_zinc_resistance_gene_C|AB505629.1|Staphylococcus_aureus_strain_JCSC6944	99.74	1935 / 1935	NODE_27_length_3069_cov_40.6723_ID_2049	910..2844
mecA_4|beta-lactam-inducible_penicillin-binding_protein_A|CP053070.1|Staphylococcus_aureus_strain_HL20709	99.9	2007 / 2007	NODE_20_length_16197_cov_21.0379_ID_2035	11230..13236
sau1-hsdS1_CC398|restriction_endonuclease_subunit_S|JF781577.1|Staphylococcus_aureus_strain_CC398	100	487 / 487	NODE_6_length_176028_cov_21.9612_ID_2007	50676..51162

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

Query/Template Length

Length of the **best matching gene** in the database/**length** of the corresponding **sequence** in the **input genome**

Analysis of Results



Input Files: *TC_2022_SA_02.fa*

Antibiotic(s): *Database*

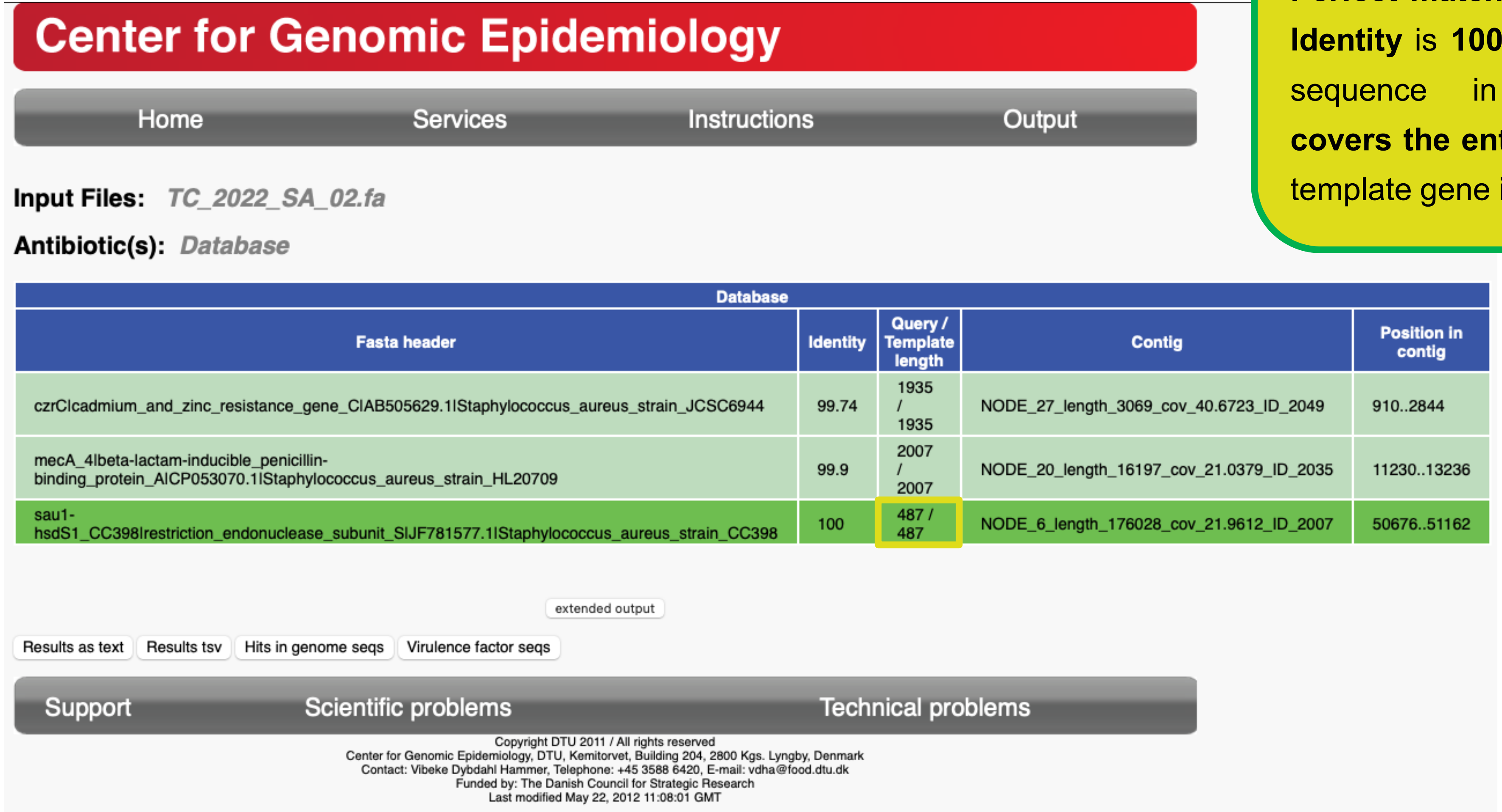
Database				
Fasta header	**Identity**	**Query / Template length**	**Contig**	**Position in contig**
czrClcadmium_and_zinc_resistance_gene_ClAB505629.1lStaphylococcus_aureus_strain_JCSC6944	99.74	1935 / 1935	NODE_27_length_3069_cov_40.6723_ID_2049	910..2844
mecA_4lbeta-lactam-inducible_penicillin-binding_protein_AlCP053070.1lStaphylococcus_aureus_strain_HL20709	99.9	2007 / 2007	NODE_20_length_16197_cov_21.0379_ID_2035	11230..13236
sau1-hsdS1_CC398lrestriction_endonuclease_subunit_SlJF781577.1lStaphylococcus_aureus_strain_CC398	100	487 / 487	NODE_6_length_176028_cov_21.9612_ID_2007	50676..51162

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

Contig and Position

Name of the contig in the query genome in which the gene was found. **Starting-end position** of the found gene in the contig

Analysis of Results



Input Files: *TC_2022_SA_02.fa*

Antibiotic(s): *Database*

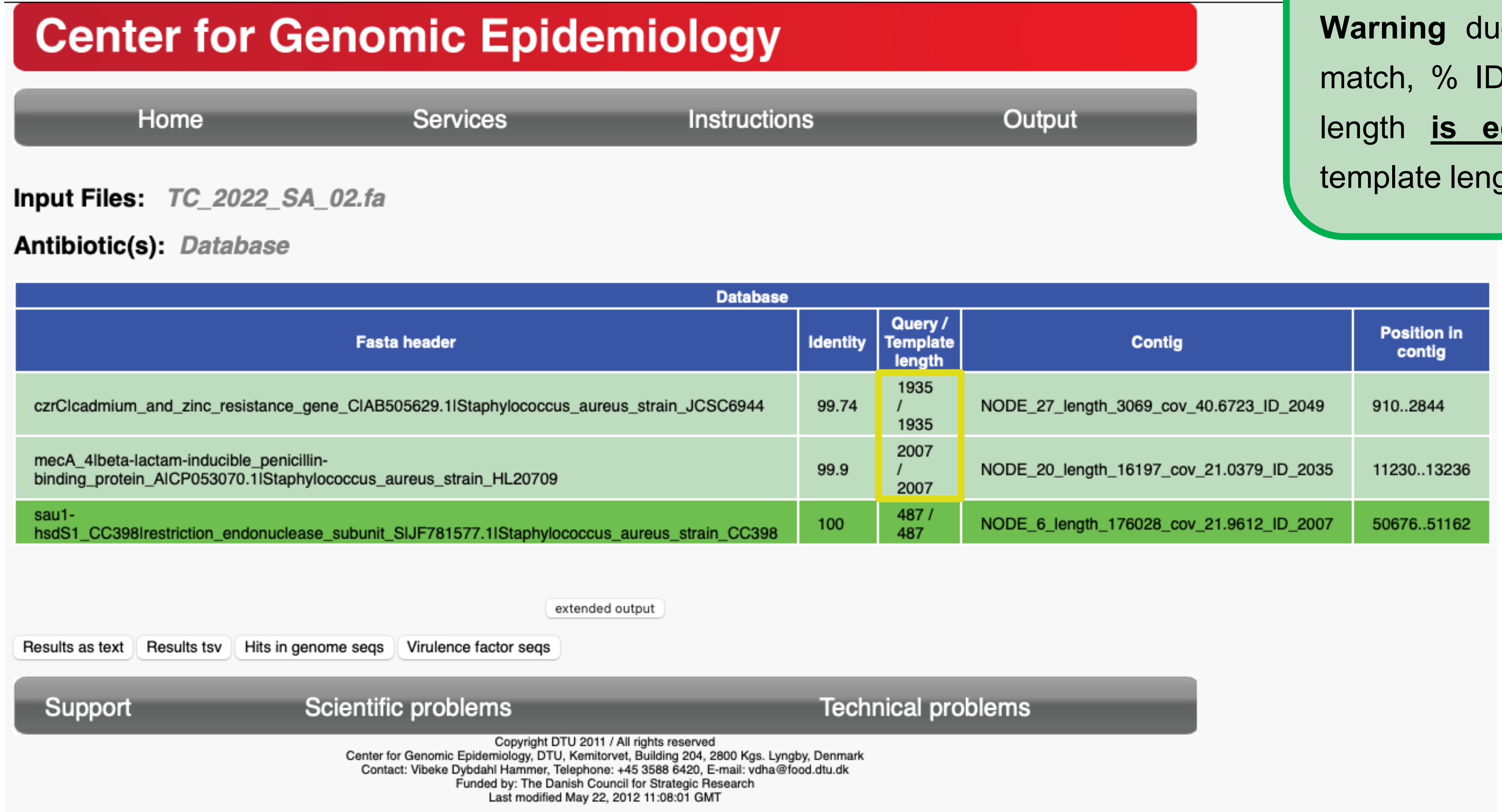
Database				
Fasta header	**Identity**	**Query / Template length**	**Contig**	**Position in contig**
czrC|cadmium_and_zinc_resistance_gene_C|AB505629.1|Staphylococcus_aureus_strain_JCSC6944	99.74	1935 / 1935	NODE_27_length_3069_cov_40.6723_ID_2049	910..2844
mecA_4|beta-lactam-inducible_penicillin-binding_protein_A|CP053070.1|Staphylococcus_aureus_strain_HL20709	99.9	2007 / 2007	NODE_20_length_16197_cov_21.0379_ID_2035	11230..13236
sau1-hsdS1_CC398|restriction_endonuclease_subunit_S|JF781577.1|Staphylococcus_aureus_strain_CC398	100	487 / 487	NODE_6_length_176028_cov_21.9612_ID_2007	50676..51162

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

The dark green color

Perfect match for a give gene. **Identity** is **100%** and the query sequence in the genome **covers the entire length** of the template gene in the database

Analysis of Results



Center for Genomic Epidemiology

Input Files: *TC_2022_SA_02.fa*

Antibiotic(s): *Database*

Database

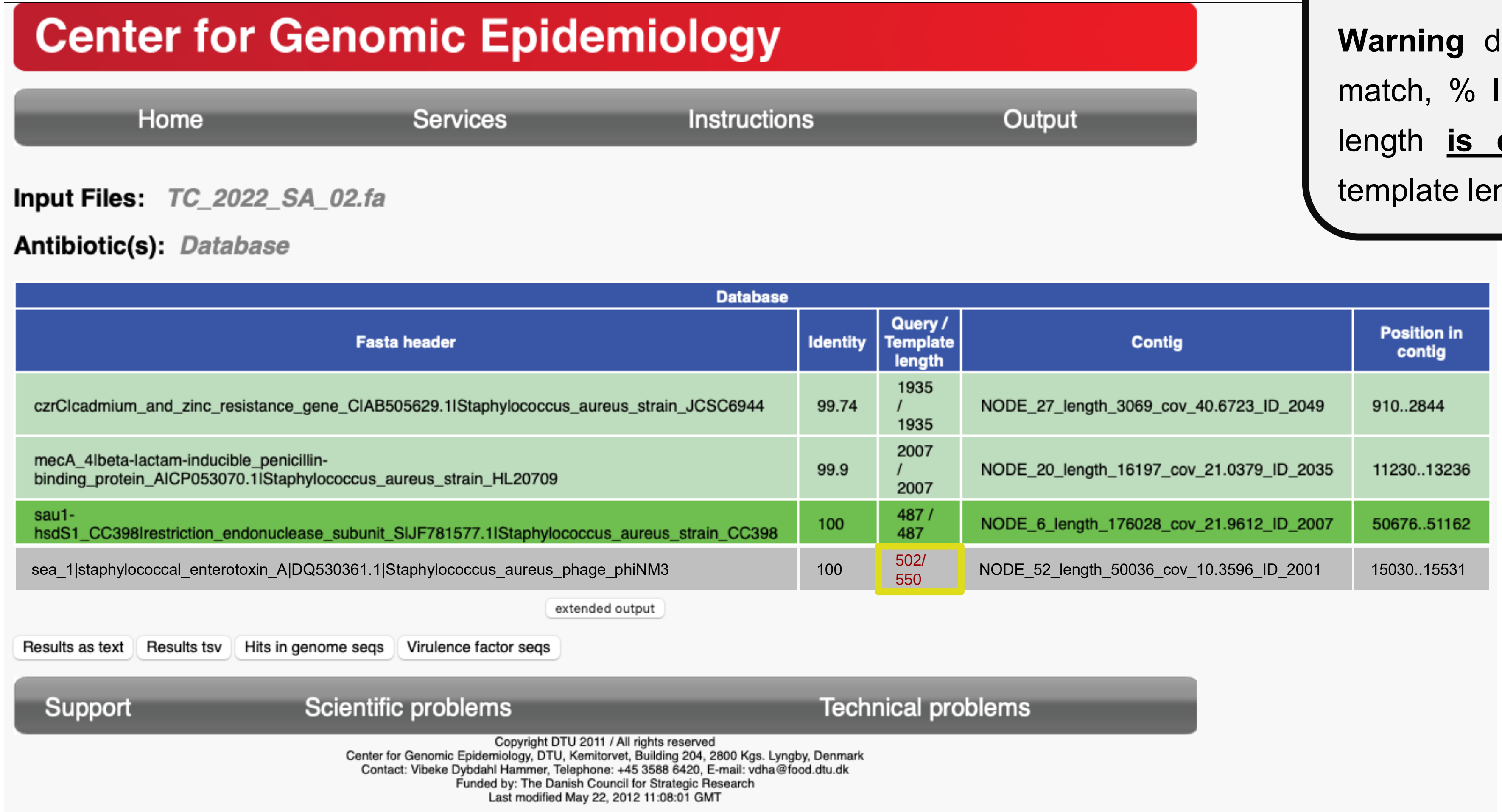
Fasta header	Identity	Query / Template length	Contig	Position in contig
czrCIcadmium_and_zinc_resistance_gene_CIAB505629.1IStaphylococcus_aureus_strain_JCSC6944	99.74	1935 / 1935	NODE_27_length_3069_cov_40.6723_ID_2049	910..2844
mecA_4Ibeta-lactam-inducible_penicillin-binding_protein_AICP053070.1IStaphylococcus_aureus_strain_HL20709	99.9	2007 / 2007	NODE_20_length_16197_cov_21.0379_ID_2035	11230..13236
sau1-hsdS1_CC398Irestriction_endonuclease_subunit_SIJF781577.1IStaphylococcus_aureus_strain_CC398	100	487 / 487	NODE_6_length_176028_cov_21.9612_ID_2007	50676..51162

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

The light green color

Warning due to a non-perfect match, % ID < 100% ➜ query length **is equal** to the gene template length

Analysis of Results



Center for Genomic Epidemiology

Input Files: *TC_2022_SA_02.fa*

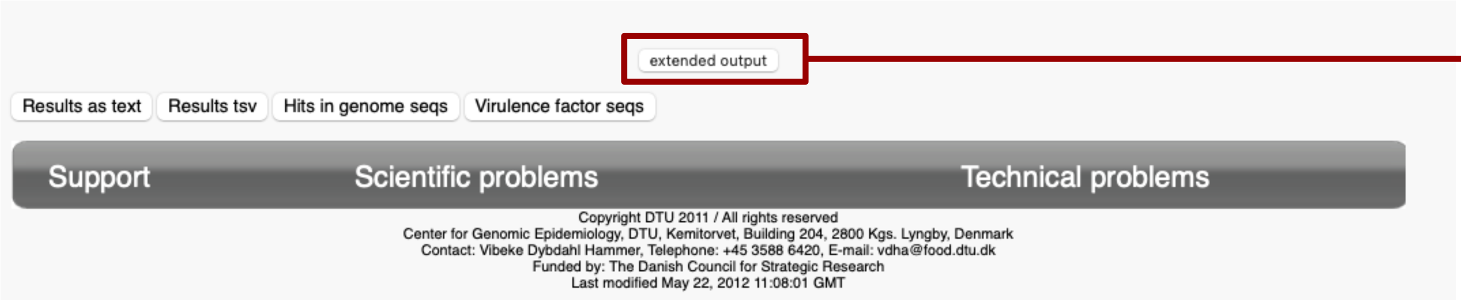
Antibiotic(s): *Database*

Database				
Fasta header	**Identity**	**Query / Template length**	**Contig**	**Position in contig**
czrClcadmium_and_zinc_resistance_gene_ClAB505629.1lStaphylococcus_aureus_strain_JCSC6944	99.74	1935 / 1935	NODE_27_length_3069_cov_40.6723_ID_2049	910..2844
mecA_4lbeta-lactam-inducible_penicillin-binding_protein_AlCP053070.1lStaphylococcus_aureus_strain_HL20709	99.9	2007 / 2007	NODE_20_length_16197_cov_21.0379_ID_2035	11230..13236
sau1-hsdS1_CC398lrestriction_endonuclease_subunit_SlJF781577.1lStaphylococcus_aureus_strain_CC398	100	487 / 487	NODE_6_length_176028_cov_21.9612_ID_2007	50676..51162
sea_1|staphylococcal_enterotoxin_A|DQ530361.1|Staphylococcus_aureus_phage_phiNM3	100	502/ 550	NODE_52_length_50036_cov_10.3596_ID_2001	15030..15531

The grey color

Warning due to a non-perfect match, % ID is 100% → query length **<u>is different</u>** than gene template length

Analysis of Results



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

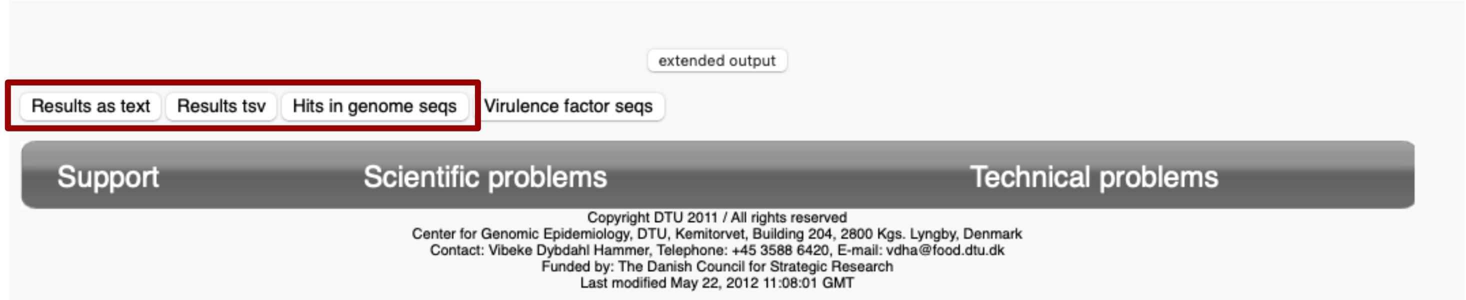
Shows the **alignments** between the template gene sequence (database) and the query gene sequence

sau1-hsdS1_CC398|restriction_endonuclease_subunit_S|JF781577.1|Staphylococcus_aureus_strain_CC398

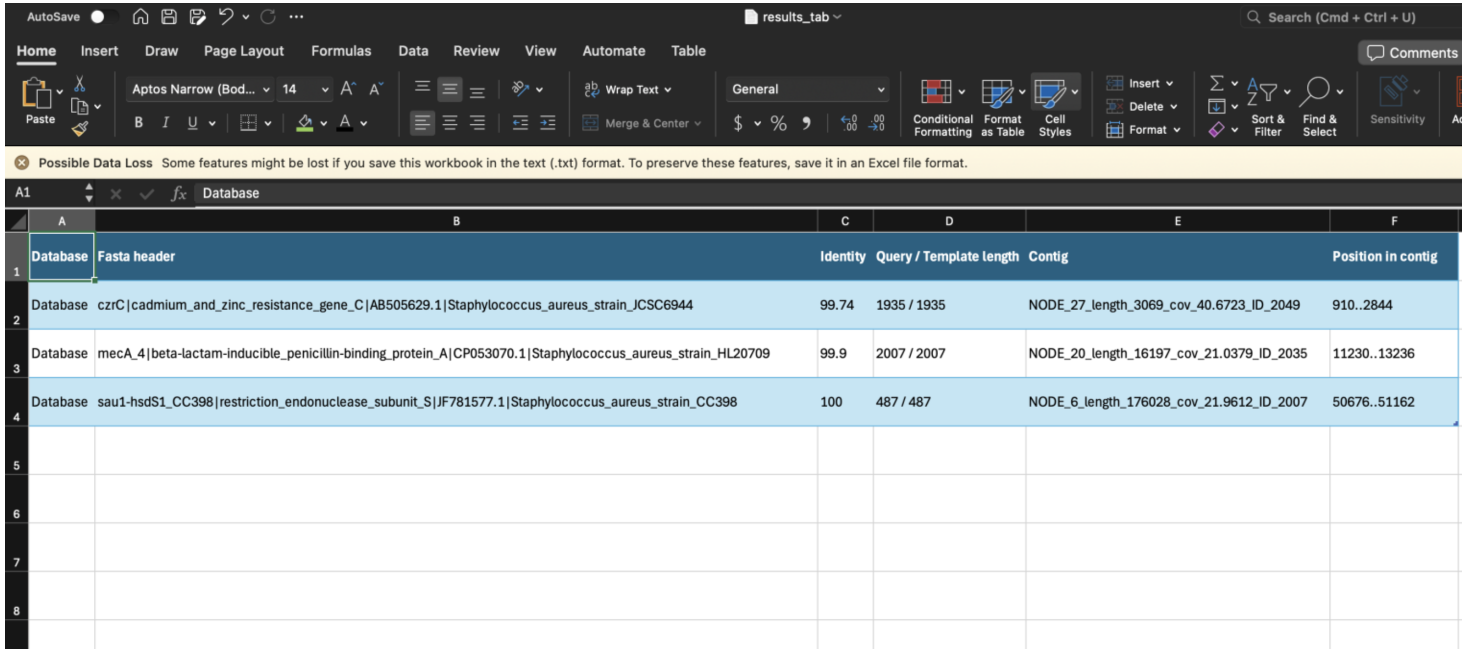
```
template   GCGAATGGGAAGAGAAGAAGCTAGGTGAGTTTGCTGGTAAAGTTACCCAAAAAAATGTTG
query      GCGAATGGGAAGAGAAGAAGCTAGGTGAGTTTGCTGGTAAAGTTACCCAAAAAAATGTTG

template   ATAAAAAATATATTGAGACATTAACTAATTCAGCTGAGTTAGGTATCATATCTCAAAAGG
query      ATAAAAAATATATTGAGACATTAACTAATTCAGCTGAGTTAGGTATCATATCTCAAAAGG

template   ATTATTTTGACAAAGAAATTTCGAATATAGATAATATTAAAAAGTACTATGTAGTTGAAG
query      ATTATTTTGACAAAGAAATTTCGAATATAGATAATATTAAAAAGTACTATGTAGTTGAAG

template   AGAATGATTTTGTTTATAACCCTAGAATGTCTAATTATGCTCCATTTGGACCAGTAAATA
query      AGAATGATTTTGTTTATAACCCTAGAATGTCTAATTTTGCTCCATTTGGACCAGTAAATA

template   GAAATAAGTTAGGGAAAAAAGGGGTCATGTCACCTCTTTATACTGTGTTTAAAATTCAAA
query      GAAATAAGTTAGGGAAAAAAGGGGTCATGTCACCTCTTTATACTGTGTTTAAAATTCAAA

template   ACATTGATTTAAACTTTATTGAGTTTTATTTTAAATCTTCAAAATGGTATAGATTTATGG
query      ACATTGATTTAAACTTTATTGAGTTTTATTTTAAATCTTCAAAATGGTATAGATTTATGG

template   CATTAAACGGTGATTCAGGTGCTCGAGCAGATAGGTTTTCTATTAAAGATAGGACATTTA
query      CATTAAACGGTGATTCAGGTGCTCGAGCAGATAGGTTTTCTATTAAAGATAGGACATTTA

template   TGGAAATGCCACTTCATATCCCATGTATGGATGAACAAATAAAAATCGGTCAGTTCTTCA
query      TGGAAATGCCACTTCATATCCCATGTATGGATGAACAAATAAAAATCGGTCAGTTCTTCA

template   GCAAACT  ATCGGACTCAAAAT
query      GCAAACT
```

1. **Green colour** - matching nucleotide
2. **Red colour** – mismatches
3. **Gray colour** – missing sequence in part of the alignment

Analysis of Results



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



Possible Data Loss Some features might be lost if you save this workbook in the text (.txt) format. To preserve these features, save it in an Excel file format.

A1 Database

Database	Fasta header	Identity	Query / Template length	Contig	Position in contig
Database	czrC|cadmium_and_zinc_resistance_gene_C|AB505629.1|Staphylococcus_aureus_strain_JCSC6944	99.74	1935 / 1935	NODE_27_length_3069_cov_40.6723_ID_2049	910..2844
Database	mecA_4|beta-lactam-inducible_penicillin-binding_protein_A|CP053070.1|Staphylococcus_aureus_strain_HL20709	99.9	2007 / 2007	NODE_20_length_16197_cov_21.0379_ID_2035	11230..13236
Database	sau1-hsdS1_CC398|restriction_endonuclease_subunit_S|JF781577.1|Staphylococcus_aureus_strain_CC398	100	487 / 487	NODE_6_length_176028_cov_21.9612_ID_2007	50676..51162

Results download:

- **Results as text:** a *text* file containing the result table and alignments ➡️ can be read by most text editors
- **Results as tsv:** a *tsv* file containing the result table ➡️ can be read by **Excel**
- **Hits in genome seqs:** a *fasta* file containing the best matching sequences from the query genome

Thank you for your attention! 😊

Open to questions/comments/suggestions

JOANA MOURÃO

EURL-AR