

# Package ‘AssessORFData’

January 22, 2026

**Type** Package

**Title** Data and Files for the AssessORF Package

**Version** 1.28.0

**Date** 2019-04-07

**Description** This package provides access to mapping and results objects generated by the AssessORF package, as well as the genome sequences for the strains corresponding to those objects.

**Depends** R (>= 3.5.0), RSQLite (>= 1.1)

**Imports** DECIPHER, utils

**Suggests** AssessORF, BiocStyle, knitr, rmarkdown

**biocViews** OrganismData, Bacillus\_subtilis\_Data, Escherichia\_coli\_Data,  
Pseudomonas\_aeruginosa\_Data, Staphylococcus\_aureus\_Data,  
Genome, Proteome, SequencingData

**License** GPL-3

**Encoding** UTF-8

**LazyData** FALSE

**NeedsCompilation** no

**RoxygenNote** 7.1.1

**VignetteBuilder** knitr

**git\_url** <https://git.bioconductor.org/packages/AssessORFData>

**git\_branch** RELEASE\_3\_22

**git\_last\_commit** 2cd0c9f

**git\_last\_commit\_date** 2025-10-29

**Repository** Bioconductor 3.22

**Date/Publication** 2026-01-22

**Author** Deepank Korandla [aut],  
Nicholas Cooley [cre] (ORCID: <<https://orcid.org/0000-0002-6029-304X>>)

**Maintainer** Nicholas Cooley <npc19@pitt.edu>

## Contents

AP1 . . . . .	2
AssessORF_StrainIDs . . . . .	4
ATCC11842 . . . . .	4
ATCC13032 . . . . .	6
ATCC17978 . . . . .	7
ATCC700084 . . . . .	9
BW25113 . . . . .	10
CCMP1375 . . . . .	12
CECT5344 . . . . .	13
CNRZ327 . . . . .	15
COH1 . . . . .	16
D_UW_3_CX . . . . .	18
EGD_e . . . . .	19
GetDataMapObj . . . . .	21
GetGeneSources . . . . .	21
GetResultsObj . . . . .	22
GetStrainIDs . . . . .	23
H37Rv . . . . .	23
HG001 . . . . .	25
Houston_1 . . . . .	26
II1403 . . . . .	28
K_12_MG1655 . . . . .	29
LAL14_1 . . . . .	31
MG1363 . . . . .	32
MGAS5005 . . . . .	34
NCIB_3610 . . . . .	35
PAO1 . . . . .	37
SaveGenomeToPath . . . . .	38
SL1344 . . . . .	39
Strain10403S . . . . .	41
Strain168 . . . . .	42
TCH1516 . . . . .	44
<b>Index</b>	<b>46</b>

AP1

Assessment *Objects for Streptococcus pyogenes strain AP1*

### Description

Objects of class Assessment and either subclass DataMap or subclass Results for *Streptococcus pyogenes* strain AP1

### Usage

```
data(AP1_PreSaved_DataMapObj)
data(AP1_PreSaved_ResultsObj_GenBank)
data(AP1_PreSaved_ResultsObj_GeneMarkS2)
data(AP1_PreSaved_ResultsObj_Glimmer)
data(AP1_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of Assessment objects contain, please see the **AssessORF** package.

## Details

'AP1\_PreSaved\_DataMapObj' is an object of subclass DataMap.

'AP1\_PreSaved\_ResultsObj\_GenBank' is an object of subclass Results with predicted genes from the GenBank database.

'AP1\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass Results with predicted genes from the program GeneMarkS-2.

'AP1\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass Results with predicted genes from the program Glimmer.

'AP1\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass Results with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: AP1
- Species: *S. pyogenes*

## Mapping object

The mapping object, 'AP1\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD006345. The related genomes used to determine evolutionary conservation all came from the family, Streptococcaceae excluding anomalous genomes and non-complete *Streptococcus pneumoniae* genomes. Links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the MapAssessmentData function from the **AssessORF** package.

## Results objects

The 4 results objects, 'AP1\_PreSaved\_ResultsObj\_GenBank', 'AP1\_PreSaved\_ResultsObj\_GeneMarkS2', 'AP1\_PreSaved\_ResultsObj\_Glimmer', and 'AP1\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("AP1", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD006345>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucore/CP007537.1>

---

AssessORF_StrainIDs	<i>Vector of strain IDs used in the package</i>
---------------------	-------------------------------------------------

---

## Description

Vector of strain IDs, which describes the strains for which the package has data

## Usage

```
data(AssessORF_StrainIDs)
```

## Format

Character vector of length 26

---

ATCC11842	<i>Assessment Objects for Lactobacillus delbrueckii subsp. bulgaricus strain ATCC 11842</i>
-----------	---------------------------------------------------------------------------------------------

---

## Description

Objects of class *Assessment* and either subclass *DataMap* or subclass *Results* for *Lactobacillus delbrueckii* subsp. *bulgaricus* strain ATCC 11842

## Usage

```
data(ATCC11842_PreSaved_DataMapObj)
data(ATCC11842_PreSaved_ResultsObj_GenBank)
data(ATCC11842_PreSaved_ResultsObj_GeneMarks2)
data(ATCC11842_PreSaved_ResultsObj_Glimmer)
data(ATCC11842_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of *Assessment* objects contain, please see the **AssessORF** package.

## Details

'ATCC11842\_PreSaved\_DataMapObj' is an object of subclass DataMap.

'ATCC11842\_PreSaved\_ResultsObj\_GenBank' is an object of subclass Results with predicted genes from the GenBank database.

'ATCC11842\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass Results with predicted genes from the program GeneMarkS-2.

'ATCC11842\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass Results with predicted genes from the program Glimmer.

'ATCC11842\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass Results with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: ATCC11842
- Species: *L. delbrueckii bulgaricus*

## Mapping object

The mapping object, 'ATCC11842\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD006551. The related genomes used to determine evolutionary conservation all came from the genus *Lactobacillus*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the MapAssessmentData function from the **AssessORF** package.

## Results objects

The 4 results objects, 'ATCC11842\_PreSaved\_ResultsObj\_GenBank', 'ATCC11842\_PreSaved\_ResultsObj\_GeneMarkS2', 'ATCC11842\_PreSaved\_ResultsObj\_Glimmer', and 'ATCC11842\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("ATCC11842", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX006551>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: [https://www.ncbi.nlm.nih.gov/nucleotide/NC\\_008054.1](https://www.ncbi.nlm.nih.gov/nucleotide/NC_008054.1)

---

ATCC13032	Assessment <i>Objects for Corynebacterium glutamicum strain ATCC 13032</i>
-----------	----------------------------------------------------------------------------

---

## Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Corynebacterium glutamicum* strain ATCC 13032

## Usage

```
data(ATCC13032_PreSaved_DataMapObj)
data(ATCC13032_PreSaved_ResultsObj_GenBank)
data(ATCC13032_PreSaved_ResultsObj_GeneMarkS2)
data(ATCC13032_PreSaved_ResultsObj_Glimmer)
data(ATCC13032_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

## Details

'ATCC13032\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'ATCC13032\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'ATCC13032\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'ATCC13032\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'ATCC13032\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: ATCC13032
- Species: *C. glutamicum*

## Mapping object

The mapping object, 'ATCC13032\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD005812. The related genomes used to determine evolutionary conservation came from the genera *Corynebacterium*, *Dietzia*, and *Tsukamurella*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

## Results objects

The 4 results objects, 'ATCC13032\_PreSaved\_ResultsObj\_GenBank', 'ATCC13032\_PreSaved\_ResultsObj\_GeneMarkS2', 'ATCC13032\_PreSaved\_ResultsObj\_Glimmer', and 'ATCC13032\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("ATCC13032", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD005812>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nuccore/BA000036.3>

---

ATCC17978

Assessment *Objects for Acinetobacter baumannii strain ATCC 17978*

---

## Description

Objects of class Assessment and either subclass DataMap or subclass Results for *Acinetobacter baumannii* strain ATCC 17978

## Usage

```
data(ATCC17978_PreSaved_DataMapObj)
data(ATCC17978_PreSaved_ResultsObj_GenBank)
data(ATCC17978_PreSaved_ResultsObj_GeneMarkS2)
data(ATCC17978_PreSaved_ResultsObj_Glimmer)
data(ATCC17978_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of Assessment objects contain, please see the **AssessORF** package.

## Details

'ATCC17978\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'ATCC17978\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'ATCC17978\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'ATCC17978\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'ATCC17978\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- `StrainID`: ATCC17978
- `Species`: *A. baumannii*

## Mapping object

The mapping object, 'ATCC17978\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The proteomics data was generated in-house and collected as part of the work for the AssessORF paper. The raw mass spectra data can be found at ProteomeXchange dataset PXD012539. The related genomes used to determine evolutionary conservation all came from the genus *Acinetobacter*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

## Results objects

The 4 results objects, 'ATCC17978\_PreSaved\_ResultsObj\_GenBank', 'ATCC17978\_PreSaved\_ResultsObj\_GeneMarkS2', 'ATCC17978\_PreSaved\_ResultsObj\_Glimmer', and 'ATCC17978\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the `AssessGenes` function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The `GeneLeftPos`, `GeneRightPos`, and the `GeneStrand` within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The `GeneSource` list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use `SaveGenomeToPath("ATCC17978", <INSERT FILE PATH HERE>)` to save the genome sequence for the strain to your local machine.



## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD012539>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucleotide/CP000521.1>

---

ATCC700084	Assessment <i>Objects for Mycobacterium smegmatis strain ATCC 700084</i>
------------	--------------------------------------------------------------------------

---

## Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Mycobacterium smegmatis* strain ATCC 700084

## Usage

```
data(ATCC700084_PreSaved_DataMapObj)
data(ATCC700084_PreSaved_ResultsObj_GenBank)
data(ATCC700084_PreSaved_ResultsObj_GeneMarkS2)
data(ATCC700084_PreSaved_ResultsObj_Glimmer)
data(ATCC700084_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

## Details

'ATCC700084\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'ATCC700084\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'ATCC700084\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'ATCC700084\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'ATCC700084\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: ATCC700084
- Species: *M. smegmatis*

## Mapping object

The mapping object, 'ATCC700084\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD003500. The related genomes used to determine evolutionary conservation all came from the genus *Mycobacterium*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

## Results objects

The 4 results objects, 'ATCC700084\_PreSaved\_ResultsObj\_GenBank', 'ATCC700084\_PreSaved\_ResultsObj\_GeneMarkS2', 'ATCC700084\_PreSaved\_ResultsObj\_Glimmer', and 'ATCC700084\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the `AssessGenes` function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The `GeneLeftPos`, `GeneRightPos`, and the `GeneStrand` within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The `GeneSource` list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use `SaveGenomeToPath("ATCC700084", <INSERT FILE PATH HERE>)` to save the genome sequence for the strain to your local machine. Please note that there is no genome sequence available for strain ATCC 700084 so the reference genome from strain MC2 155 was used instead.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD003500>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: [https://www.ncbi.nlm.nih.gov/nucleotide/NC\\_008596.1](https://www.ncbi.nlm.nih.gov/nucleotide/NC_008596.1)

---

BW25113

Assessment *Objects for Escherichia coli strain BW25113*

---

## Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Escherichia coli* strain BW25113

## Usage

```
data(BW25113_PreSaved_DataMapObj)
data(BW25113_PreSaved_ResultsObj_GenBank)
data(BW25113_PreSaved_ResultsObj_GeneMarkS2)
data(BW25113_PreSaved_ResultsObj_Glimmer)
data(BW25113_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

## Details

'BW25113\_PreSaved\_DataMapObj' is an object of subclass DataMap.

'BW25113\_PreSaved\_ResultsObj\_GenBank' is an object of subclass Results with predicted genes from the GenBank database.

'BW25113\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass Results with predicted genes from the program GeneMarkS-2.

'BW25113\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass Results with predicted genes from the program Glimmer.

'BW25113\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass Results with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: BW25113
- Species: E. coli

## Mapping object

The mapping object, 'BW25113\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD000498. The related genomes used to determine evolutionary conservation all came from the genus *Escherichia*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the MapAssessmentData function from the **AssessORF** package.

## Results objects

The 4 results objects, 'BW25113\_PreSaved\_ResultsObj\_GenBank', 'BW25113\_PreSaved\_ResultsObj\_GeneMarkS2', 'BW25113\_PreSaved\_ResultsObj\_Glimmer', and 'BW25113\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("BW25113", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX000498>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: [https://www.ncbi.nlm.nih.gov/nucleotide/NZ\\_CP009273.1](https://www.ncbi.nlm.nih.gov/nucleotide/NZ_CP009273.1)

---

CCMP1375	Assessment <i>Objects for Prochlorococcus marinus subsp. marinus strain CCMP1375</i>
----------	--------------------------------------------------------------------------------------

---

## Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Prochlorococcus marinus* subsp. *marinus* strain CCMP1375

## Usage

```
data(CCMP1375_PreSaved_DataMapObj)
data(CCMP1375_PreSaved_ResultsObj_GenBank)
data(CCMP1375_PreSaved_ResultsObj_GeneMarkS2)
data(CCMP1375_PreSaved_ResultsObj_Glimmer)
data(CCMP1375_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

## Details

'CCMP1375\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'CCMP1375\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'CCMP1375\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'CCMP1375\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'CCMP1375\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: CCMP1375
- Species: *P. marinus*

## Mapping object

The mapping object, 'CCMP1375\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD005745. The related genomes used to determine evolutionary conservation all came from the order *Synechococcales*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

## Results objects

The 4 results objects, 'CCMP1375\_PreSaved\_ResultsObj\_GenBank', 'CCMP1375\_PreSaved\_ResultsObj\_GeneMarkS2', 'CCMP1375\_PreSaved\_ResultsObj\_Glimmer', and 'CCMP1375\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("CCMP1375", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD005745>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: [https://www.ncbi.nlm.nih.gov/nucleotide/NC\\_005042.1](https://www.ncbi.nlm.nih.gov/nucleotide/NC_005042.1)

---

CECT5344	Assessment Objects for <i>Pseudomonas pseudoalcaligenes</i> strain CECT 5344
----------	------------------------------------------------------------------------------

---

## Description

Objects of class Assessment and either subclass DataMap or subclass Results for *Pseudomonas pseudoalcaligenes* strain CECT 5344

## Usage

```
data(CECT5344_PreSaved_DataMapObj)
data(CECT5344_PreSaved_ResultsObj_GenBank)
data(CECT5344_PreSaved_ResultsObj_GeneMarkS2)
data(CECT5344_PreSaved_ResultsObj_Glimmer)
data(CECT5344_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of Assessment objects contain, please see the **AssessORF** package.

## Details

'CECT5344\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'CECT5344\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'CECT5344\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'CECT5344\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'CECT5344\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: CECT5344
- Species: *P. pseudoalcaligenes*

## Mapping object

The mapping object, 'CECT5344\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD005745. The related genomes used to determine evolutionary conservation all came from the genus *Pseudomonas*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

## Results objects

The 4 results objects, 'CECT5344\_PreSaved\_ResultsObj\_GenBank', 'CECT5344\_PreSaved\_ResultsObj\_GeneMarkS2', 'CECT5344\_PreSaved\_ResultsObj\_Glimmer', and 'CECT5344\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the `AssessGenes` function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The `GeneLeftPos`, `GeneRightPos`, and the `GeneStrand` within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The `GeneSource` list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use `SaveGenomeToPath("CECT5344", <INSERT FILE PATH HERE>)` to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX005745>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucleotide/HG916826.1>

---

CNRZ327	Assessment <i>Objects for Lactobacillus delbrueckii subsp. lactis strain LBCNRZ327_V11</i>
---------	--------------------------------------------------------------------------------------------

---

## Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Lactobacillus delbrueckii subsp. lactis* strain LBCNRZ327\_V11

## Usage

```
data(CNRZ327_PreSaved_DataMapObj)
data(CNRZ327_PreSaved_ResultsObj_GenBank)
data(CNRZ327_PreSaved_ResultsObj_GeneMarkS2)
data(CNRZ327_PreSaved_ResultsObj_Glimmer)
data(CNRZ327_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

## Details

'CNRZ327\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'CNRZ327\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'CNRZ327\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'CNRZ327\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'CNRZ327\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: CNRZ327
- Species: *L. delbrueckii lactis*

## Mapping object

The mapping object, 'CNRZ327\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD006551. The related genomes used to determine evolutionary conservation all came from the genus *Lactobacillus*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

## Results objects

The 4 results objects, 'CNRZ327\_PreSaved\_ResultsObj\_GenBank', 'CNRZ327\_PreSaved\_ResultsObj\_GeneMarkS2', 'CNRZ327\_PreSaved\_ResultsObj\_Glimmer', and 'CNRZ327\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("CNRZ327", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD006551>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nuccore/CCDV01000001.1>

---

COH1

---

Assessment *Objects for Streptococcus agalactiae strain COH1*


---

## Description

Objects of class Assessment and either subclass DataMap or subclass Results for *Streptococcus agalactiae* strain COH1

## Usage

```
data(COH1_PreSaved_DataMapObj)
data(COH1_PreSaved_ResultsObj_GenBank)
data(COH1_PreSaved_ResultsObj_GeneMarkS2)
data(COH1_PreSaved_ResultsObj_Glimmer)
data(COH1_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of Assessment objects contain, please see the **AssessORF** package.



## Details

'COH1\_PreSaved\_DataMapObj' is an object of subclass DataMap.

'COH1\_PreSaved\_ResultsObj\_GenBank' is an object of subclass Results with predicted genes from the GenBank database.

'COH1\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass Results with predicted genes from the program GeneMarkS-2.

'COH1\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass Results with predicted genes from the program Glimmer.

'COH1\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass Results with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: COH1
- Species: *S. agalactiae*

## Mapping object

The mapping object, 'COH1\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The proteomics data was generated in-house and collected as part of the work for the AssessORF paper. The raw mass spectra data can be found at ProteomeXchange dataset PXD012567. The related genomes used to determine evolutionary conservation all came from the family Streptococcaceae, excluding anomalous genomes and non-complete *Streptococcus pneumoniae* genomes. Links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the MapAssessmentData function from the AssessORF package.

## Results objects

The 4 results objects, 'COH1\_PreSaved\_ResultsObj\_GenBank', 'COH1\_PreSaved\_ResultsObj\_GeneMarkS2', 'COH1\_PreSaved\_ResultsObj\_Glimmer', and 'COH1\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the AssessORF package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("COH1", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX012567>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucleotide/HG939456.1>

---

D_UW_3_CX	Assessment <i>Objects for Chlamydia trachomatis strain D/UW-3/CX</i>
-----------	----------------------------------------------------------------------

---

## Description

Objects of class Assessment and either subclass DataMap or subclass Results for *Chlamydia trachomatis* strain D/UW-3/CX

## Usage

```
data(D_UW_3_CX_PreSaved_DataMapObj)
data(D_UW_3_CX_PreSaved_ResultsObj_GenBank)
data(D_UW_3_CX_PreSaved_ResultsObj_GeneMarkS2)
data(D_UW_3_CX_PreSaved_ResultsObj_Glimmer)
data(D_UW_3_CX_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of Assessment objects contain, please see the **AssessORF** package.

## Details

'D\_UW\_3\_CX\_PreSaved\_DataMapObj' is an object of subclass DataMap.

'D\_UW\_3\_CX\_PreSaved\_ResultsObj\_GenBank' is an object of subclass Results with predicted genes from the GenBank database.

'D\_UW\_3\_CX\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass Results with predicted genes from the program GeneMarkS-2.

'D\_UW\_3\_CX\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass Results with predicted genes from the program Glimmer.

'D\_UW\_3\_CX\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass Results with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: D\_UW\_3\_CX
- Species: C. trachomatis

## Mapping object

The mapping object, 'D\_UW\_3\_CX\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD003883. The related genomes used to determine evolutionary conservation all came from the phylum Chlamydiae, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the MapAssessmentData function from the **AssessORF** package.

## Results objects

The 4 results objects, 'D\_UW\_3\_CX\_PreSaved\_ResultsObj\_GenBank', 'D\_UW\_3\_CX\_PreSaved\_ResultsObj\_GeneMarkS2', 'D\_UW\_3\_CX\_PreSaved\_ResultsObj\_Glimmer', and 'D\_UW\_3\_CX\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the `AssessGenes` function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The `GeneLeftPos`, `GeneRightPos`, and the `GeneStrand` within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The `GeneSource` list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use `SaveGenomeToPath("D_UW_3_CX", <INSERT FILE PATH HERE>)` to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD003883>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucleotide/1273.1>

---

EGD_e	Assessment <i>Objects for Listeria monocytogenes strain EGD-e</i>
-------	-------------------------------------------------------------------

---

## Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Listeria monocytogenes* strain EGD-e

## Usage

```
data(EGD_e_PreSaved_DataMapObj)
data(EGD_e_PreSaved_ResultsObj_GenBank)
data(EGD_e_PreSaved_ResultsObj_GeneMarkS2)
data(EGD_e_PreSaved_ResultsObj_Glimmer)
data(EGD_e_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

## Details

'EGD\_e\_PreSaved\_DataMapObj' is an object of subclass DataMap.

'EGD\_e\_PreSaved\_ResultsObj\_GenBank' is an object of subclass Results with predicted genes from the GenBank database.

'EGD\_e\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass Results with predicted genes from the program GeneMarkS-2.

'EGD\_e\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass Results with predicted genes from the program Glimmer.

'EGD\_e\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass Results with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: EGD\_e
- Species: L. monocytogenes

## Mapping object

The mapping object, 'EGD\_e\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD000890. The related genomes used to determine evolutionary conservation all came from the genus *Listeria*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the MapAssessmentData function from the **AssessORF** package.

## Results objects

The 4 results objects, 'EGD\_e\_PreSaved\_ResultsObj\_GenBank', 'EGD\_e\_PreSaved\_ResultsObj\_GeneMarkS2', 'EGD\_e\_PreSaved\_ResultsObj\_Glimmer', and 'EGD\_e\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("EGD\_e", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX000890>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: [https://www.ncbi.nlm.nih.gov/nucleotide/NC\\_003210.1](https://www.ncbi.nlm.nih.gov/nucleotide/NC_003210.1)

---

`GetDataMapObj`*Get a Data Map Object*

---

**Description**

Gets and returns the data map object for a specific strain

**Usage**

```
GetDataMapObj(strainID)
```

**Arguments**

strainID            Character string corresponding to the strain identifier.

**Details**

GetDataMapObj returns an object of class Assessment and subclass DataMap corresponding to the given strain ID. The given strain ID must be a part of the AssessORF set, and there is no partial matching. Otherwise, the function will error.

**Value**

An object of class Assessment and subclass DataMap

**Examples**

```
mapObj <- GetDataMapObj("MGAS5005")
```

---

`GetGeneSources`*Get the Gene Sources*

---

**Description**

Returns the list of gene sources used in making the results objects

**Usage**

```
GetGeneSources()
```

**Details**

GetGeneSources returns the list of gene sources (programs and databases) used in making the results objects in the AssessORF set.

**Value**

A character vector where each element corresponds to a single gene source

**Examples**

```
geneSourceSet <- GetGeneSources()
```

---

GetResultsObj	<i>Get a Results Object</i>
---------------	-----------------------------

---

**Description**

Gets and returns the results object for a specific strain-gene source combo

**Usage**

```
GetResultsObj(strainID, geneSource = "Prodigal")
```

**Arguments**

strainID	Character string corresponding to the strain identifier.
geneSource	Character string corresponding to the gene source.

**Details**

GetDataMapObj returns an object of class `Assessment` and subclass `Results` corresponding to the given strain ID and the given gene source. The given strain ID must be a part of the `AssessORF` set, and there is no partial matching. Otherwise, the function will error. The given gene source must also be a part of the `AssessORF` set, but the function ignores case when checking if the given gene source is a part of the set.

**Value**

An object of class `Assessment` and subclass `Results`

**Examples**

```
resObj1 <- GetResultsObj("MGAS5005", "Prodigal")
resObj2 <- GetResultsObj("MGAS5005", "GenBank")
resObj3 <- GetResultsObj("MGAS5005", "GeneMarkS2")
resObj4 <- GetResultsObj("MGAS5005", "Glimmer")
```

---

GetStrainIDs

*Get the Strain Identifiers*


---

### Description

Returns the list of strain identifiers for which the package has data

### Usage

```
GetStrainIDs()
```

### Details

GetStrainIDs returns the list of strains in the AssessORF set. This function is a shorter alternative to data("AssessORF\_StrainIDs").

### Value

A character vector where each element corresponds to a single strain identifier

### Examples

```
allStrainIDs <- GetStrainIDs()
```

---

H37Rv

*Assessment Objects for Mycobacterium tuberculosis strain H37Rv*


---

### Description

Objects of class Assessment and either subclass DataMap or subclass Results for *Mycobacterium tuberculosis* strain H37Rv

### Usage

```
data(H37Rv_PreSaved_DataMapObj)
data(H37Rv_PreSaved_ResultsObj_GenBank)
data(H37Rv_PreSaved_ResultsObj_GeneMarkS2)
data(H37Rv_PreSaved_ResultsObj_Glimmer)
data(H37Rv_PreSaved_ResultsObj_Prodigal)
```

### Format

All 5 objects have a list structure. For specifics on what the two types of Assessment objects contain, please see the **AssessORF** package.

## Details

'H37Rv\_PreSaved\_DataMapObj' is an object of subclass DataMap.

'H37Rv\_PreSaved\_ResultsObj\_GenBank' is an object of subclass Results with predicted genes from the GenBank database.

'H37Rv\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass Results with predicted genes from the program GeneMarkS-2.

'H37Rv\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass Results with predicted genes from the program Glimmer.

'H37Rv\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass Results with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: H37Rv
- Species: M. tuberculosis

## Mapping object

The mapping object, 'H37Rv\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD006117. The related genomes used to determine evolutionary conservation all came from the genus Mycobacterium, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the MapAssessmentData function from the **AssessORF** package.

## Results objects

The 4 results objects, 'H37Rv\_PreSaved\_ResultsObj\_GenBank', 'H37Rv\_PreSaved\_ResultsObj\_GeneMarkS2', 'H37Rv\_PreSaved\_ResultsObj\_Glimmer', and 'H37Rv\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("H37Rv", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX006117>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucleotide/AL123456.3>



---

HG001	Assessment <i>Objects for Staphylococcus aureus subsp. aureus strain HG001</i>
-------	--------------------------------------------------------------------------------

---

## Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Staphylococcus aureus* subsp. *aureus* strain HG001

## Usage

```
data(HG001_PreSaved_DataMapObj)
data(HG001_PreSaved_ResultsObj_GenBank)
data(HG001_PreSaved_ResultsObj_GeneMarkS2)
data(HG001_PreSaved_ResultsObj_Glimmer)
data(HG001_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

## Details

'HG001\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'HG001\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'HG001\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'HG001\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'HG001\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: HG001
- Species: *S. aureus*

## Mapping object

The mapping object, 'HG001\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD000702. The related genomes used to determine evolutionary conservation all came from the genus *Staphylococcus*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

## Results objects

The 4 results objects, 'HG001\_PreSaved\_ResultsObj\_GenBank', 'HG001\_PreSaved\_ResultsObj\_GeneMarkS2', 'HG001\_PreSaved\_ResultsObj\_Glimmer', and 'HG001\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the `AssessGenes` function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The `GeneLeftPos`, `GeneRightPos`, and the `GeneStrand` within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The `GeneSource` list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use `SaveGenomeToPath("HG001", <INSERT FILE PATH HERE>)` to save the genome sequence for the strain to your local machine. Note that this genome comes from accession CP018205.1, which has the same genome sequence as accession NZ\_CP018205.1. Unlike NZ\_CP018205.1 however, CP018205.1 does not have any associated GenBank genes.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD000702>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: [https://www.ncbi.nlm.nih.gov/nucleotide/NZ\\_CP018205.1](https://www.ncbi.nlm.nih.gov/nucleotide/NZ_CP018205.1)

---

Houston\_1

---

Assessment *Objects for Bartonella henselae strain Houston-1*


---

## Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Bartonella henselae* strain Houston-1

## Usage

```
data(Houston_1_PreSaved_DataMapObj)
data(Houston_1_PreSaved_ResultsObj_GenBank)
data(Houston_1_PreSaved_ResultsObj_GeneMarkS2)
data(Houston_1_PreSaved_ResultsObj_Glimmer)
data(Houston_1_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

## Details

'Houston\_1\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'Houston\_1\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'Houston\_1\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'Houston\_1\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'Houston\_1\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: Houston\_1
- Species: *B. henselae*

## Mapping object

The mapping object, 'Houston\_1\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD000153. The related genomes used to determine evolutionary conservation came from the families Bartonellaceae, Brucellaceae, Phyllobacteriaceae, Rhizobiaceae. Links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

## Results objects

The 4 results objects, 'Houston\_1\_PreSaved\_ResultsObj\_GenBank', 'Houston\_1\_PreSaved\_ResultsObj\_GeneMarkS2', 'Houston\_1\_PreSaved\_ResultsObj\_Glimmer', and 'Houston\_1\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the `AssessGenes` function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The `GeneLeftPos`, `GeneRightPos`, and the `GeneStrand` within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The `GeneSource` list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use `SaveGenomeToPath("Houston_1", <INSERT FILE PATH HERE>)` to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX000153>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucleotide/BX897699.1>

II1403

Assessment *Objects for Lactococcus lactis subsp. lactis strain II1403***Description**

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Lactococcus lactis* subsp. *lactis* strain II1403

**Usage**

```
data(II1403_PreSaved_DataMapObj)
data(II1403_PreSaved_ResultsObj_GenBank)
data(II1403_PreSaved_ResultsObj_GeneMarkS2)
data(II1403_PreSaved_ResultsObj_Glimmer)
data(II1403_PreSaved_ResultsObj_Prodigal)
```

**Format**

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

**Details**

'II1403\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'II1403\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'II1403\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'II1403\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'II1403\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: II1403
- Species: *L. lactis*

**Mapping object**

The mapping object, 'II1403\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD000494. The related genomes used to determine evolutionary conservation all came from the family, Streptococcaceae excluding anomalous genomes and non-complete *Streptococcus pneumoniae* genomes. Links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

## Results objects

The 4 results objects, 'I11403\_PreSaved\_ResultsObj\_GenBank', 'I11403\_PreSaved\_ResultsObj\_GeneMarkS2', 'I11403\_PreSaved\_ResultsObj\_Glimmer', and 'I11403\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the `AssessGenes` function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The `GeneLeftPos`, `GeneRightPos`, and the `GeneStrand` within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The `GeneSource` list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use `SaveGenomeToPath("I11403", <INSERT FILE PATH HERE>)` to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD000494>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucleotide/AE005176.1>

---

K_12_MG1655	Assessment <i>Objects for Escherichia coli strain K-12 substrain MG1655</i>
-------------	-----------------------------------------------------------------------------

---

## Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Escherichia coli* strain K-12 substrain MG1655

## Usage

```
data(K_12_MG1655_PreSaved_DataMapObj)
data(K_12_MG1655_PreSaved_ResultsObj_GenBank)
data(K_12_MG1655_PreSaved_ResultsObj_GeneMarkS2)
data(K_12_MG1655_PreSaved_ResultsObj_Glimmer)
data(K_12_MG1655_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

## Details

'K\_12\_MG1655\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'K\_12\_MG1655\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'K\_12\_MG1655\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'K\_12\_MG1655\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'K\_12\_MG1655\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: K\_12\_MG1655
- Species: E. coli

## Mapping object

The mapping object, 'K\_12\_MG1655\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD005901. The related genomes used to determine evolutionary conservation all came from the genus *Escherichia*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

## Results objects

The 4 results objects, 'K\_12\_MG1655\_PreSaved\_ResultsObj\_GenBank', 'K\_12\_MG1655\_PreSaved\_ResultsObj\_GeneMarkS2', 'K\_12\_MG1655\_PreSaved\_ResultsObj\_Glimmer', and 'K\_12\_MG1655\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the `AssessGenes` function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GeneMarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The `GeneLeftPos`, `GeneRightPos`, and the `GeneStrand` within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The `GeneSource` list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use `SaveGenomeToPath("K_12_MG1655", <INSERT FILE PATH HERE>)` to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX005901>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucleotide/CP025268.1>

---

LAL14\_1Assessment *Objects for Sulfolobus islandicus strain LAL14/1*

---

## Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Sulfolobus islandicus* strain LAL14/1

## Usage

```
data(LAL14_1_PreSaved_DataMapObj)
data(LAL14_1_PreSaved_ResultsObj_GenBank)
data(LAL14_1_PreSaved_ResultsObj_GeneMarkS2)
data(LAL14_1_PreSaved_ResultsObj_Glimmer)
data(LAL14_1_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

## Details

'LAL14\_1\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'LAL14\_1\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'LAL14\_1\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'LAL14\_1\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'LAL14\_1\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: LAL14\_1
- Species: *S. islandicus*

## Mapping object

The mapping object, 'LAL14\_1\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD003074. The related genomes used to determine evolutionary conservation all came from the phylum Crenarchaeota, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

## Results objects

The 4 results objects, 'LAL14\_1\_PreSaved\_ResultsObj\_GenBank', 'LAL14\_1\_PreSaved\_ResultsObj\_GeneMarkS2', 'LAL14\_1\_PreSaved\_ResultsObj\_Glimmer', and 'LAL14\_1\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the `AssessGenes` function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The `GeneLeftPos`, `GeneRightPos`, and the `GeneStrand` within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The `GeneSource` list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use `SaveGenomeToPath("LAL14_1", <INSERT FILE PATH HERE>)` to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD003074>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucore/CP003928.1>

---

MG1363	Assessment <i>Objects</i> for <i>Lactococcus lactis subsp. cremoris strain MG1363</i>
--------	---------------------------------------------------------------------------------------

---

## Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Lactococcus lactis* subsp. *cremoris* strain MG1363

## Usage

```
data(MG1363_PreSaved_DataMapObj)
data(MG1363_PreSaved_ResultsObj_GenBank)
data(MG1363_PreSaved_ResultsObj_GeneMarkS2)
data(MG1363_PreSaved_ResultsObj_Glimmer)
data(MG1363_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.



## Details

'MG1363\_PreSaved\_DataMapObj' is an object of subclass DataMap.

'MG1363\_PreSaved\_ResultsObj\_GenBank' is an object of subclass Results with predicted genes from the GenBank database.

'MG1363\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass Results with predicted genes from the program GeneMarkS-2.

'MG1363\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass Results with predicted genes from the program Glimmer.

'MG1363\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass Results with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: MG1363
- Species: *L. lactis*

## Mapping object

The mapping object, 'MG1363\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD011263. The related genomes used to determine evolutionary conservation all came from the family, Streptococcaceae excluding anomalous genomes and non-complete *Streptococcus pneumoniae* genomes. Links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the MapAssessmentData function from the **AssessORF** package.

## Results objects

The 4 results objects, 'MG1363\_PreSaved\_ResultsObj\_GenBank', 'MG1363\_PreSaved\_ResultsObj\_GeneMarkS2', 'MG1363\_PreSaved\_ResultsObj\_Glimmer', and 'MG1363\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("MG1363", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD011263>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nuccore/AM406671.1>

---

MGAS5005	Assessment <i>Objects for Streptococcus pyogenes strain MGAS5005</i>
----------	----------------------------------------------------------------------

---

## Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Streptococcus pyogenes* strain MGAS5005

## Usage

```
data(MGAS5005_PreSaved_DataMapObj)
data(MGAS5005_PreSaved_ResultsObj_GenBank)
data(MGAS5005_PreSaved_ResultsObj_GeneMarkS2)
data(MGAS5005_PreSaved_ResultsObj_Glimmer)
data(MGAS5005_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

## Details

'MGAS5005\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'MGAS5005\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'MGAS5005\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'MGAS5005\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'MGAS5005\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: MGAS5005
- Species: *S. pyogenes*

## Mapping object

The mapping object, 'MGAS5005\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The proteomics data was generated in-house and collected as part of the work for the AssessORF paper. The raw mass spectra data can be found at ProteomeXchange dataset PXD012568. The related genomes used to determine evolutionary conservation all came from the family, Streptococcaceae excluding anomalous genomes and non-complete *Streptococcus pneumoniae* genomes. Links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the MapAssessmentData function from the **AssessORF** package.

## Results objects

The 4 results objects, 'MGAS5005\_PreSaved\_ResultsObj\_GenBank', 'MGAS5005\_PreSaved\_ResultsObj\_GeneMarkS2', 'MGAS5005\_PreSaved\_ResultsObj\_Glimmer', and 'MGAS5005\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("MGAS5005", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX012568>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucleotide/CP000017.2>

---

NCIB_3610	Assessment Objects for <i>Bacillus subtilis</i> subsp. <i>subtilis</i> strain NCIB 3610
-----------	-----------------------------------------------------------------------------------------

---

## Description

Objects of class Assessment and either subclass DataMap or subclass Results for *Bacillus subtilis* subsp. *subtilis* strain NCIB 3610

## Usage

```
data(NCIB_3610_PreSaved_DataMapObj)
data(NCIB_3610_PreSaved_ResultsObj_GenBank)
data(NCIB_3610_PreSaved_ResultsObj_GeneMarkS2)
data(NCIB_3610_PreSaved_ResultsObj_Glimmer)
data(NCIB_3610_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of Assessment objects contain, please see the **AssessORF** package.

## Details

'NCIB\_3610\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'NCIB\_3610\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'NCIB\_3610\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'NCIB\_3610\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'NCIB\_3610\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: NCIB\_3610
- Species: *B. subtilis*

## Mapping object

The mapping object, 'NCIB\_3610\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD006444. The related genomes used to determine evolutionary conservation all came from the genus *Bacillus*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

## Results objects

The 4 results objects, 'NCIB\_3610\_PreSaved\_ResultsObj\_GenBank', 'NCIB\_3610\_PreSaved\_ResultsObj\_GeneMarkS2', 'NCIB\_3610\_PreSaved\_ResultsObj\_Glimmer', and 'NCIB\_3610\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the `AssessGenes` function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The `GeneLeftPos`, `GeneRightPos`, and the `GeneStrand` within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

### Getting the strain's genome

Use `SaveGenomeToPath("NCIB_3610", <INSERT FILE PATH HERE>)` to save the genome sequence for the strain to your local machine.

### Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD006444>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: [https://www.ncbi.nlm.nih.gov/nucore/NZ\\_CM000488.1](https://www.ncbi.nlm.nih.gov/nucore/NZ_CM000488.1)

---

PAO1

Assessment *Objects for Pseudomonas aeruginosa strain PAO1*

---

### Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Pseudomonas aeruginosa* strain PAO1

### Usage

```
data(PAO1_PreSaved_DataMapObj)
data(PAO1_PreSaved_ResultsObj_GenBank)
data(PAO1_PreSaved_ResultsObj_GeneMarkS2)
data(PAO1_PreSaved_ResultsObj_Glimmer)
data(PAO1_PreSaved_ResultsObj_Prodigal)
```

### Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

### Details

'PAO1\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'PAO1\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'PAO1\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'PAO1\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'PAO1\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: PAO1
- Species: *P. aeruginosa*

## Mapping object

The mapping object, 'PAO1\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD004560. The related genomes used to determine evolutionary conservation all came from the genus *Pseudomonas*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the MapAssessmentData function from the **AssessORF** package.

## Results objects

The 4 results objects, 'PAO1\_PreSaved\_ResultsObj\_GenBank', 'PAO1\_PreSaved\_ResultsObj\_GeneMarkS2', 'PAO1\_PreSaved\_ResultsObj\_Glimmer', and 'PAO1\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("PAO1", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX004560>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucleotide/AF004091.2>

---

SaveGenomeToPath

*Save a Strain's Genome to a Directory*

---

## Description

Saves the genome for a specified strain to a given directory

## Usage

```
SaveGenomeToPath(strainID, filePath)
```

## Arguments

strainID	Character string corresponding to the strain identifier.
filePath	Character string corresponding to the path to the file path. Must end in '.fasta'.

## Details

SaveGenomeToPath saves the genome for the specified strain ID to the given file path. If the file specified by the path already exists, it will be overwritten (with a warning). The given strain ID must be a part of the AssessORF set, and there is no partial matching. Otherwise, the function will error.

Note: there is no genome for strain ATCC700084, so a reference genome for the species (strain MC2155) is used instead.

## Value

Invisibly returns filePath

## Examples

```
tmpFile <- paste0(tempfile(), ".fasta")
SaveGenomeToPath("MGAS5005", tmpFile)
unlink(tmpFile)
```

---

SL1344	Assessment <i>Objects for Salmonella enterica subsp. enterica serovar Typhimurium strain SL1344</i>
--------	-----------------------------------------------------------------------------------------------------

---

## Description

Objects of class Assessment and either subclass DataMap or subclass Results for *Salmonella enterica* subsp. *enterica* serovar Typhimurium strain SL1344

## Usage

```
data(SL1344_PreSaved_DataMapObj)
data(SL1344_PreSaved_ResultsObj_GenBank)
data(SL1344_PreSaved_ResultsObj_GeneMarkS2)
data(SL1344_PreSaved_ResultsObj_Glimmer)
data(SL1344_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of Assessment objects contain, please see the **AssessORF** package.

## Details

'SL1344\_PreSaved\_DataMapObj' is an object of subclass DataMap.

'SL1344\_PreSaved\_ResultsObj\_GenBank' is an object of subclass Results with predicted genes from the GenBank database.

'SL1344\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass Results with predicted genes from the program GeneMarkS-2.

'SL1344\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass Results with predicted genes from the program Glimmer.

'SL1344\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: SL1344
- Species: *S. typhimurium*

### Mapping object

The mapping object, 'SL1344\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD005579. The related genomes used to determine evolutionary conservation all came from the genus *Salmonella*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

### Results objects

The 4 results objects, 'SL1344\_PreSaved\_ResultsObj\_GenBank', 'SL1344\_PreSaved\_ResultsObj\_GeneMarkS2', 'SL1344\_PreSaved\_ResultsObj\_Glimmer', and 'SL1344\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the `AssessGenes` function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The `GeneLeftPos`, `GeneRightPos`, and the `GeneStrand` within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The `GeneSource` list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

### Getting the strain's genome

Use `SaveGenomeToPath("SL1344", <INSERT FILE PATH HERE>)` to save the genome sequence for the strain to your local machine.

### Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX005579>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucleotide/FQ312003.1>



Strain10403S

Assessment *Objects for Listeria monocytogenes strain 10403S*

## Description

Objects of class Assessment and either subclass DataMap or subclass Results for *Listeria monocytogenes* strain 10403S

## Usage

```
data(Strain10403S_PreSaved_DataMapObj)
data(Strain10403S_PreSaved_ResultsObj_GenBank)
data(Strain10403S_PreSaved_ResultsObj_GeneMarkS2)
data(Strain10403S_PreSaved_ResultsObj_Glimmer)
data(Strain10403S_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of Assessment objects contain, please see the **AssessORF** package.

## Details

'Strain10403S\_PreSaved\_DataMapObj' is an object of subclass DataMap.

'Strain10403S\_PreSaved\_ResultsObj\_GenBank' is an object of subclass Results with predicted genes from the GenBank database.

'Strain10403S\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass Results with predicted genes from the program GeneMarkS-2.

'Strain10403S\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass Results with predicted genes from the program Glimmer.

'Strain10403S\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass Results with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: Strain10403S
- Species: *L. monocytogenes*

## Mapping object

The mapping object, 'Strain10403S\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD010000. The related genomes used to determine evolutionary conservation all came from the genus *Listeria*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the MapAssessmentData function from the **AssessORF** package.

## Results objects

The 4 results objects, 'Strain10403S\_PreSaved\_ResultsObj\_GenBank', 'Strain10403S\_PreSaved\_ResultsObj\_GeneMarkS2', 'Strain10403S\_PreSaved\_ResultsObj\_Glimmer', and 'Strain10403S\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("Strain10403S", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX010000>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: [https://www.ncbi.nlm.nih.gov/nucleotide/NC\\_017544.1](https://www.ncbi.nlm.nih.gov/nucleotide/NC_017544.1)

---

Strain168

Assessment *Objects for Bacillus subtilis subsp. subtilis strain 168*

---

## Description

Objects of class Assessment and either subclass DataMap or subclass Results for *Bacillus subtilis* subsp. *subtilis* strain 168

## Usage

```
data(Strain168_PreSaved_DataMapObj)
data(Strain168_PreSaved_ResultsObj_GenBank)
data(Strain168_PreSaved_ResultsObj_GeneMarkS2)
data(Strain168_PreSaved_ResultsObj_Glimmer)
data(Strain168_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of Assessment objects contain, please see the **AssessORF** package.

## Details

'Strain168\_PreSaved\_DataMapObj' is an object of subclass DataMap.

'Strain168\_PreSaved\_ResultsObj\_GenBank' is an object of subclass Results with predicted genes from the GenBank database.

'Strain168\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass Results with predicted genes from the program GeneMarkS-2.

'Strain168\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass Results with predicted genes from the program Glimmer.

'Strain168\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass Results with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: Strain168
- Species: B. subtilis

## Mapping object

The mapping object, 'Strain168\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The raw proteomics mass spectra data came from ProteomeXchange dataset PXD004565. The related genomes used to determine evolutionary conservation all came from the genus *Bacillus*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the MapAssessmentData function from the **AssessORF** package.

## Results objects

The 4 results objects, 'Strain168\_PreSaved\_ResultsObj\_GenBank', 'Strain168\_PreSaved\_ResultsObj\_GeneMarkS2', 'Strain168\_PreSaved\_ResultsObj\_Glimmer', and 'Strain168\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("Strain168", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX004565>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucleotide/CM000487.1>

---

TCH1516	Assessment <i>Objects for Staphylococcus aureus subsp. aureus strain USA300_TCH1516</i>
---------	-----------------------------------------------------------------------------------------

---

## Description

Objects of class `Assessment` and either subclass `DataMap` or subclass `Results` for *Staphylococcus aureus* subsp. *aureus* strain USA300\_TCH1516

## Usage

```
data(TCH1516_PreSaved_DataMapObj)
data(TCH1516_PreSaved_ResultsObj_GenBank)
data(TCH1516_PreSaved_ResultsObj_GeneMarkS2)
data(TCH1516_PreSaved_ResultsObj_Glimmer)
data(TCH1516_PreSaved_ResultsObj_Prodigal)
```

## Format

All 5 objects have a list structure. For specifics on what the two types of `Assessment` objects contain, please see the **AssessORF** package.

## Details

'TCH1516\_PreSaved\_DataMapObj' is an object of subclass `DataMap`.

'TCH1516\_PreSaved\_ResultsObj\_GenBank' is an object of subclass `Results` with predicted genes from the GenBank database.

'TCH1516\_PreSaved\_ResultsObj\_GeneMarkS2' is an object of subclass `Results` with predicted genes from the program GeneMarkS-2.

'TCH1516\_PreSaved\_ResultsObj\_Glimmer' is an object of subclass `Results` with predicted genes from the program Glimmer.

'TCH1516\_PreSaved\_ResultsObj\_Prodigal' is an object of subclass `Results` with predicted genes from the program Prodigal.

For all 5 objects, values of two key identifying list items within each object are listed below.

- StrainID: TCH1516
- Species: *S. aureus*

## Mapping object

The mapping object, 'TCH1516\_PreSaved\_DataMapObj', stores the mapping of proteomics evidence and evolutionary conservation evidence to the strain's genome. The proteomics data was generated in-house and collected as part of the work for the AssessORF paper. The raw mass spectra data can be found at ProteomeXchange dataset PXD012538. The related genomes used to determine evolutionary conservation all came from the genus *Staphylococcus*, and links to their sequences were downloaded from NCBI's Genome Browser. The object was built using the `MapAssessmentData` function from the **AssessORF** package.

## Results objects

The 4 results objects, 'TCH1516\_PreSaved\_ResultsObj\_GenBank', 'TCH1516\_PreSaved\_ResultsObj\_GeneMarkS2', 'TCH1516\_PreSaved\_ResultsObj\_Glimmer', and 'TCH1516\_PreSaved\_ResultsObj\_Prodigal', store how much evidence there is supporting or against each gene in a set of predicted genes for the strain's genome. The four objects were built using the AssessGenes function from the **AssessORF** package, a set of genes, and the mapping object described here.

For the GenBank object, the genes were downloaded from the corresponding record in the GenBank database. For the other three objects, GenemarkS-2 (web server), Prodigal (2.6.3), and Glimmer (3.02) were run at default settings in order to generate the set of predicted genes for the corresponding results object.

The GeneLeftPos, GeneRightPos, and the GeneStrand within each of the four results objects provide positional information on the set of predicted genes used to generate that object.

The GeneSource list item within each of the four results objects describes where the predicted genes came from (either "GenBank", "GeneMarkS2", "Glimmer", or "Prodigal" respectively).

## Getting the strain's genome

Use SaveGenomeToPath("TCH1516", <INSERT FILE PATH HERE>) to save the genome sequence for the strain to your local machine.

## Source

Proteomics data: <http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD012538>

NCBI's Genome Browser: <https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/>

GenBank record: <https://www.ncbi.nlm.nih.gov/nucore/CP000730.1>

# Index

## \* datasets

- AssessORF\_StrainIDs, [4](#)
- AP1, [2](#)
- AP1\_PreSaved\_DataMapObj (AP1), [2](#)
- AP1\_PreSaved\_ResultsObj\_GenBank (AP1), [2](#)
- AP1\_PreSaved\_ResultsObj\_GeneMarkS2 (AP1), [2](#)
- AP1\_PreSaved\_ResultsObj\_Glimmer (AP1), [2](#)
- AP1\_PreSaved\_ResultsObj\_Prodigal (AP1), [2](#)
- AssessORF\_StrainIDs, [4](#)
- ATCC11842, [4](#)
- ATCC11842\_PreSaved\_DataMapObj (ATCC11842), [4](#)
- ATCC11842\_PreSaved\_ResultsObj\_GenBank (ATCC11842), [4](#)
- ATCC11842\_PreSaved\_ResultsObj\_GeneMarkS2 (ATCC11842), [4](#)
- ATCC11842\_PreSaved\_ResultsObj\_Glimmer (ATCC11842), [4](#)
- ATCC11842\_PreSaved\_ResultsObj\_Prodigal (ATCC11842), [4](#)
- ATCC13032, [6](#)
- ATCC13032\_PreSaved\_DataMapObj (ATCC13032), [6](#)
- ATCC13032\_PreSaved\_ResultsObj\_GenBank (ATCC13032), [6](#)
- ATCC13032\_PreSaved\_ResultsObj\_GeneMarkS2 (ATCC13032), [6](#)
- ATCC13032\_PreSaved\_ResultsObj\_Glimmer (ATCC13032), [6](#)
- ATCC13032\_PreSaved\_ResultsObj\_Prodigal (ATCC13032), [6](#)
- ATCC17978, [7](#)
- ATCC17978\_PreSaved\_DataMapObj (ATCC17978), [7](#)
- ATCC17978\_PreSaved\_ResultsObj\_GenBank (ATCC17978), [7](#)
- ATCC17978\_PreSaved\_ResultsObj\_GeneMarkS2 (ATCC17978), [7](#)
- ATCC17978\_PreSaved\_ResultsObj\_Glimmer (ATCC17978), [7](#)
- ATCC17978\_PreSaved\_ResultsObj\_Prodigal (ATCC17978), [7](#)
- ATCC700084, [9](#)
- ATCC700084\_PreSaved\_DataMapObj (ATCC700084), [9](#)
- ATCC700084\_PreSaved\_ResultsObj\_GenBank (ATCC700084), [9](#)
- ATCC700084\_PreSaved\_ResultsObj\_GeneMarkS2 (ATCC700084), [9](#)
- ATCC700084\_PreSaved\_ResultsObj\_Glimmer (ATCC700084), [9](#)
- ATCC700084\_PreSaved\_ResultsObj\_Prodigal (ATCC700084), [9](#)
- BW25113, [10](#)
- BW25113\_PreSaved\_DataMapObj (BW25113), [10](#)
- BW25113\_PreSaved\_ResultsObj\_GenBank (BW25113), [10](#)
- BW25113\_PreSaved\_ResultsObj\_GeneMarkS2 (BW25113), [10](#)
- BW25113\_PreSaved\_ResultsObj\_Glimmer (BW25113), [10](#)
- BW25113\_PreSaved\_ResultsObj\_Prodigal (BW25113), [10](#)
- CCMP1375, [12](#)
- CCMP1375\_PreSaved\_DataMapObj (CCMP1375), [12](#)
- CCMP1375\_PreSaved\_ResultsObj\_GenBank (CCMP1375), [12](#)
- CCMP1375\_PreSaved\_ResultsObj\_GeneMarkS2 (CCMP1375), [12](#)
- CCMP1375\_PreSaved\_ResultsObj\_Glimmer (CCMP1375), [12](#)
- CCMP1375\_PreSaved\_ResultsObj\_Prodigal (CCMP1375), [12](#)
- CECT5344, [13](#)
- CECT5344\_PreSaved\_DataMapObj (CECT5344), [13](#)
- CECT5344\_PreSaved\_ResultsObj\_GenBank (CECT5344), [13](#)
- CECT5344\_PreSaved\_ResultsObj\_GeneMarkS2 (CECT5344), [13](#)

- CECT5344\_PreSaved\_ResultsObj\_Glimmer (CECT5344), [13](#)
- CECT5344\_PreSaved\_ResultsObj\_Prodigal (CECT5344), [13](#)
- CNRZ327, [15](#)
- CNRZ327\_PreSaved\_DataMapObj (CNRZ327), [15](#)
- CNRZ327\_PreSaved\_ResultsObj\_GenBank (CNRZ327), [15](#)
- CNRZ327\_PreSaved\_ResultsObj\_GeneMarkS2 (CNRZ327), [15](#)
- CNRZ327\_PreSaved\_ResultsObj\_Glimmer (CNRZ327), [15](#)
- CNRZ327\_PreSaved\_ResultsObj\_Prodigal (CNRZ327), [15](#)
- COH1, [16](#)
- COH1\_PreSaved\_DataMapObj (COH1), [16](#)
- COH1\_PreSaved\_ResultsObj\_GenBank (COH1), [16](#)
- COH1\_PreSaved\_ResultsObj\_GeneMarkS2 (COH1), [16](#)
- COH1\_PreSaved\_ResultsObj\_Glimmer (COH1), [16](#)
- COH1\_PreSaved\_ResultsObj\_Prodigal (COH1), [16](#)
- D\_UW\_3\_CX, [18](#)
- D\_UW\_3\_CX\_PreSaved\_DataMapObj (D\_UW\_3\_CX), [18](#)
- D\_UW\_3\_CX\_PreSaved\_ResultsObj\_GenBank (D\_UW\_3\_CX), [18](#)
- D\_UW\_3\_CX\_PreSaved\_ResultsObj\_GeneMarkS2 (D\_UW\_3\_CX), [18](#)
- D\_UW\_3\_CX\_PreSaved\_ResultsObj\_Glimmer (D\_UW\_3\_CX), [18](#)
- D\_UW\_3\_CX\_PreSaved\_ResultsObj\_Prodigal (D\_UW\_3\_CX), [18](#)
- EGD\_e, [19](#)
- EGD\_e\_PreSaved\_DataMapObj (EGD\_e), [19](#)
- EGD\_e\_PreSaved\_ResultsObj\_GenBank (EGD\_e), [19](#)
- EGD\_e\_PreSaved\_ResultsObj\_GeneMarkS2 (EGD\_e), [19](#)
- EGD\_e\_PreSaved\_ResultsObj\_Glimmer (EGD\_e), [19](#)
- EGD\_e\_PreSaved\_ResultsObj\_Prodigal (EGD\_e), [19](#)
- GetDataMapObj, [21](#)
- GetGeneSources, [21](#)
- GetResultsObj, [22](#)
- GetStrainIDs, [23](#)
- H37Rv, [23](#)
- H37Rv\_PreSaved\_DataMapObj (H37Rv), [23](#)
- H37Rv\_PreSaved\_ResultsObj\_GenBank (H37Rv), [23](#)
- H37Rv\_PreSaved\_ResultsObj\_GeneMarkS2 (H37Rv), [23](#)
- H37Rv\_PreSaved\_ResultsObj\_Glimmer (H37Rv), [23](#)
- H37Rv\_PreSaved\_ResultsObj\_Prodigal (H37Rv), [23](#)
- HG001, [25](#)
- HG001\_PreSaved\_DataMapObj (HG001), [25](#)
- HG001\_PreSaved\_ResultsObj\_GenBank (HG001), [25](#)
- HG001\_PreSaved\_ResultsObj\_GeneMarkS2 (HG001), [25](#)
- HG001\_PreSaved\_ResultsObj\_Glimmer (HG001), [25](#)
- HG001\_PreSaved\_ResultsObj\_Prodigal (HG001), [25](#)
- Houston\_1, [26](#)
- Houston\_1\_PreSaved\_DataMapObj (Houston\_1), [26](#)
- Houston\_1\_PreSaved\_ResultsObj\_GenBank (Houston\_1), [26](#)
- Houston\_1\_PreSaved\_ResultsObj\_GeneMarkS2 (Houston\_1), [26](#)
- Houston\_1\_PreSaved\_ResultsObj\_Glimmer (Houston\_1), [26](#)
- Houston\_1\_PreSaved\_ResultsObj\_Prodigal (Houston\_1), [26](#)
- I11403, [28](#)
- I11403\_PreSaved\_DataMapObj (I11403), [28](#)
- I11403\_PreSaved\_ResultsObj\_GenBank (I11403), [28](#)
- I11403\_PreSaved\_ResultsObj\_GeneMarkS2 (I11403), [28](#)
- I11403\_PreSaved\_ResultsObj\_Glimmer (I11403), [28](#)
- I11403\_PreSaved\_ResultsObj\_Prodigal (I11403), [28](#)
- K\_12\_MG1655, [29](#)
- K\_12\_MG1655\_PreSaved\_DataMapObj (K\_12\_MG1655), [29](#)
- K\_12\_MG1655\_PreSaved\_ResultsObj\_GenBank (K\_12\_MG1655), [29](#)
- K\_12\_MG1655\_PreSaved\_ResultsObj\_GeneMarkS2 (K\_12\_MG1655), [29](#)
- K\_12\_MG1655\_PreSaved\_ResultsObj\_Glimmer (K\_12\_MG1655), [29](#)

- K\_12\_MG1655\_PreSaved\_ResultsObj\_Prodigal (K\_12\_MG1655), [29](#)
- LAL14\_1, [31](#)
- LAL14\_1\_PreSaved\_DataMapObj (LAL14\_1), [31](#)
- LAL14\_1\_PreSaved\_ResultsObj\_GenBank (LAL14\_1), [31](#)
- LAL14\_1\_PreSaved\_ResultsObj\_GeneMarkS2 (LAL14\_1), [31](#)
- LAL14\_1\_PreSaved\_ResultsObj\_Glimmer (LAL14\_1), [31](#)
- LAL14\_1\_PreSaved\_ResultsObj\_Prodigal (LAL14\_1), [31](#)
- MG1363, [32](#)
- MG1363\_PreSaved\_DataMapObj (MG1363), [32](#)
- MG1363\_PreSaved\_ResultsObj\_GenBank (MG1363), [32](#)
- MG1363\_PreSaved\_ResultsObj\_GeneMarkS2 (MG1363), [32](#)
- MG1363\_PreSaved\_ResultsObj\_Glimmer (MG1363), [32](#)
- MG1363\_PreSaved\_ResultsObj\_Prodigal (MG1363), [32](#)
- MGAS5005, [34](#)
- MGAS5005\_PreSaved\_DataMapObj (MGAS5005), [34](#)
- MGAS5005\_PreSaved\_ResultsObj\_GenBank (MGAS5005), [34](#)
- MGAS5005\_PreSaved\_ResultsObj\_GeneMarkS2 (MGAS5005), [34](#)
- MGAS5005\_PreSaved\_ResultsObj\_Glimmer (MGAS5005), [34](#)
- MGAS5005\_PreSaved\_ResultsObj\_Prodigal (MGAS5005), [34](#)
- NCIB\_3610, [35](#)
- NCIB\_3610\_PreSaved\_DataMapObj (NCIB\_3610), [35](#)
- NCIB\_3610\_PreSaved\_ResultsObj\_GenBank (NCIB\_3610), [35](#)
- NCIB\_3610\_PreSaved\_ResultsObj\_GeneMarkS2 (NCIB\_3610), [35](#)
- NCIB\_3610\_PreSaved\_ResultsObj\_Glimmer (NCIB\_3610), [35](#)
- NCIB\_3610\_PreSaved\_ResultsObj\_Prodigal (NCIB\_3610), [35](#)
- PA01, [37](#)
- PA01\_PreSaved\_DataMapObj (PA01), [37](#)
- PA01\_PreSaved\_ResultsObj\_GenBank (PA01), [37](#)
- PA01\_PreSaved\_ResultsObj\_GeneMarkS2 (PA01), [37](#)
- PA01\_PreSaved\_ResultsObj\_Glimmer (PA01), [37](#)
- PA01\_PreSaved\_ResultsObj\_Prodigal (PA01), [37](#)
- SaveGenomeToPath, [38](#)
- SL1344, [39](#)
- SL1344\_PreSaved\_DataMapObj (SL1344), [39](#)
- SL1344\_PreSaved\_ResultsObj\_GenBank (SL1344), [39](#)
- SL1344\_PreSaved\_ResultsObj\_GeneMarkS2 (SL1344), [39](#)
- SL1344\_PreSaved\_ResultsObj\_Glimmer (SL1344), [39](#)
- SL1344\_PreSaved\_ResultsObj\_Prodigal (SL1344), [39](#)
- Strain10403S, [41](#)
- Strain10403S\_PreSaved\_DataMapObj (Strain10403S), [41](#)
- Strain10403S\_PreSaved\_ResultsObj\_GenBank (Strain10403S), [41](#)
- Strain10403S\_PreSaved\_ResultsObj\_GeneMarkS2 (Strain10403S), [41](#)
- Strain10403S\_PreSaved\_ResultsObj\_Glimmer (Strain10403S), [41](#)
- Strain10403S\_PreSaved\_ResultsObj\_Prodigal (Strain10403S), [41](#)
- Strain168, [42](#)
- Strain168\_PreSaved\_DataMapObj (Strain168), [42](#)
- Strain168\_PreSaved\_ResultsObj\_GenBank (Strain168), [42](#)
- Strain168\_PreSaved\_ResultsObj\_GeneMarkS2 (Strain168), [42](#)
- Strain168\_PreSaved\_ResultsObj\_Glimmer (Strain168), [42](#)
- Strain168\_PreSaved\_ResultsObj\_Prodigal (Strain168), [42](#)
- TCH1516, [44](#)
- TCH1516\_PreSaved\_DataMapObj (TCH1516), [44](#)
- TCH1516\_PreSaved\_ResultsObj\_GenBank (TCH1516), [44](#)
- TCH1516\_PreSaved\_ResultsObj\_GeneMarkS2 (TCH1516), [44](#)
- TCH1516\_PreSaved\_ResultsObj\_Glimmer (TCH1516), [44](#)
- TCH1516\_PreSaved\_ResultsObj\_Prodigal (TCH1516), [44](#)