

# Package ‘GeneSummary’

December 10, 2025

**Type** Package

**Title** RefSeq Gene Summaries

**Version** 0.99.6

**Date** 2023-09-23

**Depends** R (>= 4.0)

**Suggests** knitr, rmarkdown, markdown, testthat

**biocViews** AnnotationData, Organism, Homo\_sapiens

**Description** This package provides long description of genes collected from the RefSeq database. The text in ``COMMENT" section started with ``Summary" is extracted as the description of the gene. The long text descriptions can be used for analysis such as text mining.

**VignetteBuilder** knitr

**URL** <https://github.com/jokergoo/GeneSummary>

**BugReports** <https://github.com/jokergoo/GeneSummary/issues>

**License** MIT + file LICENSE

**NeedsCompilation** no

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

**git\_branch** devel

**git\_last\_commit** 4ac4046

**git\_last\_commit\_date** 2023-09-23

**Repository** Bioconductor 3.23

**Date/Publication** 2025-12-10

**Author** Zuguang Gu [aut, cre] (ORCID: <<https://orcid.org/0000-0002-7395-8709>>)

**Maintainer** Zuguang Gu <z.gu@dkfz.de>

## Contents

|                                  |   |
|----------------------------------|---|
| GeneSummary . . . . .            | 2 |
| loadGeneSummary . . . . .        | 2 |
| print.GeneSummary_info . . . . . | 3 |

|       |   |
|-------|---|
| Index | 4 |
|-------|---|

---

|             |                                       |
|-------------|---------------------------------------|
| GeneSummary | <i>Version and source information</i> |
|-------------|---------------------------------------|

---

**Description**

Version and source information

**Usage**

GeneSummary

**Examples**

GeneSummary

---

|                 |                                |
|-----------------|--------------------------------|
| loadGeneSummary | <i>Load gene summary table</i> |
|-----------------|--------------------------------|

---

**Description**

Load gene summary table

**Usage**

```
loadGeneSummary(organism = 9606,  
  status = c("reviewed", "predicted", "inferred", "validated", "provisional"))
```

**Arguments**

- |          |                                                                                       |
|----------|---------------------------------------------------------------------------------------|
| organism | Name of the organism (e.g. "Homo sapiens") or the corresponding taxon ID (e.g. 9606). |
| status   | Status of the review of the genes.                                                    |

**Details**

The gene summaries are extracted from RefSeq database ([https://ftp.ncbi.nih.gov/refseq/release/complete/\\*.rna.gbff.gz](https://ftp.ncbi.nih.gov/refseq/release/complete/*.rna.gbff.gz)). Gene summaries are available in the "COMMENT" section of the \*.rna.gbff.gz files.

The Perl script used to parse \*.rna.gbff.gz files is available in `system.file("extdata", "parse_refseq.pl", package = "GeneSummary")`.

**Value**

A data frame with the following columns:

**RefSeq\_accession** RefSeq accession. Note there is also the version number after the accession.

**Organism** Full organism name.

**Taxon\_ID** Taxon ID of the organism.

**Gene\_ID** Entrez gene ID.

**Review\_status** Status of the review of the gene.

**Gene\_summary** A long description of the gene.

**Examples**

```
df = loadGeneSummary(9606)
dim(df)
head(df)
```

---

```
print.GeneSummary_info
```

*Print the GeneSummary\_info object*

---

**Description**

Print the GeneSummary\_info object

**Usage**

```
## S3 method for class 'GeneSummary_info'
print(x, ...)
```

**Arguments**

|     |                            |
|-----|----------------------------|
| x   | A GeneSummary_info object. |
| ... | Other arguments            |

**Value**

No value is returned.

**Examples**

```
GeneSummary
```

# Index

GeneSummary, [2](#)

loadGeneSummary, [2](#)

print.GeneSummary\_info, [3](#)