

# Package: restez (via r-universe)

October 28, 2024

**Type** Package

**Title** Create and Query a Local Copy of 'GenBank' in R

**Version** 2.1.4.9000

**Maintainer** Joel H. Nitta <joelnitta@gmail.com>

**Description** Download large sections of 'GenBank' <<https://www.ncbi.nlm.nih.gov/genbank/>> and generate a local SQL-based database. A user can then query this database using 'restez' functions or through 'rentrez' <<https://CRAN.R-project.org/package=rentrez>> wrappers.

**URL** <https://github.com/ropensci/restez>,  
<https://docs.ropensci.org/restez/>

**BugReports** <https://github.com/ropensci/restez/issues>

**License** MIT + file LICENSE

**Encoding** UTF-8

**LazyData** true

**Depends** R (>= 3.3.0)

**Imports** utils, rentrez, DBI (>= 1.0.0), curl, cli, crayon, stringi, duckdb, fs, assertthat, ape

**Suggests** sessioninfo, testthat, knitr, R.utils, rmarkdown, mockery

**RoxygenNote** 7.2.3

**Roxygen** list(markdown = TRUE)

**Repository** <https://ropensci.r-universe.dev>

**RemoteUrl** <https://github.com/ropensci/restez>

**RemoteRef** main

**RemoteSha** 1e5c9171206024efd308aede90ba8c60b2d8c876

## Contents

add_rcrd_log	4
cat_line	4
char	5
check_connection	6
cleanup	7
connected	7
connection_get	8
count_db_ids	9
db_create	10
db_delete	12
db_download	12
db_download_intern	14
db_sqlngths_get	15
db_sqlngths_log	15
demo_db_create	16
dir_size	17
dwnld_path_get	18
dwnld_rcrd_log	18
entrez_fasta_get	19
entrez_fetch	20
entrez_gb_get	21
extract_accession	22
extract_by_patterns	23
extract_clean_sequence	24
extract_definition	25
extract_features	26
extract_inforecpart	27
extract_keywords	28
extract_locus	29
extract_organism	30
extract_seqrecpart	31
extract_sequence	32
extract_version	33
filename_log	34
file_download	34
flatfile_read	35
gbrelease_check	36
gbrelease_get	37
gbrelease_log	37
gb_build	38
gb_definition_get	39
gb_df_create	40
gb_df_generate	41
gb_extract	42
gb_fasta_get	43
gb_organism_get	44

gb_record_get . . . . .	45
gb_sequence_get . . . . .	45
gb_sql_add . . . . .	46
gb_sql_query . . . . .	47
gb_version_get . . . . .	48
has_data . . . . .	49
identify_downloadable_files . . . . .	49
is_in_db . . . . .	50
last_add_get . . . . .	51
last_dwnld_get . . . . .	52
last_entry_get . . . . .	52
latest_genbank_release . . . . .	53
latest_genbank_release_notes . . . . .	54
list_db_ids . . . . .	54
message_missing . . . . .	55
mock_def . . . . .	56
mock_gb_df_generate . . . . .	57
mock_org . . . . .	58
mock_rec . . . . .	59
mock_seq . . . . .	60
ncbi_acc_get . . . . .	60
predict_datasizes . . . . .	61
print.status . . . . .	62
readme_log . . . . .	63
record . . . . .	63
restez_connect . . . . .	64
restez_disconnect . . . . .	65
restez_path_check . . . . .	65
restez_path_get . . . . .	66
restez_path_set . . . . .	66
restez_path_unset . . . . .	67
restez_ready . . . . .	68
restez_rl . . . . .	68
restez_status . . . . .	69
search_gz . . . . .	70
seshinfo_log . . . . .	71
setup . . . . .	71
slctn_get . . . . .	72
slctn_log . . . . .	73
sql_path_get . . . . .	73
stat . . . . .	74
status_class . . . . .	75
testdatadir_get . . . . .	75

---

add_rcrd_log	<i>Log files added to the SQL database in the restez path</i>
--------------	---

---

### Description

This function is called whenever sequence files have been successfully added to the nucleotide SQL database. Row entries are added to 'add\_lot.tsv' in the user's restez path containing the filename, GB release numbers and the time of successful adding. The log is to help users keep track of when sequences have been added.

### Usage

```
add_rcrd_log(fl)
```

### Arguments

fl	filename, character
----	---------------------

### See Also

Other private: [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_infocpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

cat_line	<i>Cat lines</i>
----------	------------------

---

### Description

Helper function for printing lines to console. Automatically formats lines by adding newlines.

### Usage

```
cat_line(...)
```

**Arguments**

...                   Text to print, character

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

char

*Print green*

---

**Description**

Print to console green text to indicate a name/filepath/text

**Usage**

char(x)

**Arguments**

x                   Text to print, character

**Value**

coloured character encoding, character

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#),

gbrelease\_get(), gbrelease\_log(), has\_data(), identify\_downloadable\_files(), last\_add\_get(), last\_dwnld\_get(), last\_entry\_get(), latest\_genbank\_release\_notes(), latest\_genbank\_release(), message\_missing(), mock\_def(), mock\_gb\_df\_generate(), mock\_org(), mock\_rec(), mock\_seq(), predict\_datasizes(), readme\_log(), restez\_connect(), restez\_disconnect(), restez\_path\_check(), restez\_rl(), search\_gz(), seshinfo\_log(), setup(), slctn\_get(), slctn\_log(), sql\_path\_get(), status\_class(), stat(), testdatadir\_get()

---

check_connection	<i>Helper function to test if a stable internet connection can be established.</i>
------------------	--

---

### Description

All retrieval functions need a stable internet connection to work properly. This internal function pings the google homepage and throws an error if it cannot be reached.

### Usage

```
check_connection()
```

### Author(s)

Hajk-Georg Drost

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

cleanup	<i>Clean up test data</i>
---------	---------------------------

---

**Description**

Removes all temporary test data created.

**Usage**

```
cleanup()
```

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

connected	<i>Is restez connected?</i>
-----------	-----------------------------

---

**Description**

Returns TRUE if a restez SQL database has been connected.

**Usage**

```
connected()
```

**Value**

Logical

**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dnwld_path_get()`, `dnwld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_add()`, `gb_sql_query()`, `gbrelease_check()`, `gbrelease_get()`, `gbrelease_log()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dnwld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_def()`, `mock_gb_df_generate()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`, `readme_log()`, `restez_connect()`, `restez_disconnect()`, `restez_path_check()`, `restez_rl()`, `search_gz()`, `seshinfo_log()`, `setup()`, `slctn_get()`, `slctn_log()`, `sql_path_get()`, `status_class()`, `stat()`, `testdatadir_get()`

---

connection\_get

*Retrieve restez connection*

---

**Description**

Safely acquire the restez connection. Raises error if no connection set.

**Usage**

```
connection_get()
```

**Value**

connection

**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dnwld_path_get()`, `dnwld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_add()`, `gb_sql_query()`, `gbrelease_check()`, `gbrelease_get()`, `gbrelease_log()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dnwld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_def()`, `mock_gb_df_generate()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`, `readme_log()`, `restez_connect()`, `restez_disconnect()`, `restez_path_check()`, `restez_rl()`, `search_gz()`, `seshinfo_log()`, `setup()`, `slctn_get()`, `slctn_log()`, `sql_path_get()`, `status_class()`, `stat()`, `testdatadir_get()`



---

count_db_ids	<i>Return the number of ids</i>
--------------	---------------------------------

---

### Description

Return the number of ids in a user's restez database.

### Usage

```
count_db_ids(db = "nucleotide")
```

### Arguments

db                    character, database name

### Details

Requires an open connection. If no connection or db 0 is returned.

### Value

integer

### See Also

Other database: [db\\_create\(\)](#), [db\\_delete\(\)](#), [db\\_download\(\)](#), [demo\\_db\\_create\(\)](#), [is\\_in\\_db\(\)](#), [list\\_db\\_ids\(\)](#)

### Examples

```
library(restez)
restez_path_set(filepath = tempdir())
demo_db_create(n = 5)
(count_db_ids())

# delete demo after example
db_delete(everything = TRUE)
```

---

db_create	<i>Create new NCBI database</i>
-----------	---------------------------------

---

### Description

Create a new local SQL database from downloaded files. Currently only GenBank/nucleotide/nuccore database is supported.

### Usage

```
db_create(
  db_type = "nucleotide",
  min_length = 0,
  max_length = NULL,
  acc_filter = NULL,
  invert = FALSE,
  alt_restez_path = NULL,
  scan = FALSE
)
```

### Arguments

db_type	character, database type
min_length	Minimum sequence length, default 0.
max_length	Maximum sequence length, default NULL.
acc_filter	Character vector; accessions to include or exclude from the database as specified by invert.
invert	Logical vector of length 1; if TRUE, accessions in acc_filter will be excluded from the database; if FALSE, only accessions in acc_filter will be included in the database. Default FALSE.
alt_restez_path	Alternative restez path if you would like to use the downloads from a different restez path.
scan	Logical vector of length 1; should the sequence file be scanned for accessions in acc_filter prior to processing? Requires zgrep to be installed (so does not work on Windows). Only used if acc_filter is not NULL and invert is FALSE. Default FALSE.

### Details

All .seq.gz files are added to the database by default. A user can specify minimum/maximum sequence lengths or accession numbers to limit the sequences to be added to the database – smaller databases are faster to search. The final selection of sequences is the result of applying all filters (acc\_filter, min\_length, max\_length) in combination.

The scan option can decrease the time needed to build a database if only a small number of sequences should be written to the database compared to the number of the sequences downloaded

from GenBank; i.e., if many of the files downloaded from GenBank do not contain any sequences that should be written to the database. When set to TRUE, if a file does not contain any of the accessions in `acc_filter`, further processing of that file will be skipped and none of the sequences it contains will be added to the database.

Alternatively, a user can use the `alt_restez_path` to add the files from an alternative `reste` file path. For example, you may wish to have a database of all environmental sequences but then an additional smaller one of just the sequences with lengths below 100 bp. Instead of having to download all environmental sequences twice, you can generate multiple `reste` databases using the same downloaded files from a single `reste` path.

This function will not overwrite a pre-existing database. Old databases must be deleted before a new one can be created. Use `db_delete()` with `everything=FALSE` to delete an SQL database.

Connections/disconnections to the database are made automatically.

### See Also

Other database: [count\\_db\\_ids\(\)](#), [db\\_delete\(\)](#), [db\\_download\(\)](#), [demo\\_db\\_create\(\)](#), [is\\_in\\_db\(\)](#), [list\\_db\\_ids\(\)](#)

### Examples

```
## Not run:
# Example of general usage
library(restez)
reste_path_set(filepath = 'path/for/downloads/and/database')
db_download()
db_create()

# Example of using `acc_filter`
#
# Download files to temporary directory
temp_dir <- paste0(tempdir(), "/reste", collapse = "")
dir.create(temp_dir)
reste_path_set(filepath = temp_dir)
# Choose GenBank domain 20 ('unannotated'), the smallest
db_download(preselection = 20)
# Only include three accessions in database
db_create(
  acc_filter = c("AF000122", "AF000123", "AF000124")
)
list_db_ids()
db_delete()
unlink(temp_dir)

## End(Not run)
```

---

`db_delete`*Delete database*

---

**Description**

Delete the local SQL database and/or restez folder.

**Usage**

```
db_delete(everything = FALSE)
```

**Arguments**

`everything`      T/F, delete the whole restez folder as well?

**Details**

Any connected database will be automatically disconnected.

**See Also**

Other database: [count\\_db\\_ids\(\)](#), [db\\_create\(\)](#), [db\\_download\(\)](#), [demo\\_db\\_create\(\)](#), [is\\_in\\_db\(\)](#), [list\\_db\\_ids\(\)](#)

**Examples**

```
library(restez)
fp <- tempdir()
restez_path_set(filepath = fp)
demo_db_create(n = 10)
db_delete(everything = FALSE)
# Will not run: gb_sequence_get(id = 'demo_1')
# only the SQL database is deleted
db_delete(everything = TRUE)
# Now returns NULL
(restez_path_get())
```

---

`db_download`*Download database*

---

**Description**

Download .seq.tar files from the latest GenBank release.

## Usage

```
db_download(  
  db = "nucleotide",  
  overwrite = FALSE,  
  preselection = NULL,  
  max_tries = 1  
)
```

## Arguments

db	Database type, only 'nucleotide' currently available.
overwrite	T/F, overwrite pre-existing downloaded files?
preselection	Character vector of length 1; GenBank domains to download. If not specified (default), a menu will be provided for selection. To specify, provide either a single number or a single character string of numbers separated by spaces, e.g. "19 20" for 'Phage' (19) and 'Unannotated' (20).
max_tries	Numeric vector of length 1; maximum number of times to attempt downloading database (default 1).

## Details

In default mode, the user interactively selects the parts (i.e., "domains") of GenBank to download (e.g. primates, plants, bacteria ...). Alternatively, the selected domains can be provided as a character string to preselection.

The max\_tries argument is useful for large databases that may otherwise fail due to periodic lapses in internet connectivity. This value can be set to Inf to continuously try until the database download succeeds (not recommended if you do not have an internet connection!).

## Value

T/F, if all files download correctly, TRUE else FALSE.

## See Also

[ncbi\\_acc\\_get\(\)](#)

Other database: [count\\_db\\_ids\(\)](#), [db\\_create\(\)](#), [db\\_delete\(\)](#), [demo\\_db\\_create\(\)](#), [is\\_in\\_db\(\)](#), [list\\_db\\_ids\(\)](#)

## Examples

```
## Not run:  
library(restez)  
restez_path_set(filepath = 'path/for/downloads')  
db_download()  
  
## End(Not run)
```

---

db\_download\_intern      *Download database (internal version)*

---

### Description

Download .seq.tar files from the latest GenBank release. The user interactively selects the parts of GenBank to download (e.g. primates, plants, bacteria ...). This is an internal function so the download can be wrapped in while() to enable persistent downloading.

### Usage

```
db_download_intern(db = "nucleotide", overwrite = FALSE, preselection = NULL)
```

### Arguments

db	Database type, only 'nucleotide' currently available.
overwrite	T/F, overwrite pre-existing downloaded files?
preselection	Character vector of length 1; GenBank domains to download. If not specified (default), a menu will be provided for selection. To specify, provide either a single number or a single character string of numbers separated by spaces, e.g. "19 20" for 'Phage' (19) and 'Unannotated' (20).

### Details

The downloaded files will appear in the restez filepath under downloads.

### Value

T/F, if all files download correctly, TRUE else FALSE.

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

db_sqlngths_get	<i>Return the minimum and maximum sequence lengths in db</i>
-----------------	--

---

**Description**

Returns the maximum and minimum sequence lengths as set by the user upon db creation.

**Usage**

```
db_sqlngths_get()
```

**Details**

If no file found, returns empty character vector.

**Value**

vector of integers

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

db_sqlngths_log	<i>Log the min and max sequence lengths</i>
-----------------	---

---

**Description**

Log the min and maximum sequence length used in the created db.

**Usage**

```
db_sqlngths_log(min_lngth, max_lngth)
```

**Arguments**

min_lngth	Minimum length
max_lngth	Maximum length

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_infocpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

demo_db_create	<i>Create demo database</i>
----------------	-----------------------------

---

**Description**

Creates a local mock SQL database from package test data for demonstration purposes. No internet connection required.

**Usage**

```
demo_db_create(db_type = "nucleotide", n = 100)
```

**Arguments**

db_type	character, database type
n	integer, number of mock sequences

**See Also**

Other database: [count\\_db\\_ids\(\)](#), [db\\_create\(\)](#), [db\\_delete\(\)](#), [db\\_download\(\)](#), [is\\_in\\_db\(\)](#), [list\\_db\\_ids\(\)](#)



**Examples**

```

library(restz)
# set the restz path to a temporary dir
restz_path_set(filepath = tempdir())
# create demo database
demo_db_create(n = 5)
# in the demo, IDs are 'demo_1', 'demo_2' ...
(gb_sequence_get(id = 'demo_1'))

# Delete a demo database after an example
db_delete(everything = TRUE)

```

dir\_size

*Calculate the size of a directory***Description**

Returns the size of directory in GB

**Usage**

```
dir_size(fp)
```

**Arguments**

fp                   File path, character

**Value**

numeric

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_infocpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restz\\_connect\(\)](#), [restz\\_disconnect\(\)](#), [restz\\_path\\_check\(\)](#), [restz\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

dwnld_path_get	<i>Get dwnld path</i>
----------------	-----------------------

---

### Description

Return path to folder where raw .seq files are stored.

### Usage

```
dwnld_path_get()
```

### Value

character

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_infocpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

dwnld_rcrd_log	<i>Log a downloaded file in the restez path</i>
----------------	---

---

### Description

This function is called whenever a file is successfully downloaded. A row entry is added to the 'download\_log.tsv' in the user's restez path containing the file name, the GB release number and the time of successfully download. The log is to help users keep track of when they downloaded files and to determine if the downloaded files are out of date.

### Usage

```
dwnld_rcrd_log(f1)
```

**Arguments**

f1                    file name, character

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

entrez\_fasta\_get            *Get Entrez fasta*

---

**Description**

Return fasta format as expected from an Entrez call. If not all IDs are returned, will run `rentrez::entrez_fetch`.

**Usage**

```
entrez_fasta_get(id, ...)
```

**Arguments**

id                    vector, unique ID(s) for record(s)  
 ...                   arguments passed on to `rentrez`

**Value**

character string containing the file created

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#),

```
extract_version(), file_download(), filename_log(), flatfile_read(), gb_build(), gb_df_create(),
gb_df_generate(), gb_sql_add(), gb_sql_query(), gbrelease_check(), gbrelease_get(),
gbrelease_log(), has_data(), identify_downloadable_files(), last_add_get(), last_dwnld_get(),
last_entry_get(), latest_genbank_release_notes(), latest_genbank_release(), message_missing(),
mock_def(), mock_gb_df_generate(), mock_org(), mock_rec(), mock_seq(), predict_datasizes(),
readme_log(), restez_connect(), restez_disconnect(), restez_path_check(), restez_rl(),
search_gz(), seshinfo_log(), setup(), slctn_get(), slctn_log(), sql_path_get(), status_class(),
stat(), testdatadir_get()
```

---

entrez\_fetch

*Entrez\_fetch*


---

### Description

Wrapper for `rentrez::entrez_fetch`.

### Usage

```
entrez_fetch(db, id = NULL, rettype, retmode = "", ...)
```

### Arguments

<code>db</code>	character, name of the database
<code>id</code>	vector, unique ID(s) for record(s)
<code>rettype</code>	character, data format
<code>retmode</code>	character, data mode
<code>...</code>	Arguments to be passed on to <code>rentrez</code>

### Details

Attempts to first search local database with user-specified parameters, if the record is missing in the database, the function then calls `rentrez::entrez_fetch` to search GenBank remotely.

`rettype='fasta'` and `rettype='gb'` are respectively equivalent to `gb_fasta_get()` and `gb_record_get()`.

### Value

character string containing the file created

### Supported return types and modes

XML `retmode` is not supported. Rettypes `'seqid'`, `'ft'`, `'acc'` and `'uilst'` are also not supported.

### Note

It is advisable to call `restez` and `rentrez` functions with `'::'` notation rather than `library()` calls to avoid namespace issues. e.g. `restez::entrez_fetch()`.

**See Also**[rentrez::entrez\\_fetch\(\)](#)**Examples**

```

library(restez)
reste_path_set(tempdir())
demo_db_create(n = 5)
# return fasta record
fasta_res <- entrez_fetch(db = 'nucleotide',
                          id = c('demo_1', 'demo_2'),
                          rettype = 'fasta')

cat(fasta_res)
# return whole GB record in text format
gb_res <- entrez_fetch(db = 'nucleotide',
                      id = c('demo_1', 'demo_2'),
                      rettype = 'gb')

cat(gb_res)
# NOT RUN
# whereas these request would go through rentrez
# fasta_res <- entrez_fetch(db = 'nucleotide',
#                           id = c('S71333', 'S71334'),
#                           rettype = 'fasta')
# gb_res <- entrez_fetch(db = 'nucleotide',
#                        id = c('S71333', 'S71334'),
#                        rettype = 'gb')

# delete demo after example
db_delete(everything = TRUE)

```

---

`entrez_gb_get`*Get Entrez GenBank record*

---

**Description**

Return gb and gbwithparts format as expected from an Entrez call. If not all IDs are returned, will run `rentrez::entrez_fetch`.

**Usage**

```
entrez_gb_get(id, ...)
```

**Arguments**

<code>id</code>	vector, unique ID(s) for record(s)
<code>...</code>	arguments passed on to <code>rentrez</code>

**Value**

character string containing the file created

**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dwld_path_get()`, `dwld_rcrd_log()`, `entrez_fasta_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_add()`, `gb_sql_query()`, `gbrelease_check()`, `gbrelease_get()`, `gbrelease_log()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dwld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_def()`, `mock_gb_df_generate()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`, `readme_log()`, `restez_connect()`, `restez_disconnect()`, `restez_path_check()`, `restez_rl()`, `search_gz()`, `seshinfo_log()`, `setup()`, `slctn_get()`, `slctn_log()`, `sql_path_get()`, `status_class()`, `stat()`, `testdatadir_get()`

---

extract_accession	<i>Extract accession</i>
-------------------	--------------------------

---

**Description**

Return accession ID from GenBank record

**Usage**

```
extract_accession(record)
```

**Arguments**

record	GenBank record in text format, character
--------	--

**Details**

If element is not found, "" returned.

**Value**

character

**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dwld_path_get()`, `dwld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_add()`, `gb_sql_query()`, `gbrelease_check()`, `gbrelease_get()`,

gbrelease\_log(), has\_data(), identify\_downloadable\_files(), last\_add\_get(), last\_dwnld\_get(), last\_entry\_get(), latest\_genbank\_release\_notes(), latest\_genbank\_release(), message\_missing(), mock\_def(), mock\_gb\_df\_generate(), mock\_org(), mock\_rec(), mock\_seq(), predict\_datasizes(), readme\_log(), restez\_connect(), restez\_disconnect(), restez\_path\_check(), restez\_rl(), search\_gz(), seshinfo\_log(), setup(), slctn\_get(), slctn\_log(), sql\_path\_get(), status\_class(), stat(), testdatadir\_get()

---

extract\_by\_patterns     *Extract by keyword*

---

### Description

Search through GenBank record for a keyword and return text up to the end\_pattern.

### Usage

```
extract_by_patterns(record, start_pattern, end_pattern = "\n")
```

### Arguments

record	GenBank record in text format, character
start_pattern	REGEX pattern indicating the point to start extraction, character
end_pattern	REGEX pattern indicating the point to stop extraction, character

### Details

The start\_pattern should be any of the capitalized elements in a GenBank record (e.g. LOCUS, DESCRIPTION, ACCESSION). The end\_pattern depends on how much of the selected element a user wants returned. By default, the extraction will stop at the next newline. If keyword or end pattern not found, returns NULL.

### Value

character or NULL

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#),

[readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

extract\_clean\_sequence

*Extract clean sequence from sequence part*

## Description

Return clean sequence from seqrecpart of a GenBank record

## Usage

```
extract_clean_sequence(seqrecpart, max_len = 1e+08)
```

## Arguments

seqrecpart	Sequence part of a GenBank record, character
max_len	Number: maximum number of characters allowed in a single record before splitting the record into parts. Does not affect output, but only internal calculations, so generally should not be changed. Default = 1e8.

## Details

If element is not found, ” returned.

## Value

character

## See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)



---

extract_definition	<i>Extract definition</i>
--------------------	---------------------------

---

### Description

Return definition from GenBank record.

### Usage

```
extract_definition(record)
```

### Arguments

record            GenBank record in text format, character

### Details

If element is not found, "" returned.

### Value

character

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

extract_features	<i>Extract features</i>
------------------	-------------------------

---

### Description

Return feature table as list from GenBank record

### Usage

```
extract_features(record)
```

### Arguments

record            GenBank record in text format, character

### Details

If element is not found, empty list returned.

### Value

list of lists

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

extract\_inforecpart     *Extract the information record part*

---

### Description

Return information part from GenBank record

### Usage

```
extract_inforecpart(record)
```

### Arguments

record                GenBank record in text format, character

### Details

If element is not found, "" returned.

### Value

character

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

extract_keywords	<i>Extract keywords</i>
------------------	-------------------------

---

### Description

Return keywords as list from GenBank record

### Usage

```
extract_keywords(record)
```

### Arguments

record            GenBank record in text format, character

### Details

If element is not found, ” returned.

### Value

character vector

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_infocpart\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

extract_locus	<i>Extract locus</i>
---------------	----------------------

---

### Description

Return locus information from GenBank record

### Usage

```
extract_locus(record)
```

### Arguments

record            GenBank record in text format, character

### Details

If element is not found, "" returned.

### Value

named character vector

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_infocpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

extract_organism	<i>Extract organism</i>
------------------	-------------------------

---

### Description

Return organism name from GenBank record

### Usage

```
extract_organism(record)
```

### Arguments

record            GenBank record in text format, character

### Details

If element is not found, "" returned.

### Value

character

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_infocpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

extract_seqrepart	<i>Extract the sequence record part</i>
-------------------	---

---

### Description

Return sequence part from GenBank record

### Usage

```
extract_seqrepart(record)
```

### Arguments

record            GenBank record in text format, character

### Details

If element is not found, "" returned.

### Value

character

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_infocpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

extract_sequence	<i>Extract sequence</i>
------------------	-------------------------

---

**Description**

Return sequence from GenBank record

**Usage**

```
extract_sequence(record)
```

**Arguments**

record	GenBank record in text format, character
--------	--

**Details**

If element is not found, "" returned.

**Value**

character

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_infocpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)



---

extract_version	<i>Extract version</i>
-----------------	------------------------

---

### Description

Return accession + version ID from GenBank record

### Usage

```
extract_version(record)
```

### Arguments

record            GenBank record in text format, character

### Details

If element is not found, "" returned.

### Value

character

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_infocpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

filename_log	<i>Write filenames to log files</i>
--------------	-------------------------------------

---

**Description**

Record a filename in a log file along with GB release and time.

**Usage**

```
filename_log(fl, fp)
```

**Arguments**

fl	file name, character
fp	filepath to log file, character

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

file_download	<i>Download a file</i>
---------------	------------------------

---

**Description**

Download a GenBank .seq.tar file. Check the file has downloaded properly. If not, returns FALSE. If overwrite is true, any previous file will be overwritten.

**Usage**

```
file_download(fl, overwrite = FALSE)
```

**Arguments**

fl                    character, base filename (e.g. gbpri9.seq) to be downloaded  
 overwrite            T/F

**Value**

T/F

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_infocpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

flatfile\_read                    *Read flatfile sequence records*

---

**Description**

Read records from a .seq file.

**Usage**

flatfile\_read(flpth)

**Arguments**

flpth                    Path to .seq file

**Value**

list of GenBank records in text format

**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dwncld_path_get()`, `dwncld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_add()`, `gb_sql_query()`, `gbrelease_check()`, `gbrelease_get()`, `gbrelease_log()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dwncld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_def()`, `mock_gb_df_generate()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`, `readme_log()`, `restez_connect()`, `restez_disconnect()`, `restez_path_check()`, `restez_rl()`, `search_gz()`, `seshinfo_log()`, `setup()`, `slctn_get()`, `slctn_log()`, `sql_path_get()`, `status_class()`, `stat()`, `testdatadir_get()`

---

gbrelease\_check

*Check if the last GenBank release number is the latest*


---

**Description**

Returns TRUE if the GenBank release number is the most recent GenBank release available.

**Usage**

```
gbrelease_check()
```

**Value**

logical

**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dwncld_path_get()`, `dwncld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_add()`, `gb_sql_query()`, `gbrelease_get()`, `gbrelease_log()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dwncld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_def()`, `mock_gb_df_generate()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`, `readme_log()`, `restez_connect()`, `restez_disconnect()`, `restez_path_check()`, `restez_rl()`, `search_gz()`, `seshinfo_log()`, `setup()`, `slctn_get()`, `slctn_log()`, `sql_path_get()`, `status_class()`, `stat()`, `testdatadir_get()`

---

gbrelease_get	<i>Get the GenBank release number in the restez path</i>
---------------	--

---

**Description**

Returns the GenBank release number. Returns empty character if none found.

**Usage**

```
gbrelease_get()
```

**Details**

If no file found, returns empty character vector.

**Value**

character

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

gbrelease_log	<i>Log the GenBank release number in the restez path</i>
---------------	--

---

**Description**

This function is called whenever `db_download` is run. It logs the GB release number in the 'gb\_release.txt' in the user's restez path. The log is to help users keep track of whether their database is out of date.

**Usage**

```
gbrelease_log(release)
```

**Arguments**

release            GenBank release number, character

**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dwnld_path_get()`, `dwnld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_add()`, `gb_sql_query()`, `gbrelease_check()`, `gbrelease_get()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dwnld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_def()`, `mock_gb_df_generate()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`, `readme_log()`, `restez_connect()`, `restez_disconnect()`, `restez_path_check()`, `restez_rl()`, `search_gz()`, `seshinfo_log()`, `setup()`, `slctn_get()`, `slctn_log()`, `sql_path_get()`, `status_class()`, `stat()`, `testdatadir_get()`

---

gb\_build

*Read and add .seq files to database*


---

**Description**

Given a list of `seq_files`, read and add the contents of the files to a SQL-like database. If any errors during the process, FALSE is returned.

**Usage**

```
gb_build(
  dpth,
  seq_files,
  max_length,
  min_length,
  acc_filter = NULL,
  invert = FALSE,
  scan = FALSE
)
```

**Arguments**

dpth            Download path (where `seq_files` are stored)

seq\_files        .seq.tar seq file names

max\_length      Maximum sequence length, default NULL.

min\_length      Minimum sequence length, default 0.



**Value**

named vector of definitions, if no results found NULL

**See Also**

[ncbi\\_acc\\_get\(\)](#)

Other get: [gb\\_fasta\\_get\(\)](#), [gb\\_organism\\_get\(\)](#), [gb\\_record\\_get\(\)](#), [gb\\_sequence\\_get\(\)](#), [gb\\_version\\_get\(\)](#)

**Examples**

```
library(restez)
restez_path_set(filepath = tempdir())
demo_db_create(n = 5)
(def <- gb_definition_get(id = 'demo_1'))
(defs <- gb_definition_get(id = c('demo_1', 'demo_2'))

# delete demo after example
db_delete(everything = TRUE)
```

---

gb\_df\_create

*Create GenBank data.frame*

---

**Description**

Make data.frame from columns vectors for nucleotide entries. As part of `gb_df_generate()`.

**Usage**

```
gb_df_create(accessions, versions, organisms, definitions, sequences, records)
```

**Arguments**

accessions	character, vector of accessions
versions	character, vector of accessions + versions
organisms	character, vector of organism names
definitions	character, vector of sequence definitions
sequences	character, vector of sequences
records	character, vector of GenBank records in text format

**Value**

data.frame



**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dwncld_path_get()`, `dwncld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_generate()`, `gb_sql_add()`, `gb_sql_query()`, `gbrelease_check()`, `gbrelease_get()`, `gbrelease_log()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dwncld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_def()`, `mock_gb_df_generate()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`, `readme_log()`, `restez_connect()`, `restez_disconnect()`, `restez_path_check()`, `restez_rl()`, `search_gz()`, `seshinfo_log()`, `setup()`, `slctn_get()`, `slctn_log()`, `sql_path_get()`, `status_class()`, `stat()`, `testdatadir_get()`

---

gb\_df\_generate

*Generate GenBank records data.frame*


---

**Description**

For a list of records, construct a data.frame for insertion into SQL database.

**Usage**

```
gb_df_generate(
  records,
  min_length = 0,
  max_length = NULL,
  acc_filter = NULL,
  invert = FALSE
)
```

**Arguments**

<code>records</code>	character, vector of GenBank records in text format
<code>min_length</code>	Minimum sequence length, default 0.
<code>max_length</code>	Maximum sequence length, default NULL.
<code>acc_filter</code>	Character vector; accessions to include or exclude from the database as specified by <code>invert</code> .
<code>invert</code>	Logical vector of length 1; if TRUE, accessions in <code>acc_filter</code> will be excluded from the database; if FALSE, only accessions in <code>acc_filter</code> will be included in the database. Default FALSE.

**Details**

The resulting data.frame has five columns: accession, organism, raw\_definition, raw\_sequence, raw\_record. The prefix 'raw\_' indicates the data has been converted to the raw format, see ?char-ToRaw, in order to save on RAM. The raw\_record contains the entire GenBank record in text format.

Use acc\_filter and max and min sequence lengths to minimize the size of the database. All sequences have to be at least as long as min and less than or equal in length to max, unless max is NULL in which there is no maximum length. The final selection of sequences is the result of applying all filters (acc\_filter, min\_length, max\_length) in combination.

**Value**

data.frame, or NULL if no records pass filters

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

gb\_extract

*Extract elements of a GenBank record*

---

**Description**

Return elements of GenBank record e.g. sequence, definition ...

**Usage**

```
gb_extract(
  record,
  what = c("accession", "version", "organism", "sequence", "definition", "locus",
    "features", "keywords")
)
```

**Arguments**

record	GenBank record in text format, character
what	Which element to extract

**Details**

This function uses a REGEX to extract particular elements of a GenBank record. All of the what options return a single character with the exception of 'locus' or 'keywords' that return character vectors and 'features' that returns a list of lists for all features.

The accuracy of these functions cannot be guaranteed due to the enormity of the GenBank database. But the function is regularly tested on a range of GenBank records.

Note: all non-latin1 characters are converted to '-'.

**Value**

character or list of lists (what='features') or named character vector (what='locus')

**Examples**

```
library(restz)
data('record')
(gb_extract(record = record, what = 'locus'))
```

---

gb_fasta_get	<i>Get fasta from GenBank</i>
--------------	-------------------------------

---

**Description**

Get sequence and definition data in FASTA format. Equivalent to rettype='fasta' in `rentrez::entrez_fetch()`.

**Usage**

```
gb_fasta_get(id, width = 70)
```

**Arguments**

id	character, sequence accession ID(s)
width	integer, maximum number of characters in a line

**Value**

named vector of fasta sequences, if no results found NULL

**See Also**

[ncbi\\_acc\\_get\(\)](#)

Other get: [gb\\_definition\\_get\(\)](#), [gb\\_organism\\_get\(\)](#), [gb\\_record\\_get\(\)](#), [gb\\_sequence\\_get\(\)](#), [gb\\_version\\_get\(\)](#)

### Examples

```
library(restez)
restez_path_set(filepath = tempdir())
demo_db_create(n = 5)
(fasta <- gb_fasta_get(id = 'demo_1'))
(fastas <- gb_fasta_get(id = c('demo_1', 'demo_2'))))

# delete demo after example
db_delete(everything = TRUE)
```

---

gb_organism_get	<i>Get organism from GenBank</i>
-----------------	----------------------------------

---

### Description

Return the organism name for an accession ID.

### Usage

```
gb_organism_get(id)
```

### Arguments

id                    character, sequence accession ID(s)

### Value

named vector of definitions, if no results found NULL

### See Also

[ncbi\\_acc\\_get\(\)](#)

Other get: [gb\\_definition\\_get\(\)](#), [gb\\_fasta\\_get\(\)](#), [gb\\_record\\_get\(\)](#), [gb\\_sequence\\_get\(\)](#), [gb\\_version\\_get\(\)](#)

### Examples

```
library(restez)
restez_path_set(filepath = tempdir())
demo_db_create(n = 5)
(org <- gb_organism_get(id = 'demo_1'))
(orgs <- gb_organism_get(id = c('demo_1', 'demo_2'))))

# delete demo after example
db_delete(everything = TRUE)
```

---

gb_record_get	<i>Get record from GenBank</i>
---------------	--------------------------------

---

**Description**

Return the entire GenBank record for an accession ID. Equivalent to `rettype='gb'` in `rentrez::entrez_fetch()`.

**Usage**

```
gb_record_get(id)
```

**Arguments**

`id` character, sequence accession ID(s)

**Value**

named vector of records, if no results found NULL

**See Also**

[ncbi\\_acc\\_get\(\)](#)

Other get: [gb\\_definition\\_get\(\)](#), [gb\\_fasta\\_get\(\)](#), [gb\\_organism\\_get\(\)](#), [gb\\_sequence\\_get\(\)](#), [gb\\_version\\_get\(\)](#)

**Examples**

```
library(restez)
restez_path_set(filepath = tempdir())
demo_db_create(n = 5)
(rec <- gb_record_get(id = 'demo_1'))
(recs <- gb_record_get(id = c('demo_1', 'demo_2')))
```

  

```
# delete demo after example
db_delete(everything = TRUE)
```

---

gb_sequence_get	<i>Get sequence from GenBank</i>
-----------------	----------------------------------

---

**Description**

Return the sequence(s) for a record(s) from the accession ID(s).

**Usage**

```
gb_sequence_get(id, dnabin = FALSE)
```

**Arguments**

id character, sequence accession ID(s)  
dnabin Logical vector of length 1; should the sequences be returned using the bit-level coding scheme of the ape package? Default FALSE.

**Details**

For more information about the dnabin format, see [ape::DNabin\(\)](#).

**Value**

named vector of sequences, if no results found NULL

**See Also**

[ncbi\\_acc\\_get\(\)](#)

Other get: [gb\\_definition\\_get\(\)](#), [gb\\_fasta\\_get\(\)](#), [gb\\_organism\\_get\(\)](#), [gb\\_record\\_get\(\)](#), [gb\\_version\\_get\(\)](#)

**Examples**

```
library(restez)
restez_path_set(filepath = tempdir())
demo_db_create(n = 5)
(seq <- gb_sequence_get(id = 'demo_1'))
(seqs <- gb_sequence_get(id = c('demo_1', 'demo_2')))
(fasta_dnabin <- gb_sequence_get(id = 'demo_1', dnabin = TRUE))

# delete demo after example
db_delete(everything = TRUE)
```

---

gb\_sql\_add

*Add to GenBank SQL database*

---

**Description**

Add records data.frame to SQL-like database.

**Usage**

```
gb_sql_add(df)
```

**Arguments**

df Records data.frame

**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dwlnld_path_get()`, `dwlnld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_query()`, `gbrelease_check()`, `gbrelease_get()`, `gbrelease_log()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dwlnld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_def()`, `mock_gb_df_generate()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`, `readme_log()`, `restez_connect()`, `restez_disconnect()`, `restez_path_check()`, `restez_rl()`, `search_gz()`, `seshinfo_log()`, `setup()`, `slctn_get()`, `slctn_log()`, `sql_path_get()`, `status_class()`, `stat()`, `testdatadir_get()`

---

gb\_sql\_query

*Query the GenBank SQL*


---

**Description**

Generic query function for retrieving data from the SQL database for the get functions.

**Usage**

```
gb_sql_query(nm, id)
```

**Arguments**

nm	character, column name
id	character, sequence accession ID(s)

**Value**

data.frame

**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dwlnld_path_get()`, `dwlnld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_add()`, `gbrelease_check()`, `gbrelease_get()`, `gbrelease_log()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dwlnld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_def()`, `mock_gb_df_generate()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`,

[readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

gb_version_get	<i>Get version from GenBank</i>
----------------	---------------------------------

---

### Description

Return the accession version for an accession ID.

### Usage

```
gb_version_get(id)
```

### Arguments

id                    character, sequence accession ID(s)

### Value

named vector of versions, if no results found NULL

### See Also

[ncbi\\_acc\\_get\(\)](#)

Other get: [gb\\_definition\\_get\(\)](#), [gb\\_fasta\\_get\(\)](#), [gb\\_organism\\_get\(\)](#), [gb\\_record\\_get\(\)](#), [gb\\_sequence\\_get\(\)](#)

### Examples

```
library(restez)
restez_path_set(filepath = tempdir())
demo_db_create(n = 5)
(ver <- gb_version_get(id = 'demo_1'))
(vers <- gb_version_get(id = c('demo_1', 'demo_2')))
```

  

```
# delete demo after example
db_delete(everything = TRUE)
```



---

has_data	<i>Does the connected database have data?</i>
----------	---

---

**Description**

Returns TRUE if a restez SQL database has data.

**Usage**

```
has_data()
```

**Value**

Logical

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

```
identify_downloadable_files
```

*Identify downloadable files*

---

**Description**

Searches through the release notes for a GenBank release to find all listed .seq files. Returns a data.frame for all .seq files and their description.

**Usage**

```
identify_downloadable_files()
```

**Value**

data.frame

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

 is\_in\_db

*Is in db*


---

**Description**

Determine whether an id(s) is/are present in a database.

**Usage**

```
is_in_db(id, db = "nucleotide")
```

**Arguments**

id	character, sequence accession ID(s)
db	character, database name

**Value**

named vector of booleans

**See Also**

Other database: [count\\_db\\_ids\(\)](#), [db\\_create\(\)](#), [db\\_delete\(\)](#), [db\\_download\(\)](#), [demo\\_db\\_create\(\)](#), [list\\_db\\_ids\(\)](#)

**Examples**

```
library(restez)
# set the restez path to a temporary dir
restez_path_set(filepath = tempdir())
# create demo database
demo_db_create(n = 5)
# in the demo, IDs are 'demo_1', 'demo_2' ...
```

```
ids <- c('thisisnotanid', 'demo_1', 'demo_2')
(is_in_db(id = ids))
```

```
# delete demo after example
db_delete(everything = TRUE)
```

---

last\_add\_get

*Return date and time of the last added sequence*

---

## Description

Return the date and time of the last added sequence as determined using the 'add\_log.tsv'.

## Usage

```
last_add_get()
```

## Details

If no file found, returns empty character vector.

## Value

character

## See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

last_dwnld_get	<i>Return date and time of the last download</i>
----------------	--

---

**Description**

Return the date and time of the last download as determined using the 'download\_log.tsv'.

**Usage**

```
last_dwnld_get()
```

**Details**

If no file found, returns empty character vector.

**Value**

character

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

last_entry_get	<i>Return the last entry</i>
----------------	------------------------------

---

**Description**

Return the last entry from a tab-delimited log file.

**Usage**

```
last_entry_get(fp)
```

**Arguments**

fp                   Filepath, character

**Value**

vector

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

latest\_genbank\_release

*Retrieve latest GenBank release number*

---

**Description**

Downloads the latest GenBank release number and returns it.

**Usage**

```
latest_genbank_release()
```

**Value**

character

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#),

gb\_build(), gb\_df\_create(), gb\_df\_generate(), gb\_sql\_add(), gb\_sql\_query(), gbrelease\_check(), gbrelease\_get(), gbrelease\_log(), has\_data(), identify\_downloadable\_files(), last\_add\_get(), last\_dwnld\_get(), last\_entry\_get(), latest\_genbank\_release\_notes(), message\_missing(), mock\_def(), mock\_gb\_df\_generate(), mock\_org(), mock\_rec(), mock\_seq(), predict\_datasizes(), readme\_log(), restez\_connect(), restez\_disconnect(), restez\_path\_check(), restez\_rl(), search\_gz(), seshinfo\_log(), setup(), slctn\_get(), slctn\_log(), sql\_path\_get(), status\_class(), stat(), testdatadir\_get()

---

latest\_genbank\_release\_notes

*Download the latest GenBank Release Notes*

---

### Description

Downloads the latest GenBank release notes to a user's restez download path.

### Usage

```
latest_genbank_release_notes()
```

### See Also

Other private: add\_rcrd\_log(), cat\_line(), char(), check\_connection(), cleanup(), connected(), connection\_get(), db\_download\_intern(), db\_sqlngths\_get(), db\_sqlngths\_log(), dir\_size(), dwnld\_path\_get(), dwnld\_rcrd\_log(), entrez\_fasta\_get(), entrez\_gb\_get(), extract\_accession(), extract\_by\_patterns(), extract\_clean\_sequence(), extract\_definition(), extract\_features(), extract\_inforecpart(), extract\_keywords(), extract\_locus(), extract\_organism(), extract\_seqrecpart(), extract\_sequence(), extract\_version(), file\_download(), filename\_log(), flatfile\_read(), gb\_build(), gb\_df\_create(), gb\_df\_generate(), gb\_sql\_add(), gb\_sql\_query(), gbrelease\_check(), gbrelease\_get(), gbrelease\_log(), has\_data(), identify\_downloadable\_files(), last\_add\_get(), last\_dwnld\_get(), last\_entry\_get(), latest\_genbank\_release(), message\_missing(), mock\_def(), mock\_gb\_df\_generate(), mock\_org(), mock\_rec(), mock\_seq(), predict\_datasizes(), readme\_log(), restez\_connect(), restez\_disconnect(), restez\_path\_check(), restez\_rl(), search\_gz(), seshinfo\_log(), setup(), slctn\_get(), slctn\_log(), sql\_path\_get(), status\_class(), stat(), testdatadir\_get()

---

list\_db\_ids

*List database IDs*

---

### Description

Return a vector of all IDs in a database.

### Usage

```
list_db_ids(db = "nucleotide", n = 100)
```

**Arguments**

db                    character, database name  
n                     Maximum number of IDs to return, if NULL returns all

**Details**

Warning: can return very large vectors for large databases.

**Value**

vector of characters

**See Also**

Other database: [count\\_db\\_ids\(\)](#), [db\\_create\(\)](#), [db\\_delete\(\)](#), [db\\_download\(\)](#), [demo\\_db\\_create\(\)](#), [is\\_in\\_db\(\)](#)

**Examples**

```
library(restez)
restez_path_set(filepath = tempdir())
demo_db_create(n = 5)
# Warning: not recommended for real databases
# with potentially millions of IDs
all_ids <- list_db_ids()

# What shall we do with these IDs?
# ... how about make a mock fasta file
seqs <- gb_sequence_get(id = all_ids)
defs <- gb_definition_get(id = all_ids)
# paste together
fasta_seqs <- paste0('>', defs, '\n', seqs)
fasta_file <- paste0(fasta_seqs, collapse = '\n')
cat(fasta_file)

# delete after example
db_delete(everything = TRUE)
```

---

message\_missing                    *Produce message of missing IDs*

---

**Description**

Sends message to console stating number of missing IDs.

**Usage**

```
message_missing(n)
```

**Arguments**

n                    Number of missing IDs

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

mock\_def

*Mock def*

---

**Description**

Make a mock sequence definition. Designed to be part of a loop.

**Usage**

```
mock_def(i)
```

**Arguments**

i                    integer, iterator

**Value**

character



**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dwlnld_path_get()`, `dwlnld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_add()`, `gb_sql_query()`, `gbrelease_check()`, `gbrelease_get()`, `gbrelease_log()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dwlnld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_gb_df_generate()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`, `readme_log()`, `restez_connect()`, `restez_disconnect()`, `restez_path_check()`, `restez_rl()`, `search_gz()`, `seshinfo_log()`, `setup()`, `slctn_get()`, `slctn_log()`, `sql_path_get()`, `status_class()`, `stat()`, `testdatadir_get()`

---

`mock_gb_df_generate`     *Generate mock GenBank records data.frame*

---

**Description**

Make a mock nucleotide data.frame for entry into a demonstration SQL database.

**Usage**

```
mock_gb_df_generate(n)
```

**Arguments**

`n`                    integer, number of entries

**Value**

data.frame

**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dwlnld_path_get()`, `dwlnld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_add()`, `gb_sql_query()`, `gbrelease_check()`, `gbrelease_get()`, `gbrelease_log()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dwlnld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_def()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`, `readme_log()`, `restez_connect()`, `restez_disconnect()`, `restez_path_check()`, `restez_rl()`,

[search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

mock\_org

*Mock org*

---

## Description

Make a mock sequence organism. Designed to be part of a loop.

## Usage

```
mock_org(i)
```

## Arguments

`i` integer, iterator

## Value

character

## See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_infocpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

mock_rec	<i>Mock rec</i>
----------	-----------------

---

### Description

Create a mock GenBank record for demo-ing and testing purposes. Designed to be part of a loop. Accession, organism... etc. are optional arguments.

### Usage

```
mock_rec(
  i,
  definition = NULL,
  accession = NULL,
  version = NULL,
  organism = NULL,
  sequence = NULL
)
```

### Arguments

i	integer, iterator
definition	character
accession	character
version	character
organism	character
sequence	character

### Value

character

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

mock_seq	<i>Mock seq</i>
----------	-----------------

---

**Description**

Make a mock sequence. Designed to be part of a loop.

**Usage**

```
mock_seq(i, sqlngth = 10)
```

**Arguments**

i	integer, iterator
sqlngth	integer, sequence length

**Value**

character

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

ncbi_acc_get	<i>Get accession numbers by querying NCBI GenBank</i>
--------------	---

---

**Description**

The query string can be formatted using [GenBank advanced query terms](#) to obtain accession numbers corresponding to a specific set of criteria.

**Usage**

```
ncbi_acc_get(query, strict = TRUE, drop_ver = TRUE)
```

**Arguments**

query	Character vector of length 1; query string to search GenBank.
strict	Logical vector of length 1; should an error be issued if the number of unique accessions retrieved does not match the number of hits from GenBank? Default TRUE.
drop_ver	Logical vector of length 1; should the version part of the accession number (e.g., '.1' in 'AB001538.1') be dropped? Default TRUE.

**Details**

Note this queries NCBI GenBank, not the local database generated with restez.

It can be used either to restrict the accessions used to construct the local database (`acc_filter` argument of `db_create()`) or to specify accessions to read from the local database (`id` argument of `gb_fasta_get()` and other `gb_*_get()` functions).

**Value**

Character vector; accession numbers resulting from query.

**See Also**

[db\\_create\(\)](#), [gb\\_fasta\\_get\(\)](#)

**Examples**

```
## Not run:  
# requires an internet connection  
cmin_accs <- ncbi_acc_get("Crepidomanes minutum")  
length(cmin_accs)  
head(cmin_accs)  
  
## End(Not run)
```

---

predict\_datasizes      *Print file size predictions to screen*

---

**Description**

Predicts the file sizes of the downloads and the database from the GenBank filesize information. Conversion factors are based on previous restez downloads.

**Usage**

```
predict_datasizes(uncompressed_filesize)
```

**Arguments**

uncompressed\_filesize  
 GBs of the stated filesize, numeric

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

 print.status

*Print method for status class*


---

**Description**

Prints to screen the three sections of the status class. Not meant to be used interactively.

**Usage**

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

**Arguments**

x                    Status object  
 ...                 Other arguments (not used by this function)

---

readme_log	<i>Create README in restez_path</i>
------------	-------------------------------------

---

**Description**

Write notes for the curious sorts who peruse the restez\_path.

**Usage**

```
readme_log()
```

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_infocpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

record	<i>Example GenBank record</i>
--------	-------------------------------

---

**Description**

Example GenBank record in text format for demonstration purposes.

**Usage**

```
data("record")
```

**Format**

A large character object containing record information and DNA sequence.

**Source**

<https://www.ncbi.nlm.nih.gov/nuccore/AY952423.1>

**References**

GenBank

**Examples**

```
data(record)
cat(record)
```

---

restez_connect	<i>Connect to the restez database</i>
----------------	---------------------------------------

---

**Description**

Sets a connection to the local database.

**Usage**

```
restez_connect(read_only = FALSE)
```

**Arguments**

read_only	Logical; should the connection be made in read-only mode? Read-only mode is required for multiple R processes to access the database simultaneously. Default FALSE.
-----------	---

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)



---

restez_disconnect	<i>Disconnect from restez database</i>
-------------------	--

---

**Description**

Safely disconnect from the restez connection

**Usage**

```
restez_disconnect()
```

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

restez_path_check	<i>Check restez filepath</i>
-------------------	------------------------------

---

**Description**

Raises error if restez path does not exist.

**Usage**

```
restez_path_check()
```

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#),

```
gb_build(), gb_df_create(), gb_df_generate(), gb_sql_add(), gb_sql_query(), gbrelease_check(),
gbrelease_get(), gbrelease_log(), has_data(), identify_downloadable_files(), last_add_get(),
last_dwnld_get(), last_entry_get(), latest_genbank_release_notes(), latest_genbank_release(),
message_missing(), mock_def(), mock_gb_df_generate(), mock_org(), mock_rec(), mock_seq(),
predict_datasizes(), readme_log(), restez_connect(), restez_disconnect(), restez_rl(),
search_gz(), seshinfo_log(), setup(), slctn_get(), slctn_log(), sql_path_get(), status_class(),
stat(), testdatadir_get()
```

---

restez_path_get	<i>Get restez path</i>
-----------------	------------------------

---

### Description

Return filepath to where the restez database is stored.

### Usage

```
restez_path_get()
```

### Value

character

### See Also

Other setup: [restez\\_path\\_set\(\)](#), [restez\\_path\\_unset\(\)](#), [restez\\_ready\(\)](#), [restez\\_status\(\)](#)

### Examples

```
library(restez)
# set a restez path with a tempdir
restez_path_set(filepath = tempdir())
# check what the set path is
(restez_path_get())
```

---

restez_path_set	<i>Set restez path</i>
-----------------	------------------------

---

### Description

Specify the filepath for the local GenBank database.

### Usage

```
restez_path_set(filepath)
```

## Arguments

filepath            character, valid filepath to the folder where the database should be stored.

## Details

Adds 'restez\_path' to options(). In this path the folder 'restez' will be created and all downloaded and database files will be stored there.

## See Also

Other setup: [restez\\_path\\_get\(\)](#), [restez\\_path\\_unset\(\)](#), [restez\\_ready\(\)](#), [restez\\_status\(\)](#)

## Examples

```
## Not run:  
library(restez)  
restez_path_set(filepath = 'path/to/where/you/want/files/to/download')  
  
## End(Not run)
```

---

restez_path_unset	<i>Unset restez path</i>
-------------------	--------------------------

---

## Description

Set the restez path to NULL

## Usage

```
restez_path_unset()
```

## See Also

Other setup: [restez\\_path\\_get\(\)](#), [restez\\_path\\_set\(\)](#), [restez\\_ready\(\)](#), [restez\\_status\(\)](#)

---

restez_ready	<i>Is restez ready?</i>
--------------	-------------------------

---

**Description**

Returns TRUE if a restez SQL database is available. Use `restez_status()` for more information.

**Usage**

```
restez_ready()
```

**Value**

Logical

**See Also**

Other setup: [restez\\_path\\_get\(\)](#), [restez\\_path\\_set\(\)](#), [restez\\_path\\_unset\(\)](#), [restez\\_status\(\)](#)

**Examples**

```
library(restez)
fp <- tempdir()
restez_path_set(filepath = fp)
demo_db_create(n = 5)
(restez_ready())
db_delete(everything = TRUE)
(restez_ready())
```

---

restez_rl	<i>Restez readline</i>
-----------	------------------------

---

**Description**

Wrapper for base readline.

**Usage**

```
restez_rl(prompt)
```

**Arguments**

prompt            character, display text

**Value**

character

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

restez\_status

*Check restez status*


---

**Description**

Report to console current setup status of restez.

**Usage**

```
restez_status(gb_check = FALSE)
```

**Arguments**

`gb_check` Check whether last download was from latest GenBank release? Default FALSE.

**Details**

Set `gb_check=TRUE` to see if your downloads are up-to-date.

**Value**

Status class

**See Also**

Other setup: [restez\\_path\\_get\(\)](#), [restez\\_path\\_set\(\)](#), [restez\\_path\\_unset\(\)](#), [restez\\_ready\(\)](#)

**Examples**

```
library(restez)
fp <- tempdir()
restez_path_set(filepath = fp)
demo_db_create(n = 5)
restez_status()
db_delete(everything = TRUE)
# Errors:
# restez_status()
```

---

search\_gz

*Scan a gzipped file for text*


---

**Description**

Scans a zipped file for text strings and returns TRUE if any are present.

**Usage**

```
search_gz(terms, path)
```

**Arguments**

terms	Character vector; search terms (most likely GenBank accession numbers)
path	Path to the gzipped file to scan

**Value**

Logical

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

seshinfo_log	<i>Log the system session information in restez path</i>
--------------	--

---

**Description**

Records the session and system information to file.

**Usage**

```
seshinfo_log()
```

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

setup	<i>Set up test common test data</i>
-------	-------------------------------------

---

**Description**

Creates temporary test folders.

**Usage**

```
setup()
```

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#),

gb\_build(), gb\_df\_create(), gb\_df\_generate(), gb\_sql\_add(), gb\_sql\_query(), gbrelease\_check(), gbrelease\_get(), gbrelease\_log(), has\_data(), identify\_downloadable\_files(), last\_add\_get(), last\_dwnld\_get(), last\_entry\_get(), latest\_genbank\_release\_notes(), latest\_genbank\_release(), message\_missing(), mock\_def(), mock\_gb\_df\_generate(), mock\_org(), mock\_rec(), mock\_seq(), predict\_datasizes(), readme\_log(), restez\_connect(), restez\_disconnect(), restez\_path\_check(), restez\_rl(), search\_gz(), seshinfo\_log(), slctn\_get(), slctn\_log(), sql\_path\_get(), status\_class(), stat(), testdatadir\_get()

---

slctn\_get

*Retrieve GenBank selections made by user*

---

### Description

Returns the selections made by the user.

### Usage

```
slctn_get()
```

### Details

If no file found, returns empty character vector.

### Value

character vector

### See Also

Other private: add\_rcrd\_log(), cat\_line(), char(), check\_connection(), cleanup(), connected(), connection\_get(), db\_download\_intern(), db\_sqlngths\_get(), db\_sqlngths\_log(), dir\_size(), dwnld\_path\_get(), dwnld\_rcrd\_log(), entrez\_fasta\_get(), entrez\_gb\_get(), extract\_accession(), extract\_by\_patterns(), extract\_clean\_sequence(), extract\_definition(), extract\_features(), extract\_inforecpart(), extract\_keywords(), extract\_locus(), extract\_organism(), extract\_seqrecpart(), extract\_sequence(), extract\_version(), file\_download(), filename\_log(), flatfile\_read(), gb\_build(), gb\_df\_create(), gb\_df\_generate(), gb\_sql\_add(), gb\_sql\_query(), gbrelease\_check(), gbrelease\_get(), gbrelease\_log(), has\_data(), identify\_downloadable\_files(), last\_add\_get(), last\_dwnld\_get(), last\_entry\_get(), latest\_genbank\_release\_notes(), latest\_genbank\_release(), message\_missing(), mock\_def(), mock\_gb\_df\_generate(), mock\_org(), mock\_rec(), mock\_seq(), predict\_datasizes(), readme\_log(), restez\_connect(), restez\_disconnect(), restez\_path\_check(), restez\_rl(), search\_gz(), seshinfo\_log(), setup(), slctn\_log(), sql\_path\_get(), status\_class(), stat(), testdatadir\_get()



---

 slctn\_log

*Log the GenBank selection made by a user*


---

**Description**

This function is called whenever a user makes a selection with the [db\\_download\(\)](#). It records GenBank numbers selections.

**Usage**

```
slctn_log(selection)
```

**Arguments**

selection      selected GenBank sequences, named vector

**See Also**

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

 sql\_path\_get

*Get SQL path*


---

**Description**

Return path to where SQL database is stored.

**Usage**

```
sql_path_get()
```

**Value**

character

**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dwld_path_get()`, `dwld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_add()`, `gb_sql_query()`, `gbrelease_check()`, `gbrelease_get()`, `gbrelease_log()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dwld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_def()`, `mock_gb_df_generate()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`, `readme_log()`, `restez_connect()`, `restez_disconnect()`, `restez_path_check()`, `restez_rl()`, `search_gz()`, `seshinfo_log()`, `setup()`, `slctn_get()`, `slctn_log()`, `status_class()`, `stat()`, `testdatadir_get()`

---

stat

*Print blue*


---

**Description**

Print to console blue text to indicate a number/statistic.

**Usage**

```
stat(...)
```

**Arguments**

... Any number of text arguments to print, character

**Value**

coloured character encoding, character

**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dwld_path_get()`, `dwld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_add()`, `gb_sql_query()`, `gbrelease_check()`, `gbrelease_get()`, `gbrelease_log()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dwld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_def()`, `mock_gb_df_generate()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`, `readme_log()`, `restez_connect()`, `restez_disconnect()`, `restez_path_check()`,

[restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [status\\_class\(\)](#), [testdatadir\\_get\(\)](#)

---

status\_class

*Generate a list class for storing status information*

---

### Description

Creates a three-part list for holding information on the status of the restez file path.

### Usage

status\_class()

### Value

Status class

### See Also

Other private: [add\\_rcrd\\_log\(\)](#), [cat\\_line\(\)](#), [char\(\)](#), [check\\_connection\(\)](#), [cleanup\(\)](#), [connected\(\)](#), [connection\\_get\(\)](#), [db\\_download\\_intern\(\)](#), [db\\_sqlngths\\_get\(\)](#), [db\\_sqlngths\\_log\(\)](#), [dir\\_size\(\)](#), [dwnld\\_path\\_get\(\)](#), [dwnld\\_rcrd\\_log\(\)](#), [entrez\\_fasta\\_get\(\)](#), [entrez\\_gb\\_get\(\)](#), [extract\\_accession\(\)](#), [extract\\_by\\_patterns\(\)](#), [extract\\_clean\\_sequence\(\)](#), [extract\\_definition\(\)](#), [extract\\_features\(\)](#), [extract\\_inforecpart\(\)](#), [extract\\_keywords\(\)](#), [extract\\_locus\(\)](#), [extract\\_organism\(\)](#), [extract\\_seqrecpart\(\)](#), [extract\\_sequence\(\)](#), [extract\\_version\(\)](#), [file\\_download\(\)](#), [filename\\_log\(\)](#), [flatfile\\_read\(\)](#), [gb\\_build\(\)](#), [gb\\_df\\_create\(\)](#), [gb\\_df\\_generate\(\)](#), [gb\\_sql\\_add\(\)](#), [gb\\_sql\\_query\(\)](#), [gbrelease\\_check\(\)](#), [gbrelease\\_get\(\)](#), [gbrelease\\_log\(\)](#), [has\\_data\(\)](#), [identify\\_downloadable\\_files\(\)](#), [last\\_add\\_get\(\)](#), [last\\_dwnld\\_get\(\)](#), [last\\_entry\\_get\(\)](#), [latest\\_genbank\\_release\\_notes\(\)](#), [latest\\_genbank\\_release\(\)](#), [message\\_missing\(\)](#), [mock\\_def\(\)](#), [mock\\_gb\\_df\\_generate\(\)](#), [mock\\_org\(\)](#), [mock\\_rec\(\)](#), [mock\\_seq\(\)](#), [predict\\_datasizes\(\)](#), [readme\\_log\(\)](#), [restez\\_connect\(\)](#), [restez\\_disconnect\(\)](#), [restez\\_path\\_check\(\)](#), [restez\\_rl\(\)](#), [search\\_gz\(\)](#), [seshinfo\\_log\(\)](#), [setup\(\)](#), [slctn\\_get\(\)](#), [slctn\\_log\(\)](#), [sql\\_path\\_get\(\)](#), [stat\(\)](#), [testdatadir\\_get\(\)](#)

---

testdatadir\_get

*Get test data directory*

---

### Description

Get the folder containing test data.

### Usage

testdatadir\_get()

**See Also**

Other private: `add_rcrd_log()`, `cat_line()`, `char()`, `check_connection()`, `cleanup()`, `connected()`, `connection_get()`, `db_download_intern()`, `db_sqlngths_get()`, `db_sqlngths_log()`, `dir_size()`, `dwlnld_path_get()`, `dwlnld_rcrd_log()`, `entrez_fasta_get()`, `entrez_gb_get()`, `extract_accession()`, `extract_by_patterns()`, `extract_clean_sequence()`, `extract_definition()`, `extract_features()`, `extract_inforecpart()`, `extract_keywords()`, `extract_locus()`, `extract_organism()`, `extract_seqrecpart()`, `extract_sequence()`, `extract_version()`, `file_download()`, `filename_log()`, `flatfile_read()`, `gb_build()`, `gb_df_create()`, `gb_df_generate()`, `gb_sql_add()`, `gb_sql_query()`, `gbrelease_check()`, `gbrelease_get()`, `gbrelease_log()`, `has_data()`, `identify_downloadable_files()`, `last_add_get()`, `last_dwlnld_get()`, `last_entry_get()`, `latest_genbank_release_notes()`, `latest_genbank_release()`, `message_missing()`, `mock_def()`, `mock_gb_df_generate()`, `mock_org()`, `mock_rec()`, `mock_seq()`, `predict_datasizes()`, `readme_log()`, `restez_connect()`, `restez_disconnect()`, `restez_path_check()`, `restez_rl()`, `search_gz()`, `seshinfo_log()`, `setup()`, `slctn_get()`, `slctn_log()`, `sql_path_get()`, `status_class()`, `stat()`

# Index

- \* **database**
  - count\_db\_ids, 9
  - db\_create, 10
  - db\_delete, 12
  - db\_download, 12
  - demo\_db\_create, 16
  - is\_in\_db, 50
  - list\_db\_ids, 54
- \* **datasets**
  - record, 63
- \* **entrez**
  - entrez\_fetch, 20
- \* **get**
  - gb\_definition\_get, 39
  - gb\_fasta\_get, 43
  - gb\_organism\_get, 44
  - gb\_record\_get, 45
  - gb\_sequence\_get, 45
  - gb\_version\_get, 48
- \* **parse**
  - gb\_extract, 42
- \* **private**
  - add\_rcrd\_log, 4
  - cat\_line, 4
  - char, 5
  - check\_connection, 6
  - cleanup, 7
  - connected, 7
  - connection\_get, 8
  - db\_download\_intern, 14
  - db\_sqlngths\_get, 15
  - db\_sqlngths\_log, 15
  - dir\_size, 17
  - dwnld\_path\_get, 18
  - dwnld\_rcrd\_log, 18
  - entrez\_fasta\_get, 19
  - entrez\_gb\_get, 21
  - extract\_accession, 22
  - extract\_by\_patterns, 23
  - extract\_clean\_sequence, 24
  - extract\_definition, 25
  - extract\_features, 26
  - extract\_inforecpart, 27
  - extract\_keywords, 28
  - extract\_locus, 29
  - extract\_organism, 30
  - extract\_seqrecpart, 31
  - extract\_sequence, 32
  - extract\_version, 33
  - file\_download, 34
  - filename\_log, 34
  - flatfile\_read, 35
  - gb\_build, 38
  - gb\_df\_create, 40
  - gb\_df\_generate, 41
  - gb\_sql\_add, 46
  - gb\_sql\_query, 47
  - gbrelease\_check, 36
  - gbrelease\_get, 37
  - gbrelease\_log, 37
  - has\_data, 49
  - identify\_downloadable\_files, 49
  - last\_add\_get, 51
  - last\_dwnld\_get, 52
  - last\_entry\_get, 52
  - latest\_genbank\_release, 53
  - latest\_genbank\_release\_notes, 54
  - message\_missing, 55
  - mock\_def, 56
  - mock\_gb\_df\_generate, 57
  - mock\_org, 58
  - mock\_rec, 59
  - mock\_seq, 60
  - predict\_datasizes, 61
  - readme\_log, 63
  - restez\_connect, 64
  - restez\_disconnect, 65
  - restez\_path\_check, 65

- restez\_rl, 68
- search\_gz, 70
- sesinfo\_log, 71
- setup, 71
- slctn\_get, 72
- slctn\_log, 73
- sql\_path\_get, 73
- stat, 74
- status\_class, 75
- testdatadir\_get, 75
- \* setup**
  - restez\_path\_get, 66
  - restez\_path\_set, 66
  - restez\_path\_unset, 67
  - restez\_ready, 68
  - restez\_status, 69
- add\_rcrd\_log, 4, 5–8, 14–19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- ape::DNAbin(), 46
- cat\_line, 4, 4, 5–8, 14–19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- char, 4, 5, 5, 6–8, 14–19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- check\_connection, 4, 5, 6, 7, 8, 14–19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- cleanup, 4–6, 7, 8, 14–19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- connected, 4–7, 7, 8, 14–19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- connection\_get, 4–8, 8, 14–19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- count\_db\_ids, 9, 11–13, 16, 50, 55
- db\_create, 9, 10, 12, 13, 16, 50, 55
- db\_create(), 61
- db\_delete, 9, 11, 12, 13, 16, 50, 55
- db\_delete(), 11
- db\_download, 9, 11, 12, 12, 16, 50, 55
- db\_download(), 73
- db\_download\_intern, 4–8, 14, 15–19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- db\_sqlngths\_get, 4–8, 14, 15, 16–19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- db\_sqlngths\_log, 4–8, 14, 15, 15, 17–19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- demo\_db\_create, 9, 11–13, 16, 50, 55
- dir\_size, 4–8, 14–16, 17, 18, 19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- dwnld\_path\_get, 4–8, 14–17, 18, 19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- dwnld\_rcrd\_log, 4–8, 14–18, 18, 19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- entrez\_fasta\_get, 4–8, 14–19, 19, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- entrez\_fetch, 20
- entrez\_gb\_get, 4–8, 14–19, 21, 22–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- extract\_accession, 4–8, 14–19, 22, 22, 23–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- extract\_by\_patterns, 4–8, 14–19, 22, 23, 24–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- extract\_clean\_sequence, 4–8, 14–19, 22, 23, 24, 25–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- extract\_definition, 4–8, 14–19, 22–24, 25, 26–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- extract\_features, 4–8, 14–19, 22–25, 26, 27–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- extract\_inforecpart, 4–8, 14–19, 22–26, 27, 28–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- extract\_keywords, 4–8, 14–19, 22–27, 28, 29–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- extract\_locus, 4–8, 14–19, 22–28, 29, 30–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- extract\_organism, 4–8, 14–19, 22–29, 30, 31–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- extract\_seqrecpart, 4–8, 14–19, 22–30, 31, 32–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76

- extract\_sequence, 4–8, 14–19, 22–31, 32, 33–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- extract\_version, 4–8, 14–20, 22–32, 33, 34–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- file\_download, 4–8, 14–20, 22–34, 34, 36–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- filename\_log, 4–8, 14–20, 22–33, 34, 35–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- flatfile\_read, 4–8, 14–20, 22–35, 35, 36–39, 41, 42, 47, 49–54, 56–60, 62–65, 69–76
- gb\_build, 4–8, 14–20, 22–38, 38, 41, 42, 47, 49–54, 56–60, 62–66, 69–76
- gb\_definition\_get, 39, 43–46, 48
- gb\_df\_create, 4–8, 14–20, 22–39, 40, 42, 47, 49–54, 56–60, 62–66, 69–76
- gb\_df\_generate, 4–8, 14–20, 22–39, 41, 41, 47, 49–54, 56–60, 62–66, 69–76
- gb\_extract, 42
- gb\_fasta\_get, 40, 43, 44–46, 48
- gb\_fasta\_get(), 20, 61
- gb\_organism\_get, 40, 43, 44, 45, 46, 48
- gb\_record\_get, 40, 43, 44, 45, 46, 48
- gb\_record\_get(), 20
- gb\_sequence\_get, 40, 43–45, 45, 48
- gb\_sql\_add, 4–8, 14–20, 22–39, 41, 42, 46, 47, 49–54, 56–60, 62–66, 69–76
- gb\_sql\_query, 4–8, 14–20, 22–39, 41, 42, 47, 49–54, 56–60, 62–66, 69–76
- gb\_version\_get, 40, 43–46, 48
- gbrelease\_check, 4–8, 14–20, 22–36, 36, 37–39, 41, 42, 47, 49–54, 56–60, 62–66, 69–76
- gbrelease\_get, 4–8, 14–20, 22–36, 37, 38, 39, 41, 42, 47, 49–54, 56–60, 62–66, 69–76
- gbrelease\_log, 4–8, 14–20, 22–37, 37, 39, 41, 42, 47, 49–54, 56–60, 62–66, 69–76
- has\_data, 4–8, 14–20, 22–39, 41, 42, 47, 49, 50–54, 56–60, 62–66, 69–76
- identify\_downloadable\_files, 4–8, 14–20, 22–39, 41, 42, 47, 49, 49, 51–54, 56–60, 62–66, 69–76
- is\_in\_db, 9, 11–13, 16, 50, 55
- last\_add\_get, 4–8, 14–20, 22–39, 41, 42, 47, 49, 50, 51, 52–54, 56–60, 62–66, 69–76
- last\_dwnld\_get, 4–8, 14–20, 22–39, 41, 42, 47, 49–51, 52, 53, 54, 56–60, 62–66, 69–76
- last\_entry\_get, 4–8, 14–20, 22–39, 41, 42, 47, 49–52, 52, 54, 56–60, 62–66, 69–76
- latest\_genbank\_release, 4–8, 14–20, 22–39, 41, 42, 47, 49–53, 53, 54, 56–60, 62–66, 69–76
- latest\_genbank\_release\_notes, 4–8, 14–20, 22–39, 41, 42, 47, 49–54, 54, 56–60, 62–66, 69–76
- list\_db\_ids, 9, 11–13, 16, 50, 54
- message\_missing, 4–8, 14–20, 22–39, 41, 42, 47, 49–54, 55, 57–60, 62–66, 69–76
- mock\_def, 4–8, 14–20, 22–39, 41, 42, 47, 49–54, 56, 56, 57–60, 62–66, 69–76
- mock\_gb\_df\_generate, 4–8, 14–20, 22–39, 41, 42, 47, 49–54, 56, 57, 57, 58–60, 62–66, 69–76
- mock\_org, 4–8, 14–20, 22–39, 41, 42, 47, 49–54, 56, 57, 58, 59, 60, 62–66, 69–76
- mock\_rec, 4–8, 14–20, 22–39, 41, 42, 47, 49–54, 56–58, 59, 60, 62–66, 69–76
- mock\_seq, 4–8, 14–20, 22–39, 41, 42, 47, 49–54, 56–59, 60, 62–66, 69–76
- ncbi\_acc\_get, 60
- ncbi\_acc\_get(), 13, 40, 43–46, 48
- predict\_datasizes, 4–8, 14–20, 22–39, 41, 42, 47, 49–54, 56–60, 61, 63–66, 69–76
- print.status, 62
- readme\_log, 4–8, 14–20, 22–39, 41, 42, 47–54, 56–60, 62, 63, 64–66, 69–76
- record, 63
- rentrez::entrez\_fetch(), 21, 43, 45

restez\_connect, 4–8, 14–20, 22–39, 41, 42,  
47–54, 56–60, 62, 63, 64, 65, 66,  
69–76

restez\_disconnect, 4–8, 14–20, 22–39, 41,  
42, 47–54, 56–60, 62–64, 65, 66,  
69–76

restez\_path\_check, 4–8, 14–20, 22–39, 41,  
42, 47–54, 56–60, 62–65, 65, 69–76

restez\_path\_get, 66, 67–69

restez\_path\_set, 66, 66, 67–69

restez\_path\_unset, 66, 67, 67, 68, 69

restez\_ready, 66, 67, 68, 69

restez\_rl, 4–8, 14–20, 22–39, 41, 42, 47–54,  
56–60, 62–66, 68, 70–76

restez\_status, 66–68, 69

search\_gz, 4–8, 14–20, 22–39, 41, 42, 47–54,  
56–60, 62–66, 69, 70, 71–76

seshinfo\_log, 4–8, 14–20, 22–39, 41, 42,  
47–54, 56–60, 62–66, 69, 70, 71,  
72–76

setup, 4–8, 14–20, 22–39, 41, 42, 47–54,  
56–60, 62–66, 69–71, 71, 72–76

slctn\_get, 4–8, 14–20, 22–39, 41, 42, 47–54,  
56–60, 62–66, 69–72, 72, 73–76

slctn\_log, 4–8, 14–20, 22–39, 41, 42, 47–54,  
56–60, 62–66, 69–72, 73, 74–76

sql\_path\_get, 4–8, 14–20, 22–39, 41, 42,  
47–54, 56–60, 62–66, 69–73, 73, 75,  
76

stat, 4–8, 14–20, 22–39, 41, 42, 47–54,  
56–60, 62–66, 69–74, 74, 75, 76

status\_class, 4–8, 14–20, 22–39, 41, 42,  
47–54, 56–60, 62–66, 69–75, 75, 76

testdatadir\_get, 4–8, 14–20, 22–39, 41, 42,  
47–54, 56–60, 62–66, 69–75, 75