

# Package ‘aliases2entrez’

February 15, 2021

**Title** Converts Human gene symbols to entrez IDs

**Version** 0.1.2

**Description** Queries multiple resources authors HGNC (2019) <<https://www.genenames.org>>, authors limma (2015) <[doi:10.1093/nar/gkv007](https://doi.org/10.1093/nar/gkv007)> to find the correspondence between evolving nomenclature of human gene symbols, aliases, previous symbols or synonyms with stable, curated gene entrezID from NCBI database. This allows fast, accurate and up-to-date correspondence between human gene expression datasets from various date and platform (e.g: gene symbol: BRCA1 - ID: 672).

**BugReports** <https://github.com/peyronlab/aliases2entrez/issues>

**Imports** doParallel, limma, utils, org.Hs.eg.db, AnnotationDbi, parallel, foreach, readr, RCurl

**License** MIT + file LICENSE

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.1.1

**Suggests** spelling

**Language** en-US

**NeedsCompilation** no

**Author** Raphael Bonnet [aut, cre] (Université Côte d’Azur),  
Lee Mariault [ctb] (Université Côte d’Azur),  
Jean-François Peyron [aut] (Inserm)

**Maintainer** Raphael Bonnet <[raphael.bonnet@univ-cotedazur.fr](mailto:raphael.bonnet@univ-cotedazur.fr)>

**Repository** CRAN

**Date/Publication** 2021-02-15 10:30:06 UTC

## R topics documented:

convert_symbols . . . . .	2
update_symbols . . . . .	3

---

convert_symbols	<i>Multi resources gene symbols conversion to entrez ID (Human)</i>
-----------------	---

---

### Description

This function is used to convert gene symbols, previous symbols or aliases to gene entrez ID

It performs :

- a gene query to limma::alias2Symbol to map gene alias to official symbols
- looks for LOC\* symbols
- tries to find correspondence within HGNC database
- queries org.Hs.eg.db
- checks again with adaptive symbol parsing (e.g. transforms BRCA-1 to BRCA1)

### Usage

```
convert_symbols(symbols, HGNC, c=1)
```

### Arguments

symbols	gene matrix from which rownames (gene symbols) are extracted
HGNC	HGNC correspondence file.
c	number of cores to use for parallel processes

### Value

returns a vector containing IDs if match were found or NA if unknown or withdrawn symbol

### Examples

```
# import the correspondence file
file <- system.file("extdata", "HGNC.txt", package = "aliases2entrez")
HGNC <- read.delim(file)
# alternatively update a new one with update_symbols()
symbols <- c("BRCA1", "TP53")
# run the main function
ids <- convert_symbols(symbols, HGNC)
```

---

update_symbols	<i>Update last HGNC correspondence database</i>
----------------	---

---

**Description**

This function is used to update gene symbol correspondence from HGNC database

**Usage**

```
update_symbols(url=NULL)
```

**Arguments**

url                    user can provide url (default is NULL)

**Value**

returns a data.frame containing gene symbols with status, previous symbols and synonyms as well as their corresponding entrezIDs

**Examples**

```
HGNC <- update_symbols()
```

# Index

`convert_symbols`, [2](#)

`update_symbols`, [3](#)