

Package ‘UniprotR’

August 19, 2019

Title Retrieving Information of Proteins from Uniprot

Version 1.0.3

Author Ali Mostafa [aut],
Mohamed Soudy [aut, cre]

Maintainer Mohamed Soudy <MohmedSoudy2009@gmail.com>

Description Connect to Uniprot <<https://www.uniprot.org/>> to retrieve information about proteins using their accession number such information could be name or taxonomy information.

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1

Imports utils , grDevices , graphics, httr , plyr , scales , stats ,
magrittr , magick , data.tree , ggplot2

NeedsCompilation no

Repository CRAN

Date/Publication 2019-08-19 09:00:02 UTC

R topics documented:

ConstructLocTree	2
ConvertID	3
GetAccessionList	4
GetCross_references_Information	4
GetExpression	5
GetFamily_Domains	6
GetGeneral_Information	6
GetMiscellaneous	7
GetNamesTaxa	8
GetPathology_Biotech	9
GetpdbStructure	9
GetProteinAnnotate	10
GetProteinFunction	11

GetProteinGOInfo	11
GetProteinInteractions	12
GetproteinNetwork	13
GetPTM_Processing	13
GetPublication	14
GetSequences	15
GetSubcellular_location	16
GetTaxaidentifier	16
GetTaxonomic_lineage	17
HandleBadRequests	18
PlotProteinGO_bio	19
PlotProteinGO_cel	19
PlotProteinGO_molc	20
PlotProteinTaxa	20
PlotSummaryInfo	21
UniprotR	21
Index	23

ConstructLocTree	<i>Connect and parse UniProt information.</i>
------------------	---

Description

This Function is used to plot location's Tree in the data of the accession/s in the chromosomes.

Usage

```
ConstructLocTree (ProteinDataObject,directorypath = NULL)
```

Arguments

ProteinDataObject
input a Dataframe returned from GetNamesTaxa function

directorypath path to save txt file containig results retured by the function.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

ConvertID *Connect and parse UniProt database identifiers information.*

Description

The function is work to convert the UniProtKB AC/ID to any database identifiers available by the UniProtKB. For more information about available database identifiers see https://www.uniprot.org/help/api_idmapping.

Usage

```
ConvertID(ProteinAccList , ID_from = "ACC+ID" , ID_to = NULL
, directorypath = NULL)
```

Arguments

ProteinAccList Vector of UniProt Accession/s

ID_from string of database identifier abbreviation, from which the Accession/ID will be converted

ID_to string of database identifier abbreviation, to which the Accession/ID will be converted. default is all database identifier available in UniProtKB

directorypath path to save excel file containig results returned by the function.

Value

DataFrame where column one contains the Accession/ID before conversion and other columns contains the Accession/ID after conversion

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- ConvertID("014520" , "ACC+ID" , "EMBL")
```

GetAccessionList *Connect and parse UniProt information.*

Description

This function can be used to get a list of UniProt Accession/s from a csv file.

Usage

```
GetAccessionList(DataObjPath)
```

Arguments

DataObjPath input path of excel file

Value

a vector of UniProt Accession/s

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

GetCross_references_Information
Connect and parse UniProt Cross-references information.

Description

The function is work to retrieve Cross-references data from UniProt for a list of proteins accessions.

For more information about what included in the Cross-references data see https://www.uniprot.org/help/uniprotkb_column_r

Usage

```
GetCross_references_Information(ProteinAcclList , directorypath = NULL)
```

Arguments

ProteinAcclList Vector of UniProt Accession/s

directorypath path to save excel file containig results returened by the function.

Value

DataFrame where rows names are the accession and columns contains the Cross-references Information of protein from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetCross_references_Information("014520")
```

GetExpression	<i>Connect and parse UniProt Expression information.</i>
---------------	--

Description

The function is work to retrieve Expression data from UniProt for a list of proteins accessions. For more information about what included in the Expression data see https://www.uniprot.org/help/uniprotkb_column_names.

Usage

```
GetExpression(ProteinAcclList , directorypath = NULL)
```

Arguments

ProteinAcclList Vector of UniProt Accession/s
directorypath path to save excel file containig results returned by the function.

Value

DataFrame where rows names are the accession and columns contains the information retrieved from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetExpression("014520")
```

GetFamily_Domains *Connect and parse UniProt Family Domains information.*

Description

The function is work to retrieve Family Domains data from UniProt for a list of proteins accessions. For more information about what included in the Family Domains data see https://www.uniprot.org/help/uniprotkb_colu

Usage

```
GetFamily_Domains(ProteinAccList , directorypath = NULL)
```

Arguments

ProteinAccList Vector of UniProt Accession/s
directorypath path to save excel file containig results returened by the function.

Value

DataFrame where rows names are the accession and columns contains the information retrieved from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetFamily_Domains("014520")
```

GetGeneral_Information *Connect and parse UniProt General Information.*

Description

The function is work to retrieve General Information data from UniProt for a list of proteins accessions. For more information about what included in the General Information data see https://www.uniprot.org/help/uniprotkb_

Usage

```
GetGeneral_Information(ProteinAccList , directorypath = NULL)
```

Arguments

ProteinAccList Vector of UniProt Accession/s
directorypath path to save excel file containig results returened by the function.

Value

DataFrame where rows names are the accession and columns contains the General Information of protein from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetGeneral_Information("014520")
```

GetMiscellaneous *Connect and parse UniProt Miscellaneous information.*

Description

The function is work to retrieve Miscellaneous data from UniProt for a list of proteins accessions. For more information about what included in the Miscellaneous data see https://www.uniprot.org/help/uniprotkb_column

Usage

```
GetMiscellaneous(ProteinAccList , directorypath = NULL)
```

Arguments

ProteinAccList Vector of UniProt Accession/s
directorypath path to save excel file containig results returened by the function.

Value

DataFrame where rows names are the accession and columns contains the information retrieved from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetMiscellaneous("014520")
```

GetNamesTaxa

Connect and parse UniProt Names Taxa information.

Description

The function is work to retrieve Names Taxa data from UniProt for a list of proteins accessions. For more information about what included in the NamesTaxa data see https://www.uniprot.org/help/uniprotkb_column_names.

Usage

```
GetNamesTaxa(ProteinAcclList , directorypath = NULL)
```

Arguments

ProteinAcclList Vector of UniProt Accession/s

directorypath path to save excel file containig results returened by the function.

Value

DataFrame where rows names are the accession and columns contains the information of protein name & taxonomy from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetNamesTaxa("014520")
```

GetPathology_Biotech *Connect and parse UniProt Pathology_Biotech information.*

Description

The function is work to retrieve Pathology_Biotech data from UniProt for a list of proteins accessions. For more information about what included in the Pathology_Biotech data see https://www.uniprot.org/help/uniprotkb_c

Usage

```
GetPathology_Biotech(ProteinAccList , directorypath = NULL)
```

Arguments

ProteinAccList Vector of UniProt Accession/s
directorypath path to save excel file containig results returned by the function.

Value

DataFrame where rows names are the accession and columns contains the information retrieved from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetPathology_Biotech("014520")
```

GetpdbStructure *Connect and parse UniProt information.*

Description

The function is work to retrieving GetpdbStructure and download it to user directory.

Usage

```
GetpdbStructure(ProteinAccList , directorypath = NULL)
```

Arguments

ProteinAcclList input a vector of UniProt Accession/s
directorypath path to save excel file containig results returned by the function.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

GetProteinAnnotate *Connect and parse UniProt information.*

Description

The function is work to retrieve user-defined information data from UniProt for a list of proteins accessions For more information see https://www.uniprot.org/help/uniprotkb_column_names

Usage

```
GetProteinAnnotate(ProteinAcclList , columns)
```

Arguments

ProteinAcclList a vector of UniProt Accession/s
columns a vector of UniProtKB column names

Value

DataFrame where rows names are the accession and columns contains the information retrieved from the UniProt

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetProteinInteractions("014520")
```

GetProteinFunction *Connect and parse UniProt Protein Function information.*

Description

The function is work to retrieve Protein Function data from UniProt for a list of proteins accessions. For more information about what included in the Protein Function data see https://www.uniprot.org/help/uniprotkb_col

Usage

```
GetProteinFunction(ProteinAccList , directorypath = NULL)
```

Arguments

ProteinAccList Vector of UniProt Accession/s
directorypath path to save excel file containig results returned by the function.

Value

DataFrame where rows names are the accession and columns contains the information of protein function roles from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetProteinFunction("014520")
```

GetProteinGOInfo *Connect and parse UniProt proteins gene ontology information.*

Description

The function is work to retrieve proteins gene ontology data from UniProt for a list of proteins accessions. For more information about what included in the proteins gene ontology data see https://www.uniprot.org/help/uniprotkb_col

Usage

```
GetProteinGOInfo(ProteinAccList , directorypath = NULL)
```

Arguments

ProteinAccList Vector of UniProt Accession/s
directorypath path to save excel file containig results returened by the function.

Value

DataFrame where rows names are the accession and columns contains the information of Gene ontology of protein from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetProteinG0Info("014520")
```

GetProteinInteractions

Connect and parse UniProt protein Interactions information.

Description

The function is work to retrieve protein Interactions data from UniProt for a list of proteins accessions. For more information about what included in the protein Interactions data see https://www.uniprot.org/help/uniprotkb_c

Usage

```
GetProteinInteractions(ProteinAccList , directorypath = NULL)
```

Arguments

ProteinAccList Vector of UniProt Accession/s
directorypath path to save excel file containig results returened by the function.

Value

DataFrame where rows names are the accession and columns contains the information of protein Interactions from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetProteinInteractions("ProteinAccession")
```

GetproteinNetwork *Connect and parse stringdb information.*

Description

This function is connecting to stringdb and retrieve all possible interactions for the searched protein/s.

Usage

```
GetproteinNetwork(ProteinAcclList , directorypath = NULL)
```

Arguments

ProteinAcclList input a vector of UniProt Accession/s

directorypath path to save excel file containig results returned by the function.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

GetPTM_Processsing *Connect and parse UniProt PTM_Processsing information.*

Description

The function is work to retrieve PTM_Processsing data from UniProt for a list of proteins accessions. For more information about what included in the PTM_Processsing data see https://www.uniprot.org/help/uniprotkb_co

Usage

```
GetPTM_Processsing(ProteinAcclList, directorypath = NULL)
```

Arguments

ProteinAcclList Vector of UniProt Accession/s

directorypath path to save excel file containig results returned by the function

Value

DataFrame where rows names are the accession and columns contains the information retrieved from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetPTM_Processsing("014520" )
```

GetPublication	<i>Connect and parse UniProt Publication about a protein information.</i>
----------------	---

Description

The function is work to retrieve Publication data from UniProt for a list of proteins accessions. For more information about what included in the Publication data see https://www.uniprot.org/help/uniprotkb_column_names.

Usage

```
GetPublication(ProteinAccList , directorypath = NULL)
```

Arguments

ProteinAccList Vector of UniProt Accession/s

directorypath path to save excel file containig results returned by the function.

Value

DataFrame where rows names are the accession and columns contains the Publication of protein from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetPublication("014520")
```

GetSequences

Connect and parse UniProt Sequences information.

Description

The function is work to retrieve Sequences data from UniProt for a list of proteins accessions. For more information about what included in the Sequences data see https://www.uniprot.org/help/uniprotkb_column_names.

Usage

```
GetSequences(ProteinAcclList, directorypath = NULL)
```

Arguments

ProteinAcclList Vector of UniProt Accession/s

directorypath path to save excel file containig results returened by the function

Value

DataFrame where rows names are the accession and columns contains the information retrieved from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetSequences("014520")
```

GetSubcellular_location

Connect and parse UniProt protein Subcellular location information.

Description

The function is work to retrieve protein Subcellular location data from UniProt for a list of proteins accessions. For more information about what included in the protein Subcellular location data see https://www.uniprot.org/help/uniprotkb_column_names.

Usage

```
GetSubcellular_location(ProteinAccList, directorypath = NULL)
```

Arguments

ProteinAccList Vector of UniProt Accession/s.

directorypath path to save excel file containig results returened by the function.

Value

DataFrame where rows names are the accession and columns contains the information about Sub-cellular location of protein from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

GetTaxaidentifier

Connect and parse UniProt protein taxonomic information.

Description

The function is work to retrieve Miscellaneous data from UniProt for a list of proteins accessions. For more information about what included in the Miscellaneous data see https://www.uniprot.org/help/uniprotkb_column_names.

Usage

```
GetTaxaidentifier(ProteinAccList, directorypath = NULL)
```

Arguments

ProteinAccList Vector of UniProt Accession/s
directorypath path to save excel file containig results returened by the function.

Value

DataFrame where rows names are the accession and columns contains the Taxonomic Information of protein from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetTaxaidentifier("014520")
```

GetTaxonomic_lineage *Connect and parse UniProt Taxonomic_lineage information.*

Description

The function is work to retrieve Taxonomic_lineage data from UniProt for a list of proteins accessions. For more information about what included in the Taxonomic_lineage data see https://www.uniprot.org/help/uniprotkb_

Usage

```
GetTaxonomic_lineage(ProteinAccList, directorypath = NULL)
```

Arguments

ProteinAccList Vector of UniProt Accession/s
directorypath path to save excel file containig results returened by the function.

Value

DataFrame where rows names are the accession and columns contains the information retrieved from the UniProt

Note

The function also, Creates a csv file with the retrieved information.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
Obj <- GetTaxonomic_lineage("014520")
```

HandleBadRequests	<i>Handle bad requests This Function is used to handle possible errors when trying to get url response.</i>
-------------------	---

Description

Handle bad requests This Function is used to handle possible errors when trying to get url response.

Usage

```
HandleBadRequests(RequestCode)
```

Arguments

RequestCode Response returned from url

Value

None

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Examples

```
HandleBadRequests(400)
```

PlotProteinGO_bio *Connect and parse UniProt information.*

Description

This Function is used to plot biological process data retrieved from UniprotR Function "GetProteinGOInfo".

Usage

```
PlotProteinGO_bio(GO_df , dir_path = NA)
```

Arguments

GO_df Dataframe.
dir_path path to save excel file containig results returened by the function (default = NA).

Note

if no dir_path was given (default = NA) the function will only view the plot and will not save it

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

PlotProteinGO_cel *Connect and parse UniProt information.*

Description

This Function is used to plot cellular components data retrieved from UniprotR Function "GetProteinGOInfo".

Usage

```
PlotProteinGO_cel(GO_df , dir_path = NA)
```

Arguments

GO_df Dataframe.
dir_path path to save excel file containig results returened by the function (default = NA).

Note

if no dir_path was given (default = NA) the function will only view the plot and Will not save it

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

PlotProteinGO_molc *Connect and parse UniProt information.*

Description

This Function is used to plot molecular functions data retrieved from UniprotR Function "GetProteinGOInfo".

Usage

```
PlotProteinGO_molc(GO_df , dir_path = NA)
```

Arguments

GO_df	Dataframe.
dir_path	path to save excel file containig results returened by the function (default = NA).

Note

if no dir_path was given (default = NA) the function will only view the plot and will not save it

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

PlotProteinTaxa *Connect and parse UniProt information.*

Description

This Function is used to plot different taxas found of the accessions.

Usage

```
PlotProteinTaxa(ProteinDataObject , directorypath = NULL)
```

Arguments

ProteinDataObject
input a Dataframe of proteins as rownames.
directorypath path to save excel file containig results returened by the function.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

PlotSummaryInfo *Connect and parse UniProt information.*

Description

This Function is used to plot location's frequency in the data of the accession/s in the chromosomes.

Usage

PlotSummaryInfo (ProteinDataObject,directorypath = NULL)

Arguments

ProteinDataObject
input a Dataframe returned from GetNamesTaxa function
directorypath path to save excel file containig results returened by the function.

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

UniprotR *UniprotR: A package for retrieving proteins information from UniProt.*

Description

UniprotR provides different functions to retrieve proteins information provided by the UniProt. UniprotR could be divided into two categories one for retrieving and one for visualizing the data.

The first category for retrieving data

ConstructLocTree
GetCross_references_Information
GetExpression
GetFamily_Domains
GetGeneral_Information
GetMiscellaneous
GetNamesTaxa
GetPathology_Biotech
GetProteinFunction
GetProteinGOInfo
GetProteinInteractions
GetProteoAnnotator
GetPTM_Processsing
GetPublication
GetSequences
GetSubcellular_location
GetTaxaidentifier
GetTaxonomic_lineage
each function has a description that could be reached by »> ?Function_Name

The second category for visualizing data

PlotProteinTaxa
PlotSummaryInfo
PlotProteinGO_bio
PlotProteinGO_cel
PlotProteinGO_molc
each function has a description that could be reached by »> ?Function_Name

Author(s)

Mohmed Soudy <Mohamed.soudy@57357.com> and Ali Mostafa <ali.mo.anwar@std.agr.cu.edu.eg>

Index

ConstructLocTree, [2](#)
ConvertID, [3](#)

GetAccessionList, [4](#)
GetCross_references_Information, [4](#)
GetExpression, [5](#)
GetFamily_Domains, [6](#)
GetGeneral_Information, [6](#)
GetMiscellaneous, [7](#)
GetNamesTaxa, [8](#)
GetPathology_Biotech, [9](#)
GetpdbStructure, [9](#)
GetProteinAnnotate, [10](#)
GetProteinFunction, [11](#)
GetProteinGOInfo, [11](#)
GetProteinInteractions, [12](#)
GetproteinNetwork, [13](#)
GetPTM_Processing, [13](#)
GetPublication, [14](#)
GetSequences, [15](#)
GetSubcellular_location, [16](#)
GetTaxaidentifier, [16](#)
GetTaxonomic_lineage, [17](#)

HandleBadRequests, [18](#)

PlotProteinGO_bio, [19](#)
PlotProteinGO_cel, [19](#)
PlotProteinGO_molc, [20](#)
PlotProteinTaxa, [20](#)
PlotSummaryInfo, [21](#)

UniprotR, [21](#)
UniprotR-package (UniprotR), [21](#)