

Package ‘UniprotR’

August 27, 2023

Title Retrieving Information of Proteins from Uniprot

Version 2.3.0

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

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, For detailed information kindly read the publication <<https://www.sciencedirect.com/science/article/pii/S1874391919303859>>.

License GPL-3

Encoding UTF-8

RoxygenNote 7.2.3

Imports utils , grDevices , stats , grid , graphics , httr , plyr ,
dplyr , scales , magrittr , magick , data.tree , ggplot2 ,
tidyverse , gridExtra , ggpubr , curl , networkD3 , stringr ,
qdapRegex , htmlwidgets , alakazam , gprofiler2 , progress , ggsci

URL <https://github.com/Proteomicslab57357/UniprotR>

BugReports <https://github.com/Proteomicslab57357/UniprotR/issues>

NeedsCompilation no

Repository CRAN

Date/Publication 2023-08-27 14:10:02 UTC

R topics documented:

ConstructGenesTree	3
ConstructLocTree	3
ConvertID	4
Enrichment.BP	5
Enrichment.CC	5
Enrichment.KEGG	6
Enrichment.MF	6

Enrichment.REAC	7
Get.diseases	7
GetAccessionList	8
GetExpression	8
GetFamily_Domains	9
GetGeneral_Information	10
GetMiscellaneous	11
GetNamesTaxa	11
GetPathology_Biotech	12
GetpdbStructure	13
GetProteinAnnotate	13
GetProteinFunction	14
GetProteinGOInfo	15
GetProteinInteractions	15
GetproteinNetwork	16
GetproteinNetwork_all	17
GetProteomeFasta	17
GetProteomeInfo	18
GetPTM_Processing	18
GetPublication	19
GETSeqFastaUniprot	20
GetSeqLength	20
GetSequenceIso	21
GetSequences	22
GetStructureInfo	22
GetSubcellular_location	23
Goparse	24
HandleBadRequests	24
Pathway.Enr	25
Plot.GOMolecular	25
Plot.GOSubCellular	26
PlotAcidity	26
PlotCharge	27
PlotChromosomeInfo	27
PlotEnrichedGO	28
PlotEnrichedPathways	28
PlotGenesNetwork	29
PlotGOAll	30
PlotGOBiological	30
PlotGoInfo	31
PlotGoterms	31
PlotGravy	32
PlotPhysicochemical	32
PlotproteinExist	33
PlotProteinGO_bio	33
PlotProteinGO_cel	34
PlotProteinGO_molc	34
PlotproteinStatus	35

<i>ConstructGenesTree</i>	3
PlotProteinTaxa	35
ProcessAcc	36
Index	37

ConstructGenesTree *Connect and parse UniProt information.*

Description

This Function is used to plot Genes Tree in the data of the accession/s.

Usage

`ConstructGenesTree(ProteinDataObject,directorypath = NULL)`

Arguments

ProteinDataObject input a Dataframe returned from GetNamesTaxa function
 directorypath path to save txt 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>

ConstructLocTree *Connect and parse UniProt information.*

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 returened 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/id_mapping. see https://raw.githubusercontent.com/MohmedSoudy/UniprotR/master/uniprot_ids.csv

Usage

```
ConvertID(ProteinAcclList , ID_from = "UniProtKB_AC-ID" , ID_to = NULL
, directorypath = NULL)
```

Arguments

ProteinAcclList 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>

Enrichment.BP *Connect and parse UniProt information*

Description

This function is used for Enrichment analysis of biological process of given list of genes or proteins

Usage

```
Enrichment.BP(Accs,OS="hsapiens",p_value=0.05,directorypath=NULL)
```

Arguments

Accs	Vector of UniProt Accession/s or genes
OS	organism name Example: human - 'hsapiens', mouse - 'mmusculus'
p_value	custom p-value threshold for significance, default = 0.05
directorypath	Path to save output plot

Author(s)

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

Enrichment.CC *Connect and parse UniProt information*

Description

This function is used for Enrichment analysis of cellular component of given list of genes or proteins

Usage

```
Enrichment.CC(Accs,OS="hsapiens",p_value=0.05,directorypath=NULL)
```

Arguments

Accs	Vector of UniProt Accession/s or genes
OS	organism name Example: human - 'hsapiens', mouse - 'mmusculus'
p_value	custom p-value threshold for significance, default = 0.05
directorypath	Path to save output plot

Author(s)

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

Enrichment.KEGG *Connect and parse UniProt information*

Description

This function is used for Enrichment analysis of given list of genes or proteins from KEGG database

Usage

```
Enrichment.KEGG(Accs,OS="hsapiens",p_value=0.05,directorypath=NULL)
```

Arguments

Accs	Vector of UniProt Accession/s or genes
OS	organism name Example: human - 'hsapiens', mouse - 'mmusculus'
p_value	custom p-value threshold for significance, default = 0.05
directorypath	Path to save output plot

Author(s)

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

Enrichment.MF *Connect and parse UniProt information*

Description

This function is used for Enrichment analysis of Molecular function of given list of genes or proteins

Usage

```
Enrichment.MF(Accs,OS="hsapiens",p_value=0.05,directorypath=NULL)
```

Arguments

Accs	Vector of UniProt Accession/s or genes
OS	organism name Example: human - 'hsapiens', mouse - 'mmusculus'
p_value	custom p-value threshold for significance, default = 0.05
directorypath	Path to save output plot

Author(s)

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

Enrichment.REAC *Connect and parse UniProt information*

Description

This function is used for Enrichment analysis of given list of genes or proteins from REACTOME

Usage

```
Enrichment.REAC(Accs,OS="hsapiens",p_value=0.05,directorypath=NULL)
```

Arguments

Accs	Vector of UniProt Accession/s or genes
OS	organism name Example: human - 'hsapiens', mouse - 'mmusculus'
p_value	custom p-value threshold for significance, default = 0.05
directorypath	Path to save output plot

Author(s)

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

Get.diseases *Connect and parse UniProt information.*

Description

This Function is used to get diseases associated with proteins.

Usage

```
Get.diseases(Pathology_object , directorypath = NULL)
```

Arguments

Pathology_object	Dataframe retrieved from UniprotR Function "GetPathology_Biotech"
directorypath	path to save Output file

Author(s)

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

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>

GetExpression *Connect and parse UniProt Expression information.*

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/return_fields.

Usage

```
GetExpression(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>

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/return_fields.

Usage

```
GetFamily_Domains(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>

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/return_fields

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/return_fields.

Usage

```
GetMiscellaneous(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 <- 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/return_fields.

Usage

```
GetNamesTaxa(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 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/return_fields

Usage

```
GetPathology_Biotech(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 <- 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

ProteinAccList input a vector of UniProt Accession/s
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>

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(ProteinAccList , columns)
```

Arguments

ProteinAccList 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/return_fields.

Usage

```
GetProteinFunction(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 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/return_fields

Usage

```
GetProteinGOInfo(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 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 <- GetProteinGOInfo("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/return_fields

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(ProteinAccList , directorypath = NULL)
```

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

Arguments

ProteinAccList input a vector of UniProt Accession/s
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>

GetproteinNetwork_all *Connect and parse stringdb information.*

Description

This function is connecting to stringdb and retrieve PPI between input list

Usage

```
GetproteinNetwork_all(ProteinAcclist , directorypath = NULL, SpeciesID = 9606)
```

```
GetproteinNetwork_all(ProteinAcclist , directorypath = NULL,SpeciesID = 9606)
```

Arguments

ProteinAcclist input a vector of UniProt Accession/s

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

SpeciesID Taxonomic id of accession's species ex. homo sapiens 9606

Author(s)

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

GetProteomeFasta *Connect and DOWNLOAD Proteome info.*

Description

The function is work to retrieve proteome information in FASTA format based on proteome id.

Usage

```
GetProteomeFasta(ProteomeID , directorypath = NULL)
```

Arguments

ProteomeID Proteome ID from UniProt

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

Note

The function Download fasta format of proteome.

Author(s)

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

GetProteomeInfo *Connect and DOWNLOAD Proteome info.*

Description

The function is work to retrieve proteome information based on proteome id.

Usage

```
GetProteomeInfo(ProteomeID , directorypath = NULL)
```

Arguments

ProteomeID Proteome ID from UniProt
directorypath path to save CSV file containig results returened by the function.

Note

The function Download csv Info of proteome.

Author(s)

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

GetPTM_Processing *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/return_fields

Usage

```
GetPTM_Processing(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 <- GetPTM_Processing("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/return_fields.

Usage

```
GetPublication(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 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>

GETSeqFastaUniprot *Connect and parse UniProt information.*

Description

This Function is used to get Sequence information of accession/s from Uniprot as a Fasta file.

Usage

```
GETSeqFastaUniprot(Accessions,FilePath = NULL, FileName = NULL)
```

Arguments

Accessions	Vector of UniProt Accession/s
FilePath	path of directory to save the output fasta.
FileName	Name of the fasta file.

Author(s)

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

GetSeqLength *Connect and parse UniProt Sequences information.*

Description

The function is work to retrieve Sequence's Length data from Uniparc for a list of proteins accessions. This function was added to overcome the NAs returned when Uniprot database deleted the protein from the database

Usage

```
GetSeqLength(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 <- GetSeqLength("014520")
```

GetSequenceIso *Connect and parse UniProt Sequences information.*

Description

The function is work to retrieve protein's Sequence data from Uniparc for a list of proteins accessions. This function was added to handle isoformes

Usage

```
GetSequenceIso(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 <- GetSequenceIso("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/return_fields.

Usage

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

GetStructureInfo *Connect and parse UniProt protein Structure information.*

Description

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

Usage

```
GetStructureInfo(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 Structural information of protein from the UniProt

Note

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

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/return_fields.

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>

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

Description

This Function is used to parse data retrieved from UniprotR Function "GetProteinGOInfo".

Usage

```
Goparse(GOobj , index = 3)
```

Arguments

GOobj	Dataframe.
index	index of Go term in GoObj

Author(s)

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

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)
```

Pathway.Enr *Connect and parse UniProt information*

Description

This function is used for Enrichment analysis of given list of genes or proteins

Usage

```
Pathway.Enr(Accs,OS="hsapiens",p_value=0.05,directorypath=NULL)
```

Arguments

Accs	Vector of UniProt Accession/s or genes
OS	organism Example: human - 'hsapiens' for more info https://biit.cs.ut.ee/gprofiler/page/organism-list
p_value	custom p-value threshold for significance, default = 0.05
directorypath	Path to save output plot

Author(s)

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

Plot.GOMolecular *Connect and parse UniProt information.*

Description

This Function is used to plot Molecular function of proteins.

Usage

```
Plot.GOMolecular(GOObj, Top = 10, directorypath = NULL)
```

Arguments

GOObj	Dataframe returned from UniprotR Function "GetProteinGOInfo"
Top	Number of molecular functions to be visualized
directorypath	path to save Output plot.

Author(s)

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

Plot.GOSubCellular *Connect and parse UniProt information.*

Description

This Function is used to plot subcellular localization of proteins.

Usage

```
Plot.GOSubCellular(GOObj, Top = 10, directorypath = NULL)
```

Arguments

GOObj Dataframe returned from UniprotR Function "GetProteinGOInfo"
Top Number of molecular functions to be visualized
directorypath path to save Output plot.

Author(s)

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

PlotAcidity *Connect and parse UniProt information.*

Description

This Function is used to plot proteins acidity retrieved from "GetSequences" Function.

Usage

```
PlotAcidity(SeqDataObjPath , directorypath = NULL)
```

Arguments

SeqDataObjPath Dataframe retrieved from UniprotR Function "GetSequences"
directorypath path to save Output plot.

Author(s)

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

PlotCharge

Connect and parse UniProt information.

Description

This Function is used to plot proteins charge retrieved from "GetSequences" Function.

Usage

```
PlotCharge(SeqDataObjPath , directorypath = NULL)
```

Arguments

SeqDataObjPath Dataframe retrieved from UniprotR Function "GetSequences"

directorypath path to save Output plot.

Author(s)

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

PlotChromosomeInfo

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

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

Arguments

ProteinDataObject

input a Dataframe returned from GetNamesTaxa function

directorypath path to save files returned by the function.

Author(s)

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

PlotEnrichedGO *Connect and parse UniProt information*

Description

This function is used to generate a combined plot for the enriched Gene Ontology terms

Usage

```
PlotEnrichedGO(Accs,OS="hsapiens",p_value=0.05,Path=NULL,theme="aaas",width=7,height=7)
```

Arguments

Accs	Vector of UniProt Accession/s or genes
OS	organism name Example: human - 'hsapiens', mouse - 'mmusculus'
p_value	custom p-value threshold for significance, default = 0.05
Path	Path to save output plot
theme	optional parameter to generate specific theme for journals ex: "aaas", "nature", "lancet", "jama"
width	width of the generated plot
height	height of the generated plot

Author(s)

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

PlotEnrichedPathways *Connect and parse UniProt information*

Description

This function is used to generate a combined plot for the enriched pathways from KEGG and REACTOME

Usage

```
PlotEnrichedPathways(Accs,OS="hsapiens",p_value=0.05,Path=NULL,theme="aaas",w=w,h=h)
```

Arguments

Accs	Vector of UniProt Accession/s or genes
OS	organism name Example: human - 'hsapiens', mouse - 'mmusculus'
p_value	custom p-value threshold for significance, default = 0.05
Path	Path to save output plot
theme	optional parameter to generate specific theme for journals ex: "aaas", "nature", "lancet", "jama"
w	width of the generated plot
h	height of the generated plot

Author(s)

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

PlotGenesNetwork *Connect and parse UniProt information.*

Description

This Function is used to cluster proteins based on primary genes retrieved from "GetNamesTaxa" Function.

Usage

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

Arguments

ProteinDataObject	Dataframe retrieved from UniprotR Function "GetNamesTaxa"
directorypath	path to save Output plot.

Author(s)

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

PlotGOAll *Connect and parse UniProt information.*

Description

This Function is used to plot the retrieved Gene Ontology from function 'GetProteinGOInfo'.

Usage

```
PlotGOAll(GOobj, Top = 10, directorypath = NULL, width = width, height = height)
```

Arguments

GOobj	Dataframe returned from UniprotR Function "GetProteinGOInfo"
Top	Number of molecular functions to be visualized
directorypath	path to save Output plot.
width	width of the generated plot
height	height of the generated plot

Author(s)

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

PlotGOBiological *Connect and parse UniProt information.*

Description

This Function is used to plot Biological process of proteins.

Usage

```
PlotGOBiological(GOobj, Top = 10, directorypath = NULL)
```

Arguments

GOobj	Dataframe returned from UniprotR Function "GetProteinGOInfo"
Top	Number of molecular functions to be visualized
directorypath	path to save Output plot.

Author(s)

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

PlotGoInfo

Connect and parse UniProt information.

Description

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

Usage

```
PlotGoInfo(GOobj , directorypath = NULL)
```

Arguments

GOobj Dataframe retrieved from UniprotR Function "GetProteinGOInfo".
directorypath path to save excel file containig results returned by the function (default = NA).

Author(s)

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

PlotGoterms

Connect and parse UniProt information.

Description

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

Usage

```
PlotGoterms(GOobj , directorypath = NULL)
```

Arguments

GOobj Dataframe.
directorypath path to save plot returned by function (default = NA).

Author(s)

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

PlotGravy

Connect and parse UniProt information.

Description

This Function is used to plot proteins gravity index retrieved from "GetSequences" Function.

Usage

```
PlotGravy(SeqDataObjPath , directorypath = NULL)
```

Arguments

SeqDataObjPath Dataframe retrieved from UniprotR Function "GetSequences"

directorypath path to save Output plot.

Author(s)

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

PlotPhysicochemical

Connect and parse UniProt information.

Description

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

Usage

```
PlotPhysicochemical(SeqDataObjPath , directorypath = NULL)
```

Arguments

SeqDataObjPath Dataframe returned from GetSequence function.

directorypath Path to save Physicochemical properties plot.

Author(s)

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

PlotproteinExist *Connect and parse UniProt information.*

Description

This Function is used to plot protein status in the data of the accession/s.

Usage

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

Arguments

ProteinDataObject input a Dataframe returned from GetMiscellaneous function
directorypath path to save files returned by the function.

Author(s)

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

PlotProteinGO_bio *Connect and parse UniProt information.*

Description

This Function is used to plot biological process data retrieved from UniprotR Function "GetProtein-GOInfo".

Usage

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

Arguments

GO_df Dataframe.
dir_path path to save files returned 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 files returned 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 files returned 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>

PlotproteinStatus *Connect and parse UniProt information.*

Description

This Function is used to plot protein status in the data of the accession/s.

Usage

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

Arguments

ProteinDataObject
 input a Dataframe returned from GetMiscellaneous function
directorypath path to save files returned by the function.

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 files returned by the function.

Author(s)

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

ProcessAcc

Connect and parse UniProt information.

Description

This Function is used to check validity of input accessions the data of the accession/s.

Usage

ProcessAcc(Accessions)

Arguments

Accessions accession list returned from GetAccession function

Author(s)

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

Index

ConstructGenesTree, [3](#)
ConstructLocTree, [3](#)
ConvertID, [4](#)

Enrichment.BP, [5](#)
Enrichment.CC, [5](#)
Enrichment.KEGG, [6](#)
Enrichment.MF, [6](#)
Enrichment.REAC, [7](#)

Get.diseases, [7](#)
GetAccessionList, [8](#)
GetExpression, [8](#)
GetFamily_Domains, [9](#)
GetGeneral_Information, [10](#)
GetMiscellaneous, [11](#)
GetNamesTaxa, [11](#)
GetPathology_Biotech, [12](#)
GetpdbStructure, [13](#)
GetProteinAnnotate, [13](#)
GetProteinFunction, [14](#)
GetProteinGOInfo, [15](#)
GetProteinInteractions, [15](#)
GetproteinNetwork, [16](#)
GetproteinNetwork_all, [17](#)
GetProteomeFasta, [17](#)
GetProteomeInfo, [18](#)
GetPTM_Processing, [18](#)
GetPublication, [19](#)
GETSeqFastaUniprot, [20](#)
GetSeqLength, [20](#)
GetSequenceIso, [21](#)
GetSequences, [22](#)
GetStructureInfo, [22](#)
GetSubcellular_location, [23](#)
Goparse, [24](#)

HandleBadRequests, [24](#)

Pathway.Enr, [25](#)

Plot.GOMolecular, [25](#)
Plot.GOSubCellular, [26](#)
PlotAcidity, [26](#)
PlotCharge, [27](#)
PlotChromosomeInfo, [27](#)
PlotEnrichedGO, [28](#)
PlotEnrichedPathways, [28](#)
PlotGenesNetwork, [29](#)
PlotGOAll, [30](#)
PlotGOBiological, [30](#)
PlotGoInfo, [31](#)
PlotGoterms, [31](#)
PlotGravy, [32](#)
PlotPhysicochemical, [32](#)
PlotproteinExist, [33](#)
PlotProteinGO_bio, [33](#)
PlotProteinGO_cel, [34](#)
PlotProteinGO_molc, [34](#)
PlotproteinStatus, [35](#)
PlotProteinTaxa, [35](#)
ProcessAcc, [36](#)