

# LowMACAAnnotation

October 16, 2024

---

getLowMACAAnnotationData

*Retrieve Data from LowMACAAnnotation Package*

---

## Description

Functions to retrieve protein level, Pfam level and aliases data from LowMACAAnnotation Package.

## Usage

```
## Retrieve the dataset containing protein level information  
getMyUni()
```

```
## Retrieve the dataset containing Pfam level information  
getMyPfam()
```

```
## Retrieve the dataset containing Gene Symbols aliases  
getMyAlias()
```

## Value

All the functions return dataframes

## See Also

[myUni](#) [myPfam](#) [myAlias](#)

## Examples

```
myUni <- getMyUni()  
myPfam <- getMyPfam()  
myAlias <- getMyAlias()
```

---

myAlias

*Gene Symbol aliases for LowMACA package*

---

### Description

A simple parsing of the HGNC database to map aliases and previous symbols to correct official Gene Symbols

### Usage

```
data("myAlias")
```

### Format

A data frame with the following 2 variables.

Alias a character vector representing all the possible aliases and previous symbols for official Gene Symbols

Official\_Gene\_Symbol a character vector representing the approved and official Gene Symbol for HGNC database

Locus\_Group a character vector representing all the possible locus groups in HGNC database, like protein coding, RNA, pseudogene etc.

Locus\_Type a character vector representing all the possible locus types in HGNC database. It is a specification of locus group

MappedByLowMACA a character vector of yes and no if the gene is included in myUni.RData

### Source

**HGNC**

### Examples

```
#Load myAlias and show its structure
myAlias <- getMyAlias()
str(myAlias)
```

---

myPfam

*Pfam-A for LowMACA package*

---

### Description

This dataset comprises all the Pfam-A entries mapped for LowMACA

### Usage

```
data("myPfam")
```

**Format**

A data frame with the following 11 variables.

**Entry** a character vector of Uniprot entries

**Envelope\_Start** a numeric vector of starts of the pfam domain relative to the reference protein

**Envelope\_End** a numeric vector of ends of the pfam domain relative to the reference protein

**Pfam\_ID** a character vector of Pfam IDs in the form of PF##### supported by LowMACA

**Pfam\_Name** a character vector of full Pfam domain names

**Type** a character vector. One of the following: "Domain" "Family" "Repeat" or "Motif"

**Clan\_ID** a numeric vector of Clan IDs, a sort of families of Pfam domains

**Entrez** a numeric vector of Entrez IDs

**UNIPROT** a character vector of Uniprot entries in format "name\_HUMAN"

**Gene\_Symbol** a character vector of official Gene Symbols

**Pfam\_Fasta** a character vector of amino acid sequences of corresponding Pfam

**Details**

This is the result of a merge between Pfam-A, Uniprot, HGNC databases

**Source**

[Pfam website](#)

**Examples**

```
#Load and show contents of myPfam
data(myPfam)
str(myPfam)
```

---

myUni

*Uniprot database for LowMACA*

---

**Description**

This dataset comprises all the Uniprot entries mapped for LowMACA

**Usage**

```
data("myUni")
```

**Format**

A data frame with 9 variables.

**Gene\_Symbol** a character vector of official Gene Symbols

**Entrez** a numeric vector of Entrez IDs

**UNIPROT** a character vector of Uniprot entries in "name\_HUMAN" format

**Entry** a character vector of Uniprot entries

**HGNC** a character vector of gene names as HGNC numbers

**Approved\_Name** a character vector of approved extended gene names

**Protein.name** a character vector of approved extended protein names

**Chromosome** a character vector of chromosomic cytoband positions

**AMINO\_SEQ** a character vector of amino acid sequences for Uniprot entries

### Details

This dataset is a 1 to 1 mapping of genes to their canonical proteins or main proteins

### Source

[Uniprot](#)

### Examples

```
#Load dataset and show its structure
data(myUni)
str(myUni)
```

# Index

## \* datasets

myAlias, 2

myPfam, 2

myUni, 3

getLowMACAAnnotationData, 1

getMyAlias (getLowMACAAnnotationData), 1

getMyPfam (getLowMACAAnnotationData), 1

getMyUni (getLowMACAAnnotationData), 1

myAlias, 1, 2

myPfam, 1, 2

myUni, 1, 3