

# Package ‘MIGSAdata’

May 9, 2023

**Title** MIGSA vignette data

**Version** 1.24.0

**Author** Juan C. Rodriguez, Cristobal Fresno, Andrea S. Llera and Elmer A. Fernandez

**Maintainer** Juan C. Rodriguez <jcrodriguez@bdmg.com.ar>

**Description** MIGSA vignette data. The MIGSAdata package provides several data objects needed by MIGSA package to correctly generate its vignette, and follow it step by step.

**URL** <http://www.bdmg.com.ar/>

**biocViews** ExperimentData, RepositoryData, ReproducibleResearch, CancerData, BreastCancerData, MicroarrayData

**License** GPL (>= 2)

**Depends** R (>= 3.4)

**Suggests** BiocGenerics

**RoxygenNote** 5.0.1

**git\_url** <https://git.bioconductor.org/packages/MIGSAdata>

**git\_branch** RELEASE\_3\_17

**git\_last\_commit** 88d4b4a

**git\_last\_commit\_date** 2023-04-25

**Date/Publication** 2023-05-09

## R topics documented:

bcMigsResAsList . . . . .	2
mGSZspeedup . . . . .	2
MIGSAdata . . . . .	3
pbcmcData . . . . .	3
tcgaMAdata . . . . .	4
tcgaRNAseqData . . . . .	4
<b>Index</b>	<b>5</b>

---

bcMigsResAsList	<i>A MIGSA results object obtained from breast cancer data.</i>
-----------------	---

---

### Description

A MIGSA results object obtained using the datasets:

- mainz (microarray).
- nki (microarray).
- tcga (microarray).
- tcga (RNAseq).
- transbig (microarray).
- unt (microarray).
- upp (microarray).
- vdx (microarray).

### Format

A list with two main components to create the MIGSAres object

### Details

Each dataset subjects were classified using the PAM50 algorithm. For this analysis only Basal and Luminal A subjects were kept (this was the contrast used, i.e., Basal vs. LumA).

Datasets were tested for gene set enrichment over:

- the org.Hs.eg.db Gene Ontology gene sets, resulting in 14,291 as BP, 1,692 CC and 4,263 MF.
- KEGG gene sets downloaded from enrichr database resulting in 179 gene sets.

In order to get the MIGSAres object follow: `library(MIGSAdata); library(MIGSA); data(bcMigsResAsList); bcMigsRes <- MIGSA::MIGSAres.data.table(bcMigsResAsList$dframe, bcMigsResAsList$genesRank);`

---

mGSZspeedup	<i>Speedup between mGSZ and MIGSAmGSZ.</i>
-------------	--

---

### Description

A list with several benchmark timings used to show MIGSAmGSZ's speedup against mGSZ.

### Format

A list with some timing benchmarks.

---

MIGSAdata	<i>MIGSAdata package</i>
-----------	--------------------------

---

**Description**

The MIGSAdata package provides several data objects needed by MIGSA package to correctly generate its vignette, and follow it step by step.

**Author(s)**

Juan C. Rodriguez <jcrodriguez@bdmg.com.ar>, Cristobal Fresno <cfresno@bdmg.com.ar>, Andrea S. Llera <allera@leloir.org.ar> and Elmer A. Fernandez <efernandez@bdmg.com.ar>

**References**

1. Rodriguez, J. C., Gonzalez, G. A., Fresno, C., Llera, A. S., & Fernandez, E. A. (2016). Improving information retrieval in functional analysis. *Computers in Biology and Medicine*, 79, 10-20.

---

pbcmcData	<i>Pbcmc breast cancer microarray data.</i>
-----------	---

---

**Description**

Pbcmc package lets download breast cancer microarray datasets, for each, here we include its expression matrix and each subjects subtypes. Each dataset subjects were classified using the PAM50 algorithm. Only Basal and Luminal A subjects were kept. Datasets are:

- mainz.
- nki.
- transbig.
- unt.
- upp.
- vdx.

Details in how this datasets were obtained can be find in MIGSA's gettingPbcmcData vignette.

**Format**

A list containing six lists (one per dataset), each one contains gene expression as data.frame and subtypes as vector.

---

`tcgaMAdata`*TCGA breast cancer microarray data.*

---

**Description**

TCGA breast cancer microarray expression matrix and each subjects subtypes. Each dataset subjects were classified using the PAM50 algorithm. Only Basal and Luminal A subjects were kept. Details in how this dataset was obtained can be find in MIGSA's `gettingTcgaData` vignette.

**Format**

A list with gene expression as `data.frame` and subtypes as vector.

---

`tcgaRNAseqData`*TCGA breast cancer RNAseq data.*

---

**Description**

TCGA breast cancer RNAseq expression matrix and each subjects subtypes. Each dataset subjects were classified using the PAM50 algorithm. Only Basal and Luminal A subjects were kept. Details in how this dataset was obtained can be find in MIGSA's `gettingTcgaData` vignette.

**Format**

A list with gene expression as `data.frame` and subtypes as vector.

# Index

- \* **MIGSA**

  - MIGSAdata, [3](#)

- \* **data**

  - MIGSAdata, [3](#)

- \* **objects**

  - MIGSAdata, [3](#)

bcMigsResAsList, [2](#)

mGSZspeedup, [2](#)

MIGSAdata, [3](#)

pbcmcData, [3](#)

tcgMAdata, [4](#)

tcgARNaseqData, [4](#)