

# Package ‘spqnData’

September 26, 2024

**Version** 1.16.0

**Title** Data for the spqn package

**Description** Bulk RNA-seq from GTEx on 4,000 randomly selected, expressed genes. Data has been processed for co-expression analysis.

**Depends** R (>= 4.0), SummarizedExperiment

**License** Artistic-2.0

**LazyData** FALSE

**biocViews** Homo\_sapiens\_Data, ExpressionData, Tissue, RNASeqData

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

**git\_branch** RELEASE\_3\_19

**git\_last\_commit** 8f3ef94

**git\_last\_commit\_date** 2024-04-30

**Repository** Bioconductor 3.19

**Date/Publication** 2024-09-26

**Author** Yi Wang [cre, aut],  
Kasper Daniel Hansen [aut]

**Maintainer** Yi Wang <yiwangthu5@gmail.com>

## Contents

|                   |          |
|-------------------|----------|
| gtex.4k . . . . . | 2        |
| <b>Index</b>      | <b>3</b> |

---

`gtex.4k`*Example data for the spqn package.*

---

**Description**

A random sample of 4,000 expressed genes (protein-coding or lincRNAs) from GTEx v6p. The tissue is Adipose Subcutaneous.

**Usage**

```
data("gtex.4k")
```

**Format**

An object of class SummarizedExperiment.

**Details**

Data is 350 samples from GTEx v6p. The tissue is Adipose Subcutaneous.

We first selected protein-coding or lincRNAs based on the supplied annotation files. Next we kept genes with a median  $\log_2(\text{RPKM})$  expression greater than zero. This resulted in a data matrix with 12,267 genes of which 11,911 are protein-coding. We stored the mean expression value per gene in `rowData(gtex.4k)$ave_logrpkm`.

We next mean centered and variance scaled the expression values so all genes have zero mean and variance 1. We then removed 4 principal components from this data matrix using the `removePrincipalComponents` function from the WGCNA package.

Finally, we randomly selected 4,000 genes.

Additional information on the genes are present in the `rowData`. The type of gene (lincRNA or protein-coding) is present in the `gene_type` column. The average expression of each gene on the  $\log_2(\text{RPKM})$ -scale, prior to removing principal components, are present in the `ave_logrpkm` column.

**Source**

Original data from [gtexportal.org](http://gtexportal.org). A script for downloading and processing the paper is included in `scripts/gtex.Rmd`.

**Examples**

```
data(gtex.4k)
```

# Index

\* **datasets**

gtex.4k, [2](#)

gtex.4k, [2](#)