

# The bodymapRat data user's guide

Stephanie C. Hicks shicks@jimmy.harvard.edu

Kwame Okrah kwame.okrah@gmail.com

Modified: July 23, 2015. Compiled: July 23, 2015

## 1 Overview

---

The bodymapRat package contains gene expression data on 652 RNA-Seq samples from a comprehensive rat transcriptomic BodyMap study. These samples include the sequence identifier information provided in the header of the FASTQ files which can be used as a surrogate for batch. These samples have not been normalized or pre-processed.

The data are an ExpressionSet. The phenotypic information can be extracted using the pData() function and a description is listed in Table 1.

Table 1: Description of phenotypic data in the bodymapRat ExpressionSet.

Title	Description
sraExperiment	SRA Experiment ID
title	Title of sample provided by the authors
geoAccession	GEO Accession ID
BioSample	BioSample ID
avgLength	Average read length
instrument	Machine identifier (from FASTQ header)
runID	Run ID (from FASTQ header)
fcID	Flow cell ID (from FASTQ header)
fcLane	Flow cell lane (from FASTQ header)
tile	Tile (from FASTQ header)
xtile	xtile (from FASTQ header)
ytile	ytile (from FASTQ header)
organ	Body organ
sex	Gender
stage	Stage
techRep	Technical replicate number
colOrgan	Column of colors to help with plotting
rnaRIN	RIN number
barcode	barcode number

The data can be accessed as follows:

```
library(Biobase)
library(useful)
library(bodymapRat)
```

```
data(bodymapRat)
```

```
# Get the expression data
edata = exprs(bodymapRat)
corner(edata)
```

```
##                SRR1169893 SRR1169894 SRR1169895 SRR1169896 SRR1169897
## ENSRNOG000000000001      1          0          0          1          4
## ENSRNOG000000000007      1          1          0          3          0
## ENSRNOG000000000008      7          4          2          3          7
## ENSRNOG000000000009      0          0          0          0          1
## ENSRNOG000000000010      0          1          0          0          0
```

```
# Get the pheno data
pdata = pData(bodymapRat)
head(pdata)
```

```
##          sraExperiment      sraRun      title geoAccession sraSample      BioSample
## SRR1169893      SRX471368 SRR1169893 Adr_F_002_1      GSM1328469 SRS558114 SAMN02642886
## SRR1169894      SRX471368 SRR1169894 Adr_F_002_1      GSM1328469 SRS558114 SAMN02642886
## SRR1169895      SRX471369 SRR1169895 Adr_F_002_2      GSM1328470 SRS558115 SAMN02642867
## SRR1169896      SRX471369 SRR1169896 Adr_F_002_2      GSM1328470 SRS558115 SAMN02642867
## SRR1169897      SRX471370 SRR1169897 Adr_F_002_3      GSM1328471 SRS558116 SAMN02642894
## SRR1169898      SRX471370 SRR1169898 Adr_F_002_3      GSM1328471 SRS558116 SAMN02642894
##          avgLength      organ sex stage techRep colOrgan mix rnaRIN barcode instrument
## SRR1169893          50 Adrenal  F      2          1      brown  M1      9.3          11      HWI-ST845
## SRR1169894          50 Adrenal  F      2          2      brown  M1      9.3          11      HWI-ST845
## SRR1169895          50 Adrenal  F      2          1      brown  M1      9.1           5      HWI-ST845
## SRR1169896          50 Adrenal  F      2          2      brown  M1      9.1           5      HWI-ST845
## SRR1169897          50 Adrenal  F      2          1      brown  M1      9.5           3      HWI-ST1131
## SRR1169898          50 Adrenal  F      2          2      brown  M1      9.5           3      HWI-ST1195
##          runID      fcID fcLane tile xtile ytile
## SRR1169893 120326 DOVTJACXX      2 1101 1506 2000
## SRR1169894 120525 D10G7ACXX      2 1101 1394 2133
## SRR1169895 120326 DOVTJACXX      5 1101 1170 2029
## SRR1169896 120525 D10G7ACXX      5 1101 1650 2126
## SRR1169897 120424 COP4UACXX      4 1101 1675 2216
## SRR1169898 120525 COTDUACXX      4 1101 1138 2067
```

The data in this package are used as an example data set in the `qsmooth` package.

## 2 References

1. Yu et al. (2013). A rat RNA-Seq transcriptomic BodyMap across 11 organs and 4 developmental stages. *Nature Communications* **5**:3230. PMID: 24510058. PMCID: PMC3926002.

## 3 SessionInfo

```
sessionInfo()

## R version 3.2.0 (2015-04-16)
```

```
## Platform: x86_64-apple-darwin13.4.0 (64-bit)
## Running under: OS X 10.10.4 (Yosemite)
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] parallel stats graphics grDevices datasets utils methods base
##
## other attached packages:
## [1] bodymapRat_0.0.1 useful_1.1.8 ggplot2_1.0.1 Biobase_2.29.1
## [5] BiocGenerics_0.15.2 knitr_1.10.5
##
## loaded via a namespace (and not attached):
## [1] Rcpp_0.11.6 digest_0.6.8 MASS_7.3-40 grid_3.2.0 plyr_1.8.3
## [6] gtable_0.1.2 formatR_1.2 magrittr_1.5 scales_0.2.5 evaluate_0.7
## [11] highr_0.5 stringi_0.5-4 reshape2_1.4.1 proto_0.3-10 BiocStyle_1.7.4
## [16] tools_3.2.0 stringr_1.0.0 munsell_0.4.2 colorspace_1.2-6
```