# Package 'curatedBladderData'

July 11, 2024

Type Package

**Version** 1.41.0

Title Bladder Cancer Gene Expression Analysis

| <b>Date</b> 2019-03-28   |
|--|
| Author Markus Riester  |
| Maintainer Markus Riester <markus@jimmy.harvard.edu></markus@jimmy.harvard.edu>  |
| <b>Description</b> The curatedBladderData package provides relevant functions and data for gene expression analysis in patients with bladder cancer.   |
| <b>Depends</b> R (>= 2.10.0), affy   |
| Suggests BiocStyle, survival, xtable, sva, genefilter, logging   |
| License Artistic-2.0   |
| <pre>URL https://github.com/lima1/curatedBladderData</pre>   |
| biocViews ExperimentData, CancerData, OvarianCancerData, MicroarrayData, ExpressionData  |
| git_url https://git.bioconductor.org/packages/curatedBladderData   |
| git_branch devel   |
| git_last_commit 1a7442d  |
| git_last_commit_date 2024-04-30  |
| Repository Bioconductor 3.20   |
| Date/Publication 2024-07-11  |
|  |
| Contents   |
| curatedBladderData-package       2         GSE13507_eset       3         GSE1827_eset       5         GSE19915.GPL3883_eset       8         GSE19915.GPL5186_eset       11         GSE31189_eset       13         GSE31684_eset       15 |
|  |

**34** 

Index

| GSE37317_eset             | GSE32894_eset         |       | <br> |  |
|---------------------------|-----------------------|-------|------|--|
| GSE89_eset                | GSE37317_eset         |       | <br> |  |
| <del>-</del>              | GSE5287_eset          |       | <br> |  |
| PMID17099711.GPL8300_eset | GSE89_eset            |       | <br> |  |
|                           | PMID17099711.GPL8300_ | _eset | <br> |  |

curatedBladderData-package

Clinically Annotated Data for the Bladder Cancer Transcriptome

# Description

The curatedBladderData package provides manually curated clinical data, uniformly processed expression data, and convenience functions for gene expression analysis in patients with ovarian cancer.

#### **Details**

Package: curatedBladderData

Type: Package
Version: 1.41.0
Date: 2019-03-28
License: Artistic-2.0
Depends: R (>= 2.10.0), affy

#### Author(s)

Benjamin F. Ganzfried, Steve Skates, Markus Riester, Victoria Wang, Thomas Risch, Benjamin Haibe-Kains, Curtis Huttenhower, Svitlana Tyekucheva, Jie Ding, Ina Jazic, Michael Birrer, Giovanni Parmigiani, Levi Waldron

Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Harvard School of Public Health

Maintainer: Levi Waldron <levi@jimmy.harvard.edu>

# **Examples**

```
##List all datasets:
data(package="curatedBladderData")
##
```

GSE13507\_eset 3

| GSE13507_eset | Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer. |
|---------------|--|

### **Description**

While several molecular markers of bladder cancer prognosis have been identified, the limited value of current prognostic markers has created the need for new molecular indicators of bladder cancer outcomes. The aim of this study was to identify genetic signatures associated with disease prognosis in bladder cancer. We used 272 primary bladder cancer specimens for microarray analysis and real-time reverse transcriptase polymerase chain reaction (RT-PCR) analysis. Microarray gene expression analysis of randomly selected 165 primary bladder cancer specimens as an original cohort was carried out. Risk scores were applied to stratify prognosis-related gene classifiers. Prognosisrelated gene classifiers were individually analyzed with tumor invasiveness (non-muscle invasive bladder cancer [NMIBC] and muscle invasive bladder cancer [MIBC]) and prognosis. We validated selected gene classifiers using RT-PCR in the original (165) and independent (107) cohorts. Ninetyseven genes related to disease progression among NMIBC patients were identified by microarray data analysis. Eight genes, a progression-related gene classifier in NMIBC, were selected for RT-PCR. The progression-related gene classifier in patients with NMIBC was closely correlated with progression in both original and independent cohorts. Furthermore, no patient with NMIBC in the good-prognosis signature group experienced cancer progression. We identified progression-related gene classifier that has strong predictive value for determining disease outcome in NMIBC. This gene classifier could assist in selecting NMIBC patients who might benefit from more aggressive therapeutic intervention or surveillance.

#### Usage

```
data( GSE13507_eset )
```

```
experimentData(eset):

Experiment data

Experimenter name: Kim WJ, Kim EJ, Kim SK, Kim YJ et al. Predictive value of progression-related gene of Laboratory: Kim, Bae 2010

Contact information:

Title: Predictive value of progression-related gene classifier in primary non-muscle invasive bladder URL:

PMIDs: 20059769

Abstract: A 223 word abstract is available. Use 'abstract' method.

Information is available on: preprocessing notes:

platform_title:

Illumina human-6 v2.0 expression beadchip platform_shorttitle:

Illumina human-6 v2.0
```

4 GSE13507\_eset

```
platform_summary:
         illuminaHumanv2
      platform_manufacturer:
         Illumina
      platform_distribution:
         commercial
      platform_accession:
         GPL6102
      platform_technology:
         oligonucleotide beads
   Preprocessing: default
   featureData(eset):
   An object of class 'AnnotatedDataFrame'
     featureNames: A1BG A1CF ... ZZZ3 (19329 total)
     varLabels: probeset gene
     varMetadata: labelDescription
Details
   assayData: 19329 features, 255 samples
   Platform type: illuminaHumanv2
   Overall survival time-to-event summary (in years):
   Call: survfit(formula = Surv(time, cens) ~ -1)
      90 observations deleted due to missingness
            n.max n.start events median 0.95LCL 0.95UCL
   records
    165.00 165.00 165.00 69.00
                                     7.26
                                             5.53
   Available sample meta-data:
    -----
   alt_sample_name:
      Length
                           Mode
                Class
         255 character character
   sample_type:
   adjacentnormal
                        healthy
                                    metastatic
                                                        tumor
               58
                             10
                                            22
                                                          165
   summarystage:
      invasive superficial
                                 NA's
            62
                      103
                                   90
   T:
      Min. 1st Qu. Median
                             Mean 3rd Qu.
                                             Max.
                                                     NA's
     0.000 1.000 1.000
                            1.473 2.000
                                            4.000
                                                       90
```

GSE1827\_eset 5

```
N:
   0
        1
              2
                   3 NA's
133
        8
                   1 109
Μ:
   0
        1 NA's
157
            90
age:
   Min. 1st Qu.
                  Median
                             Mean 3rd Qu.
                                              Max.
                                                       NA's
  24.00
          59.00
                   66.00
                            65.18
                                    73.00
                                             88.00
                                                         90
gender:
   f
        m NA's
  30
     135
            90
adjuvant_chemo:
        y NA's
       27
138
days_to_death:
   Min. 1st Qu.
                                                       NA's
                             Mean 3rd Qu.
                  Median
                                              Max.
     31
            521
                    1113
                             1473
                                     2258
                                              4169
                                                         90
vital_status:
                       NA's
deceased
           living
      69
                96
                         90
dfs_event:
doc dod
           ned NA's
  37
       32
            96
                  90
uncurated_author_metadata:
               Class
   Length
                           Mode
      255 character character
```

GSE1827\_eset

Bladder cancer outcome and subtype classification by gene expression.

# Description

Models of bladder tumor progression have suggested that genetic alterations may determine both phenotype and clinical course. We have applied expression microarray analysis to a divergent set of bladder tumors to further elucidate the course of disease progression and to classify tumors into

GSE1827\_eset

more homogeneous and clinically relevant subgroups. cDNA microarrays containing 10,368 human gene elements were used to characterize the global gene expression patterns in 80 bladder tumors, 9 bladder cancer cell lines, and 3 normal bladder samples. Robust statistical approaches accounting for the multiple testing problem were used to identify differentially expressed genes. Unsupervised hierarchical clustering successfully separated the samples into two subgroups containing superficial (pT(a) and pT(1)) versus muscle-invasive (pT(2)-pT(4)) tumors. Supervised classification had a 90.5% success rate separating superficial from muscle-invasive tumors based on a limited subset of genes. Tumors could also be classified into transitional versus squamous subtypes (89% success rate) and good versus bad prognosis (78% success rate). The performance of our stage classifiers was confirmed in silico using data from an independent tumor set. Validation of differential expression was done using immunohistochemistry on tissue microarrays for cathepsin E, cyclin A2, and parathyroid hormone-related protein. Genes driving the separation between tumor subsets may prove to be important biomarkers for bladder cancer development and progression and eventually candidates for therapeutic targeting.

### Usage

```
data( GSE1827_eset )
```

#### **Format**

```
experimentData(eset):
Experiment data
 Experimenter name: Blaveri E, Simko JP, Korkola JE, Brewer JL, Baehner F, Mehta K, Devries S, Koppie T,
 Laboratory: Blaveri, Waldman 2005
 Contact information:
  Title: Bladder cancer outcome and subtype classification by gene expression.
 URL:
 PMIDs: 15930339
  Abstract: A 216 word abstract is available. Use 'abstract' method.
  Information is available on: preprocessing
  notes:
   platform_title:
      JAKE
   platform_shorttitle:
      JAKE
   platform_summary:
      JAKE
   platform_manufacturer:
      other
   platform_distribution:
      non-commercial
   platform_accession:
      GPL1479
   platform_technology:
      spotted DNA/cDNA
```

Preprocessing: default

GSE1827\_eset 7

```
featureData(eset):
   An object of class 'AnnotatedDataFrame'
     featureNames: A2M AADAC ... ZZEF1 (6225 total)
     varLabels: probeset gene
     varMetadata: labelDescription
Details
   assayData: 6225 features, 80 samples
   Platform type: JAKE
   Overall survival time-to-event summary (in years):
   Call: survfit(formula = Surv(time, cens) ~ -1)
            n.max n.start events median 0.95LCL 0.95UCL
   records
    80.000 80.000 80.000 44.000 2.301 0.978
   _____
   Available sample meta-data:
   -----
   alt_sample_name:
      Min. 1st Qu. Median Mean 3rd Qu.
                                           Max.
      44.0 360.0 452.5 425.1 513.2 591.0
   sample_type:
   tumor
      80
   surgery_type:
      rc turbt
      50
           30
   histological_type:
   squamous
                tcc
                 74
         6
   summarygrade:
   high low
     67
        13
   summarystage:
      invasive superficial
           53
                      27
   T:
    0 1 2 3 4
```

17 10 14 26 13

```
N:
   0
             2 NA's
        1
  29
                  34
M:
   0
        1 NA's
   3
        2
            75
age:
   Min. 1st Qu.
                 Median
                            Mean 3rd Qu.
                                             Max.
                                                      NA's
  28.00
          57.25
                   66.00
                           65.56
                                   73.00 113.00
                                                         2
gender:
f m
24 56
recurrence_status:
norecurrence
               recurrence
                                    NA's
          49
                        24
                                       7
days_to_death:
   Min. 1st Qu.
                 Median
                            Mean 3rd Qu.
                                              Max.
          217.5
                  386.0
                           842.2 1280.0
                                           4348.0
    4.0
vital_status:
deceased
           living
      44
                36
uncurated_author_metadata:
              Class
   Length
       80 character character
```

GSE19915.GPL3883\_eset Combined gene expression and genomic profiling define two intrinsic molecular subtypes of urothelial carcinoma and gene signatures for molecular grading and outcome.

# Description

In the present investigation, we sought to refine the classification of urothelial carcinoma by combining information on gene expression, genomic, and gene mutation levels. For these purposes, we performed gene expression analysis of 144 carcinomas, and whole genome array-CGH analysis and mutation analyses of FGFR3, PIK3CA, KRAS, HRAS, NRAS, TP53, CDKN2A, and TSC1 in 103 of these cases. Hierarchical cluster analysis identified two intrinsic molecular subtypes, MS1 and MS2, which were validated and defined by the same set of genes in three independent bladder cancer data sets. The two subtypes differed with respect to gene expression and mutation

profiles, as well as with the level of genomic instability. The data show that genomic instability was the most distinguishing genomic feature of MS2 tumors, and that this trait was not dependent on TP53/MDM2 alterations. By combining molecular and pathologic data, it was possible to distinguish two molecular subtypes of T(a) and T(1) tumors, respectively. In addition, we define gene signatures validated in two independent data sets that classify urothelial carcinoma into low-grade (G(1)/G(2)) and high-grade (G(3)) tumors as well as non-muscle and muscle-invasive tumors with high precisions and sensitivities, suggesting molecular grading as a relevant complement to standard pathologic grading. We also present a gene expression signature with independent prognostic effect on metastasis and disease-specific survival. We conclude that the combination of molecular and histopathologic classification systems might provide a strong improvement for bladder cancer classification and produce new insights into the development of this tumor type.(c)2010 AACR.

#### **Usage**

```
data( GSE19915.GPL3883_eset )
```

```
experimentData(eset):
Experiment data
 Experimenter name: Lindgren D, Frigyesi A, Gudjonsson S, Sj?dahl G et al. Combined gene expression and
 Laboratory: Lindgren, H?glund 2010
 Contact information:
 Title: Combined gene expression and genomic profiling define two intrinsic molecular subtypes of uroth
 URL:
 PMIDs: 20406976
 Abstract: A 247 word abstract is available. Use 'abstract' method.
  Information is available on: preprocessing
  notes:
   platform_title:
      Swegene Human 27K RAP UniGene188 array
   platform_shorttitle:
      Swegene Human 27K
   platform_summary:
   platform_manufacturer:
      other
   platform_distribution:
      non-commercial
   platform_accession:
      GPL3883
   platform_technology:
      spotted DNA/cDNA
Preprocessing: default
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: 13CDNA73 15E1.2 ... raptor | MGC14560 (10585 total)
```

varLabels: probeset gene
varMetadata: labelDescription

#### **Details**

```
assayData: 10585 features, 84 samples
Platform type:
Overall survival time-to-event summary (in years):
Call: survfit(formula = Surv(time, cens) ~ -1)
  10 observations deleted due to missingness
         n.max n.start events median 0.95LCL 0.95UCL
records
    74
           74
                  74
                         4
                                NA
                                        NA
-----
Available sample meta-data:
-----
alt_sample_name:
  Length
           Class
                      Mode
      84 character character
sample_type:
healthy tumor
     8
           76
summarystage:
  invasive superficial
                            NA's
         2
              73
                              9
T:
      1
           3 NA's
 56
     17
           2 9
G:
       2
           3 NA's
  1
      38
         13
 24
days_to_death:
  Min. 1st Qu. Median
                        Mean 3rd Qu.
                                      Max.
                                              NA's
     0
        1231 1720
                        1552 1937
                                       2429
                                               10
vital_status:
deceased
         living
                   NA's
      4
                     10
uncurated_author_metadata:
            Class
  Length
                      Mode
```

84 character character

GSE19915.GPL5186\_eset Combined gene expression and genomic profiling define two intrinsic molecular subtypes of urothelial carcinoma and gene signatures for molecular grading and outcome.

#### **Description**

In the present investigation, we sought to refine the classification of urothelial carcinoma by combining information on gene expression, genomic, and gene mutation levels. For these purposes, we performed gene expression analysis of 144 carcinomas, and whole genome array-CGH analysis and mutation analyses of FGFR3, PIK3CA, KRAS, HRAS, NRAS, TP53, CDKN2A, and TSC1 in 103 of these cases. Hierarchical cluster analysis identified two intrinsic molecular subtypes, MS1 and MS2, which were validated and defined by the same set of genes in three independent bladder cancer data sets. The two subtypes differed with respect to gene expression and mutation profiles, as well as with the level of genomic instability. The data show that genomic instability was the most distinguishing genomic feature of MS2 tumors, and that this trait was not dependent on TP53/MDM2 alterations. By combining molecular and pathologic data, it was possible to distinguish two molecular subtypes of T(a) and T(1) tumors, respectively. In addition, we define gene signatures validated in two independent data sets that classify urothelial carcinoma into low-grade (G(1)/G(2)) and high-grade (G(3)) tumors as well as non-muscle and muscle-invasive tumors with high precisions and sensitivities, suggesting molecular grading as a relevant complement to standard pathologic grading. We also present a gene expression signature with independent prognostic effect on metastasis and disease-specific survival. We conclude that the combination of molecular and histopathologic classification systems might provide a strong improvement for bladder cancer classification and produce new insights into the development of this tumor type.(c)2010 AACR.

### Usage

```
data( GSE19915.GPL5186_eset )
```

#### **Format**

notes:

```
experimentData(eset):
Experiment data
 Experimenter name: Lindgren D, Frigyesi A, Gudjonsson S, Sj?dahl G et al. Combined gene expression and
 Laboratory: Lindgren, H?glund 2010
 Contact information:
 Title: Combined gene expression and genomic profiling define two intrinsic molecular subtypes of uroth
 URL:
 PMIDs: 20406976
 Abstract: A 247 word abstract is available. Use 'abstract' method.
  Information is available on: preprocessing
```

```
platform_title:
         SWEGENE H_v3.0.1 35K
      platform_shorttitle:
         SWEGENE H_v3.0.1 35K
      platform_summary:
      platform_manufacturer:
         other
      platform_distribution:
         non-commercial
      platform_accession:
         GPL5186
      platform_technology:
         spotted oligonucleotide
   Preprocessing: default
   featureData(eset):
   An object of class 'AnnotatedDataFrame'
     featureNames: 15E1_HUMAN 38596 ... ZZZ3 (12391 total)
     varLabels: probeset gene
     varMetadata: labelDescription
Details
   assayData: 12391 features, 98 samples
   Platform type:
   Overall survival time-to-event summary (in years):
   Call: survfit(formula = Surv(time, cens) ~ -1)
      11 observations deleted due to missingness
   records n.max n.start events median 0.95LCL 0.95UCL
        87
               87
                      87
                           26
                                     NA
                                              NA
   -----
   Available sample meta-data:
   alt_sample_name:
      Length Class
                           Mode
          98 character character
   sample_type:
   healthy tumor
         7
               91
   summarystage:
                           NA's
      invasive superficial
           45 43
                                 10
```

GSE31189\_eset 13

```
T:
   Min. 1st Qu.
                  Median
                             Mean 3rd Qu.
                                                        NA's
                                               Max.
  0.000
          1.000
                   2.000
                            1.693
                                     3.000
                                              4.000
                                                          10
G:
   2
        3 NA's
  19
       71
              8
days_to_death:
   Min. 1st Qu.
                  Median
                             Mean 3rd Qu.
                                               Max.
                                                        NA's
                  1117.0
                          1067.0 1512.0
    3.0
          692.5
                                             2335.0
                                                          11
vital_status:
deceased
            living
                        NA's
      26
                61
                          11
uncurated_author_metadata:
   Length
               Class
                           Mode
       98 character character
```

GSE31189\_eset

A candidate molecular biomarker panel for the detection of bladder cancer.

# Description

Bladder cancer is among the five most common malignancies worldwide, and due to high rates of recurrence, one of the most prevalent. Improvements in noninvasive urine-based assays to detect bladder cancer would benefit both patients and health care systems. In this study, the goal was to identify urothelial cell transcriptomic signatures associated with bladder cancer. Gene expression profiling (Affymetrix U133 Plus 2.0 arrays) was applied to exfoliated urothelia obtained from a cohort of 92 subjects with known bladder disease status. Computational analyses identified candidate biomarkers of bladder cancer and an optimal predictive model was derived. Selected targets from the profiling analyses were monitored in an independent cohort of 81 subjects using quantitative real-time PCR (RT-PCR). Transcriptome profiling data analysis identified 52 genes associated with bladder cancer (P??? 0.001) and gene models that optimally predicted class label were derived. RT-PCR analysis of 48 selected targets in an independent cohort identified a 14-gene diagnostic signature that predicted the presence of bladder cancer with high accuracy. Exfoliated urothelia sampling provides a robust analyte for the evaluation of patients with suspected bladder cancer. The refinement and validation of the multigene urothelial cell signatures identified in this preliminary study may lead to accurate, noninvasive assays for the detection of bladder cancer. The development of an accurate, noninvasive bladder cancer detection assay would benefit both the patient and health care systems through better detection, monitoring, and control of disease.

14 *GSE31189\_eset* 

#### Usage

```
data( GSE31189_eset )
```

#### **Format**

```
experimentData(eset):
Experiment data
 Experimenter name: Urquidi V, Goodison S, Cai Y, Sun Y et al. A candidate molecular biomarker panel for
 Laboratory: Urquidi, Rosser 2012
  Contact information:
 Title: A candidate molecular biomarker panel for the detection of bladder cancer.
 PMIDs: 23097579
  Abstract: A 230 word abstract is available. Use 'abstract' method.
  Information is available on: preprocessing
  notes:
   platform_title:
      [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
   platform_shorttitle:
      Affymetrix HG-U133Plus2
   platform_summary:
      hgu133plus2
   platform_manufacturer:
      Affymetrix
   platform_distribution:
      commercial
   platform_accession:
      GPL570
   platform_technology:
      in situ oligonucleotide
Preprocessing: frma
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A1BG A1BG-AS1 ... ZZZ3 (19381 total)
  varLabels: probeset gene
  varMetadata: labelDescription
```

#### **Details**

assayData: 19381 features, 92 samples
Platform type: hgu133plus2
-----Available sample meta-data:

```
alt_sample_name:
   Length
              Class
                          Mode
       92 character character
sample_type:
healthy
          tumor
     40
             52
batch:
   Length
              Class
                          Mode
       92 character character
uncurated_author_metadata:
   Length
              Class
                          Mode
       92 character character
```

GSE31684\_eset

Combination of a novel gene expression signature with a clinical nomogram improves the prediction of survival in high-risk bladder cancer.

# Description

We aimed to validate and improve prognostic signatures for high-risk urothelial carcinoma of the bladder.We evaluated microarray data from 93 patients with bladder cancer managed by radical cystectomy to determine gene expression patterns associated with clinical and prognostic variables. We compared our results with published bladder cancer microarray data sets comprising 578 additional patients and with 49 published gene signatures from multiple cancer types. Hierarchical clustering was utilized to identify subtypes associated with differences in survival. We then investigated whether the addition of survival-associated gene expression information to a validated postcystectomy nomogram utilizing clinical and pathologic variables improves prediction of recurrence. Multiple markers for muscle invasive disease with highly significant expression differences in multiple data sets were identified, such as fibronectin 1 (FN1), NNMT, POSTN, and SMAD6. We identified signatures associated with pathologic stage and the likelihood of developing metastasis and death from bladder cancer, as well as with two distinct clustering subtypes of bladder cancer. Our novel signature correlated with overall survival in multiple independent data sets, significantly improving the prediction concordance of standard staging in all data sets [mean ??C-statistic: 0.14; 95% confidence interval (CI), 0.01-0.27; P < 0.001]. Tested in our patient cohort, it significantly enhanced the performance of a postoperative survival nomogram (??C-statistic: 0.08, 95% CI, -0.04-0.20; P < 0.005). Prognostic information obtained from gene expression data can aid in posttreatment prediction of bladder cancer recurrence. Our findings require further validation in external cohorts and prospectively in a clinical trial setting.

### Usage

```
data( GSE31684_eset )
```

#### **Format**

```
experimentData(eset):
Experiment data
 Experimenter name: Riester M, Taylor JM, Feifer A, Koppie T et al. Combination of a novel gene expression
 Laboratory: Riester, Michor 2012
 Contact information:
 Title: Combination of a novel gene expression signature with a clinical nomogram improves the prediction
 URL:
 PMIDs: 22228636
 Abstract: A 243 word abstract is available. Use 'abstract' method.
  Information is available on: preprocessing
  notes:
   platform_title:
      [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
   platform_shorttitle:
      Affymetrix HG-U133Plus2
   platform_summary:
      hgu133plus2
   platform_manufacturer:
      Affymetrix
   platform_distribution:
      commercial
   platform_accession:
      GPL570
   platform_technology:
      in situ oligonucleotide
Preprocessing: frma
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A1BG A1BG-AS1 ... ZZZ3 (19381 total)
  varLabels: probeset gene
  varMetadata: labelDescription
```

# **Details**

```
assayData: 19381 features, 93 samples
Platform type: hgu133plus2
Overall survival time-to-event summary (in years):
Call: survfit(formula = Surv(time, cens) ~ -1)

records n.max n.start events median 0.95LCL 0.95UCL 93.00 93.00 93.00 65.00 2.74 1.37 7.52

Available sample meta-data:
```

```
-----
alt_sample_name:
  Length Class Mode
     93 character character
sample_type:
tumor
  93
surgery_type:
rc
93
histological_type:
               tcc
ge
    cis squamous
     2 5
summarygrade:
high low
 87 6
summarystage:
  invasive superficial
      78 15
T:
0 1 2 3 4
5 10 17 42 19
N:
  0 1 NA's
 49 28 16
M:
0 1
57 36
age:
  Min. 1st Qu. Median Mean 3rd Qu.
                                   Max.
 42.00 62.00 69.00 69.11 75.00 91.00
gender:
f m
25 68
neoadjuvant_chemo:
```

n y

90 3

adjuvant\_chemo:

n y 58 35

days\_to\_tumor\_recurrence:

Min. 1st Qu. Median Mean 3rd Qu. Max. 12 170 495 1307 2574 5342

recurrence\_status:

norecurrence recurrence 54 39

days\_to\_death:

Min. 1st Qu. Median Mean 3rd Qu. Max. 12 299 953 1445 2616 5342

vital\_status:

deceased living 65 28

dfs\_event:
doc dod ned

27 38 28

smoking\_status:

current former never 19 56 18

smoking\_package\_years:

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's 5.00 22.50 40.00 44.13 60.00 120.00 22

nomogram\_score:

Min. 1st Qu. Median Mean 3rd Qu. Max. 1.00 21.11 46.60 44.99 67.29 92.15

batch:

Length Class Mode 93 character character

uncurated\_author\_metadata:

Length Class Mode 93 character character *GSE32894\_eset* 19

GSE32894\_eset

A molecular taxonomy for urothelial carcinoma.

#### **Description**

Even though urothelial cancer is the fourth most common tumor type among males, progress in treatment has been scarce. A problem in day-to-day clinical practice is that precise assessment of individual tumors is still fairly uncertain; consequently efforts have been undertaken to complement tumor evaluation with molecular biomarkers. An extension of this approach would be to base tumor classification primarily on molecular features. Here, we present a molecular taxonomy for urothelial carcinoma based on integrated genomics. We use gene expression profiles from 308 tumor cases to define five major urothelial carcinoma subtypes: urobasal A, genomically unstable, urobasal B, squamous cell carcinoma like, and an infiltrated class of tumors. Tumor subtypes were validated in three independent publically available data sets. The expression of 11 key genes was validated at the protein level by immunohistochemistry. The subtypes show distinct clinical outcomes and differ with respect to expression of cell-cycle genes, receptor tyrosine kinases particularly FGFR3, ERBB2, and EGFR, cytokeratins, and cell adhesion genes, as well as with respect to FGFR3, PIK3CA, and TP53 mutation frequency. The molecular subtypes cut across pathologic classification, and classdefining gene signatures show coordinated expression irrespective of pathologic stage and grade, suggesting the molecular phenotypes as intrinsic properties of the tumors. Available data indicate that susceptibility to specific drugs is more likely to be associated with the molecular stratification than with pathologic classification. We anticipate that the molecular taxonomy will be useful in future clinical investigations.??2012 AACR.

#### **Usage**

```
data( GSE32894_eset )
```

experimentData(eset):

```
Experiment data

Experimenter name: Sj?dahl G, Lauss M, L?vgren K, Chebil G et al. A molecular taxonomy for urothelial carbonatory: Sj?dahl, H?glund 2012

Contact information:

Title: A molecular taxonomy for urothelial carcinoma.

URL:

PMIDs: 22553347

Abstract: A 236 word abstract is available. Use 'abstract' method.

Information is available on: preprocessing notes:

platform_title:

Illumina HumanHT-12 V3.0 expression beadchip platform_shorttitle:

Illumina HumanHT-12 V3.0 platform_summary:
```

20 GSE32894\_eset

```
illuminaHumanv3
      platform_manufacturer:
         Illumina
      platform_distribution:
         commercial
      platform_accession:
         GPL6947
      platform_technology:
         oligonucleotide beads
   Preprocessing: default
   featureData(eset):
   An object of class 'AnnotatedDataFrame'
     featureNames: A1CF A2M ... ZZZ3 (15638 total)
     varLabels: probeset gene
     varMetadata: labelDescription
Details
   assayData: 15638 features, 308 samples
   Platform type: illuminaHumanv3
   Overall survival time-to-event summary (in years):
   Call: survfit(formula = Surv(time, cens) ~ -1)
      84 observations deleted due to missingness
             n.max n.start events median 0.95LCL 0.95UCL
       224
               224
                       224
                                25
                                        NA
                                                NA
                                                        NA
   Available sample meta-data:
   -----
   alt_sample_name:
      Length
                 Class
                            Mode
         308 character character
   sample_type:
   tumor
     308
   summarystage:
      invasive superficial
                                  NA's
            93
                      213
   T:
      Min. 1st Qu. Median
                              Mean 3rd Qu.
                                                      NA's
                                              Max.
    0.0000 0.0000 1.0000 0.9542 2.0000 4.0000
```

GSE37317\_eset 21

```
G:
        2
              3
                   4 NA's
   1
  48
      103
           154
N:
   0
        1
              2
                   3 NA's
  48
            10
                      245
age:
   Min. 1st Qu.
                  Median
                             Mean 3rd Qu.
                                              Max.
  20.00
          62.75
                   71.00
                            70.61
                                    79.00
                                             96.00
gender:
  f
80 228
days_to_death:
   Min. 1st Qu.
                  Median
                            Mean 3rd Qu.
                                              Max.
                                                       NA's
          552.5 1068.0
                         1214.0 1766.0
                                           3357.0
                                                         84
vital_status:
deceased
           living
                       NA's
      25
               199
                         84
dfs_event:
dod NA's
  25 283
uncurated_author_metadata:
              Class
   Length
      308 character character
```

GSE37317\_eset

Transcriptional signatures of Ral GTPase are associated with aggressive clinicopathologic characteristics in human cancer.

# Description

RalA and RalB are small GTPases that support malignant development and progression in experimental models of bladder, prostate, and squamous cancer. However, demonstration of their clinical relevance in human tumors remains lacking. Here, we developed tools to evaluate Ral protein expression, activation, and transcriptional output and evaluated their association with clinicopathologic parameters in common human tumor types. To evaluate the relevance of Ral activation and transcriptional output, we correlated RalA and RalB activation with the mutational status of key human bladder cancer genes. We also identified and evaluated a transcriptional signature of genes that correlates with depletion of RalA and RalB in vivo. The Ral transcriptional signature score, but

22 *GSE37317\_eset* 

not protein expression as evaluated by immunohistochemistry, predicted disease stage, progression to muscle invasion, and survival in human bladder cancers and metastatic and stem cell phenotypes in bladder cancer models. In prostate cancer, the Ral transcriptional signature score was associated with seminal vesicle invasion, androgen-independent progression, and reduced survival. In squamous cell carcinoma, this score was decreased in cancer tissues compared with normal mucosa, validating the experimental findings that Ral acts as a tumor suppressor in this tumor type. Together, our findings show the clinical relevance of Ral in human cancer and provide a rationale for the development of Ral-directed therapies.

#### Usage

```
data( GSE37317_eset )
```

```
experimentData(eset):
Experiment data
 Experimenter name: Smith SC, Baras AS, Owens CR, Dancik G et al. Transcriptional signatures of Ral GTPa
 Laboratory: Smith, Theodorescu 2012
 Contact information:
 Title: Transcriptional signatures of Ral GTPase are associated with aggressive clinicopathologic chara
 URL:
 PMIDs: 22586063
  Abstract: A 210 word abstract is available. Use 'abstract' method.
  Information is available on: preprocessing
  platform_title:
      [HG-U133A] Affymetrix Human Genome U133A Array
   platform_shorttitle:
      Affymetrix HG-U133A
   platform_summary:
      hgu133a
   platform_manufacturer:
      Affymetrix
   platform_distribution:
      commercial
   platform_accession:
      GPL96
   platform_technology:
      in situ oligonucleotide
Preprocessing: frma
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A1CF A2M ... ZZZ3 (13013 total)
  varLabels: probeset gene
  varMetadata: labelDescription
```

GSE5287\_eset 23

#### **Details**

```
assayData: 13013 features, 19 samples
Platform type: hgu133a
Available sample meta-data:
alt_sample_name:
  Length
              Class
                          Mode
       19 character character
sample_type:
tumor
   19
histological_type:
squamous
              tcc
       1
               18
summarystage:
   invasive superficial
         11
                       8
T:
0 1 2 3 4
4 4 4 3 4
batch:
                          Mode
  Length
              Class
       19 character character
uncurated_author_metadata:
  Length
              Class
       19 character character
```

GSE5287\_eset

Emmprin and survivin predict response and survival following cisplatin-containing chemotherapy in patients with advanced bladder cancer.

# **Description**

Cisplatin-containing chemotherapy is the standard of care for patients with locally advanced and metastatic transitional cell carcinoma of the urothelium. The response rate is approximately 50% and tumor-derived molecular prognostic markers are desirable for improved estimation of response

24 *GSE5287\_eset* 

and survival. Affymetrix GeneChip expression profiling was carried out using tumor material from 30 patients. A set of genes with an expression highly correlated to survival time after chemotherapy was identified. Two genes were selected for validation by immunohistochemistry in an independent material of 124 patients receiving cisplatin-containing therapy. Fifty-five differentially expressed genes correlated significantly to survival time. Two of the protein products (emmprin and survivin) were validated using immunohistochemistry. Multivariate analysis identified emmprin expression (hazard ratio, 2.23; P < 0.0001) and survivin expression (hazard ratio, 2.46; P < 0.0001) as independent prognostic markers for poor outcome, together with the presence of visceral metastases (hazard ratio, 2.62; P < 0.0001). In the clinical good prognostic group of patients without visceral metastases, both markers showed significant discriminating power as supplemental risk factors (P < 0.0001). Within this group of patients, the subgroups of patients with no positive, one positive, or two positive immunohistochemistry scores (emmprin and survivin) had estimated 5-year survival rates of 44.0%, 21.1%, and 0%, respectively. Response to chemotherapy could also be predicted with an odds ratio of 4.41 (95% confidence interval, 1.91-10.1) and 2.48 (95% confidence interval, 1.1-5.5) for emmprin and survivin, respectively. Emmprin and survivin proteins were identified as strong independent prognostic factors for response and survival after cisplatin-containing chemotherapy in patients with advanced bladder cancer.

# Usage

```
data( GSE5287_eset )
```

```
experimentData(eset):
Experiment data
 Experimenter name: Als AB, Dyrskj?t L, von der Maase H, Koed K et al. Emmprin and survivin predict response
 Laboratory: Als, Orntoft 2007
 Contact information:
 Title: Emmprin and survivin predict response and survival following cisplatin-containing chemotherapy
 URL:
  PMIDs: 17671123
  Abstract: A 254 word abstract is available. Use 'abstract' method.
  Information is available on: preprocessing
  notes:
   platform_title:
      [HG-U133A] Affymetrix Human Genome U133A Array
   platform_shorttitle:
      Affymetrix HG-U133A
   platform_summary:
      hgu133a
   platform_manufacturer:
      Affymetrix
   platform_distribution:
      commercial
   platform_accession:
      GPL96
   platform_technology:
```

GSE5287\_eset 25

```
in situ oligonucleotide
   Preprocessing: frma
   featureData(eset):
   An object of class 'AnnotatedDataFrame'
     featureNames: A1CF A2M ... ZZZ3 (13013 total)
     varLabels: probeset gene
     varMetadata: labelDescription
Details
   assayData: 13013 features, 30 samples
   Platform type: hgu133a
   Overall survival time-to-event summary (in years):
   Call: survfit(formula = Surv(time, cens) ~ -1)
   records n.max n.start events median 0.95LCL 0.95UCL
     30.00 30.00 30.00 25.00 4.36 3.12 7.81
   _____
   Available sample meta-data:
   _____
   alt_sample_name:
      Length Class
                          Mode
         30 character character
   sample_type:
   tumor
      30
   summarystage:
   invasive
        30
   T:
    4
   30
   neoadjuvant_chemo:
    n
   30
   adjuvant_chemo:
    У
```

30

adjuvant\_regimen:

26 GSE89\_eset

```
cisplatin
       30
days_to_death:
   Min. 1st Qu.
                  Median
                             Mean 3rd Qu.
                                              Max.
    420
           1080
                    1590
                             3160
                                     2962
                                             12600
vital_status:
deceased
           living
      25
batch:
   Length
               Class
                           Mode
       30 character character
uncurated_author_metadata:
   Length
               Class
                           Mode
       30 character character
```

GSE89\_eset

Identifying distinct classes of bladder carcinoma using microarrays.

# **Description**

Bladder cancer is a common malignant disease characterized by frequent recurrences. The stage of disease at diagnosis and the presence of surrounding carcinoma in situ are important in determining the disease course of an affected individual. Despite considerable effort, no accepted immunohistological or molecular markers have been identified to define clinically relevant subsets of bladder cancer. Here we report the identification of clinically relevant subclasses of bladder carcinoma using expression microarray analysis of 40 well characterized bladder tumors. Hierarchical cluster analysis identified three major stages, Ta, T1 and T2-4, with the Ta tumors further classified into subgroups. We built a 32-gene molecular classifier using a cross-validation approach that was able to classify benign and muscle-invasive tumors with close correlation to pathological staging in an independent test set of 68 tumors. The classifier provided new predictive information on disease progression in Ta tumors compared with conventional staging (P < 0.005). To delineate non-recurring Ta tumors from frequently recurring Ta tumors, we analyzed expression patterns in 31 tumors by applying a supervised learning classification methodology, which classified 75% of the samples correctly (P < 0.006). Furthermore, gene expression profiles characterizing each stage and subtype identified their biological properties, producing new potential targets for therapy.

#### Usage

```
data( GSE89_eset )
```

GSE89\_eset 27

#### **Format**

```
experimentData(eset):
Experiment data
 Experimenter name: Dyrskj?t L, Thykjaer T, Kruh?ffer M, Jensen JL et al. Identifying distinct classes of
 Laboratory: Dyrskjot, Orntoft 2003
  Contact information:
  Title: Identifying distinct classes of bladder carcinoma using microarrays.
  URL:
  PMIDs: 12469123
  Abstract: A 202 word abstract is available. Use 'abstract' method.
  Information is available on: preprocessing
  notes:
   platform_title:
      [Hu6800] Affymetrix Human Full Length HuGeneFL Array
   platform_shorttitle:
      Affymetrix HuGeneFL
   platform_summary:
      hu6800
   platform_manufacturer:
      Affymetrix
   platform_distribution:
      commercial
   platform_accession:
      GPL80
   platform_technology:
      in situ oligonucleotide
Preprocessing: rma
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A2M AADAC ... ZYX (5466 total)
  varLabels: probeset gene
  varMetadata: labelDescription
```

# **Details**

```
sample_type:
tumor
   40
summarystage:
   invasive superficial
          9
T:
0
   1
      2
20 11
G:
   2
        3
             4 NA's
       32
             1
                   1
uncurated_author_metadata:
   Length
              Class
       40 character character
```

PMID17099711.GPL8300\_eset

Regional copy number-independent deregulation of transcription in cancer.

#### Description

Genetic and epigenetic alterations have been identified that lead to transcriptional deregulation in cancers. Genetic mechanisms may affect single genes or regions containing several neighboring genes, as has been shown for DNA copy number changes. It was recently reported that epigenetic suppression of gene expression can also extend to a whole region; this is known as long-range epigenetic silencing. Various techniques are available for identifying regional genetic alterations, but no large-scale analysis has yet been carried out to obtain an overview of regional epigenetic alterations. We carried out an exhaustive search for regions susceptible to such mechanisms using a combination of transcriptome correlation map analysis and array CGH data for a series of bladder carcinomas. We validated one candidate region experimentally, demonstrating histone methylation leading to the loss of expression of neighboring genes without DNA methylation.

# Usage

```
data( PMID17099711.GPL8300_eset )
```

```
experimentData(eset):
Experiment data
```

```
Experimenter name: Stransky N, Vallot C, Reyal F, Bernard-Pierrot I, de Medina SG, Segraves R, de Rycke
 Laboratory: Stransky, Radvany 2006
 Contact information:
 Title: Regional copy number-independent deregulation of transcription in cancer.
 URL:
  PMIDs: 17099711
  Abstract: A 136 word abstract is available. Use 'abstract' method.
  Information is available on: preprocessing
  notes:
   platform_title:
      [HG_U95Av2] Affymetrix Human Genome U95 Version 2 Array
   platform_shorttitle:
      Affymetrix U95Av2
   platform_summary:
      hgu95av2
   platform_manufacturer:
      Affymetrix
   platform_distribution:
      commercial
   platform_accession:
      GPL8300
   platform_technology:
      NA
Preprocessing: rma
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: AADAC AAK1 ... ZZZ3 (8950 total)
  varLabels: probeset gene
  varMetadata: labelDescription
assayData: 8950 features, 30 samples
```

#### Details

```
summarystage:
   invasive superficial
         14
                     16
T:
0
   1 2
          3
10
   6
      1
substage:
   а
        b NA's
  13
        6
            11
G:
   2 3
1
8
   5 17
N:
   0
        1
             2 NA's
  18
        3
M:
   0 NA's
  26
gender:
f m
5 25
batch:
2002-01-23 2002-01-24 2002-09-20 2003-03-06
         6
                    3
                                           7
                               14
uncurated_author_metadata:
   Length
              Class
       30 character character
```

PMID17099711.GPL91\_eset

Regional copy number-independent deregulation of transcription in cancer.

# Description

Genetic and epigenetic alterations have been identified that lead to transcriptional deregulation in cancers. Genetic mechanisms may affect single genes or regions containing several neighboring genes, as has been shown for DNA copy number changes. It was recently reported that epigenetic

suppression of gene expression can also extend to a whole region; this is known as long-range epigenetic silencing. Various techniques are available for identifying regional genetic alterations, but no large-scale analysis has yet been carried out to obtain an overview of regional epigenetic alterations. We carried out an exhaustive search for regions susceptible to such mechanisms using a combination of transcriptome correlation map analysis and array CGH data for a series of bladder carcinomas. We validated one candidate region experimentally, demonstrating histone methylation leading to the loss of expression of neighboring genes without DNA methylation.

# Usage

```
data( PMID17099711.GPL91_eset )
```

```
experimentData(eset):
Experiment data
 Experimenter name: Stransky N, Vallot C, Reyal F, Bernard-Pierrot I, de Medina SG, Segraves R, de Rycke
 Laboratory: Stransky, Radvany 2006
 Contact information:
 Title: Regional copy number-independent deregulation of transcription in cancer.
 URL:
 PMIDs: 17099711
  Abstract: A 136 word abstract is available. Use 'abstract' method.
  Information is available on: preprocessing
  platform_title:
      [HG_U95A] Affymetrix Human Genome U95A Array
   platform_shorttitle:
      Affymetrix U95A
   platform_summary:
      hgu95a
   platform_manufacturer:
      Affymetrix
   platform_distribution:
      commercial
   platform_accession:
      GPL91
   platform_technology:
      NA
Preprocessing: rma
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: AADAC AAK1 ... ZZZ3 (8948 total)
  varLabels: probeset gene
  varMetadata: labelDescription
```

# **Details**

```
assayData: 8948 features, 31 samples
Platform type: hgu95a
Available sample meta-data:
unique_patient_ID:
  Length Class
     31 character character
sample_type:
healthy tumor
     5
           26
summarystage:
                     NA's
  invasive superficial
      17 9
                           5
T:
                                            NA's
  Min. 1st Qu. Median
                      Mean 3rd Qu.
                                     Max.
 0.000 1.000 2.000
                     2.077 3.000
                                    4.000
                                               5
substage:
  a b NA's
 11
      4 16
G:
      2 3 NA's
  1
     7 15 5
N:
  0
     1 2 NA's
 20
      1
           2 8
M:
  0
      1 NA's
 20
      2
gender:
f m
7 24
batch:
  Length
           Class
                     Mode
      31 character character
uncurated_author_metadata:
```

Length Class Mode 31 character character

# **Index**

```
* datasets
                                               GSE31684_eset, 15
    GSE13507_eset, 3
                                               GSE32894_eset, 19
    GSE1827_eset, 5
                                               GSE37317_eset, 21
    GSE19915.GPL3883_eset, 8
                                               GSE5287_eset, 23
                                               GSE89_eset, 26
    GSE19915.GPL5186_eset, 11
    GSE31189_eset, 13
                                               PMID17099711.GPL8300_eset, 28
    GSE31684_eset, 15
                                               PMID17099711.GPL91_eset, 30
    GSE32894_eset, 19
    GSE37317_eset, 21
    GSE5287_eset, 23
    GSE89_eset, 26
    PMID17099711.GPL8300_eset, 28
    PMID17099711.GPL91_eset, 30
 GSE13507_eset (GSE13507_eset), 3
 GSE1827_eset (GSE1827_eset), 5
 GSE19915.GPL3883_eset
        (GSE19915.GPL3883_eset), 8
 GSE19915.GPL5186_eset
        (GSE19915.GPL5186_eset), 11
 GSE31189_eset (GSE31189_eset), 13
 GSE31684_eset (GSE31684_eset), 15
 GSE32894_eset (GSE32894_eset), 19
 GSE37317_eset (GSE37317_eset), 21
 GSE5287_eset (GSE5287_eset), 23
 GSE89_eset (GSE89_eset), 26
 PMID17099711.GPL8300_eset
        (PMID17099711.GPL8300_eset), 28
 PMID17099711.GPL91_eset
        (PMID17099711.GPL91_eset), 30
curatedBladderData
        (curatedBladderData-package), 2
curatedBladderData-package, 2
GSE13507_eset, 3
GSE1827_eset, 5
GSE19915.GPL3883_eset, 8
GSE19915.GPL5186_eset, 11
GSE31189_eset, 13
```