Package 'curatedOvarianData'

November 12, 2012

Type Package

Title Clinically Annotated Data for the Ovarian Cancer Transcriptome

Version 0.99.10

Date 2012-11-06

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

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

Description The curatedOvarianData package provides data for gene expression analysis in patients with ovarian cancer.

Depends R (>= 2.10.0), affy

Imports BiocGenerics

Suggests survival, RUnit, metafor, genefilter, logging

License Artistic-2.0

URL http://bcb.dfci.harvard.edu/ovariancancer

R topics documented:

`curatedOvarianData-package`

Ovarian Cancer Gene Expression Analysis

Description

The curatedOvarianData package provides relevant functions and data for gene expression analysis in patients with ovarian cancer.

Details

Package:	curatedOvarianData
Type:	Package
Version:	0.99.10
Date:	2012-11-06
License:	Artistic-2.0
Depends:	R (>= 2.10.0), affy

In the "Available sample meta-data" sections of each dataset, please refer to the following key.

For "subtype": ser=serous;endo=endometrioid;endo_clearcell=mixture of ser+endo. Other includes sarcomatoid, endometroid, papillary.

For "primarysite" and for "arrayedsite": ov=ovary;ft=fallopian tube.

For "summarygrade": low = 1, 2, LMP. High= 3,4,23.

For "summarystage": early = 1,2, 12. late=3,4,23,34.

For "tumorstage": FIGO Stage (I-IV, but coded here as 1-4 to ensure correct ordering in factors). If multiple stages given (eg 34), use the highest.

For "T": TNM Stage (1-3). If multiple stages given (eg 34), use the highest.

For "substage": substage (abcd). For cases like ab, bc, use highest given.

For "G": Grade (1-4): If multiple given, ie 12, 23, use highest given.

For "N": N (0/1): degree of spread to regional lymph nodes.

For "M": M (0/1): presence of metastasis.

For "pltx": patient treated with platin.

For "tax": patient treated with taxol.

For "neo": patient treated with neoadjuvant treatment.

For "primary_therapy_outcome_success": response to any kind of therapy (including radiation only).

For "chemo_response": platinum resistance: refractory=3mo or less, resistant=6mo or less, sensitive=12mo or higher.

For "inferred_chemo_response": inferred platinum resistance: refractory=death in 6mo or less, sensitive=survival for 12mo or more.

For "debulking": amount of residual disease (optimal = <1mm, suboptimal=>1mm).

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="curatedOvarianData")
```

`E.MTAB.386_eset` *Angiogenic mRNA and microRNA gene expression signatures predict overall survival in serous ovarian cancer (mRNA data only).*

Description

Ovarian cancer is the fifth leading cause of cancer death for women in the U.S. and the seventh most fatal worldwide. Although ovarian cancer is notable for its initial sensitivity to platinum-based therapies, the vast majority of women eventually recurs and succumbs to increasingly platinum-resistant disease. Modern, targeted cancer drugs intervene in cell signaling pathways by compensating for their deregulation, and identifying these key mechanisms and pathways would greatly advance our ability to treat disease. In order to shed light on the molecular diversity of ovarian cancer, we performed comprehensive transcriptional profiling on 129 high grade, late stage serous ovarian cancers.

We implemented a novel, re-sampling based version of the ISIS class discovery algorithm (rISIS: robust ISIS) and applied it to the entire set of ovarian cancer transcriptional profiles. rISIS identified a novel stratification of this disease into two groups with significantly different overall survival. Gene set enrichment analysis found strong support for the stratification by extracellular matrix, cell adhesion, and angiogenesis genes. Application of this 'angiogenesis' signature to independent, published ovarian cancer gene expression data confirms its prognostic potential. Additional support for this stratification is provided by micro-RNA expression profiles which exhibit statistically significant expression differences between the groups, and additional mechanistic analyses have allowed development of hypotheses relevant to directed therapeutic intervention for specific subclasses of the disease. In particular, the subgroup stratification we discovered may be relevant for identifying which patients may be best suited for anti-angiogenic therapies that are now being tested in clinical trials.

Usage

```
data(E.MTAB.386_eset)
```

Format

Authors: Stefan Bentink, John Quackenbush, Jian-Bing Fan, Michelle S. Hirsch, Kristina Holton, Renee Rubio, Craig April, Jing Chen, Eliza Wickham-Garcia, Joyce Liu, Aedin Culhane, Ronny Drapkin, and Ursula A. Matulonis.

Contact: John Quackenbush, Stefan Bentink <bentink@jimmy.harvard.edu>, Ursula A Matulonis <ursula_matulonis@dfci.harvard.edu>

Title: Angiogenic mRNA and microRNA gene expression signatures predict overall survival in serous ovarian cancer.

Url: http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-386

PubMedID: 22348002

Details

Platforms used: A-MEXP-1811 - Illumina Human microRNA MI_V1_R0_XS0000122-MAP & A-MEXP-931 - Illumina HumanRef-8 v2 Expression BeadChip.

Overall survival time-to-event summary (in years):

Call: survfit(formula = Surv(time, cens) ~ -1)

Available sample meta-data:

---

alt_sample_name: Length = 141. Class = character. Mode = character.

sample_type: tumor = 141.

subtype: ser = 141.

primarysite: ov = 141.

summarystage: late = 140. NA's = 1.

tumorstage: 3 = 118. 4 = 22. NA's = 1.

substage: a = 4. b = 14. c = 101. NA's = 22.

age_at_initial_pathologic_diagnosis: Min. = 21.21 Median = 66.42 Mean = 60.84 Max. = 95.13

days_to_death: Min. = 3.9 Median = 920.1 Mean = 1039.36 Max. = 2724

vital_status: deceased = 81. living = 60.

debulking: optimal = 34. suboptimal = 107.

Source

http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-386

Examples

```
data(E.MTAB.386_eset)
## maybe str(GSE12418_eset) ; plot(GSE12418_eset) ...
if(require(affy)){
  summary(E.MTAB.386_eset$os_binary)
}
```

GSE12418_eset *Expression analysis of stage III serous ovarian adenocarcinoma distinguishes a sub-group of survivors.*

Description

In order to find novel candidate biomarkers to improve the outcome for patients with ovarian cancer, this study analyzed differences in gene expressions in 54 stage III serous ovarian adenocarcinomas with oligonucleotide microarrays containing 27,000 unique probes. The microarray data was verified with quantitative real time polymerase chain reaction for the genes TACC1, MUC5B and PRAME.

Usage

```
data(GSE12418_eset)
```

Format

Authors: Partheen K, Levan K, Osterberg L, Horvath G.

Lab: Department of Oncology, Goteborg University, SE-413 45 Goteborg, Sweden.

Contact: karolina.partheen@oncology.gu.se

Title: Expression analysis of stage III serous ovarian adenocarcinoma distinguishes a sub-group of survivors.

Url: http://pubget.com/search?q=16996261

PubMedID: 16996261.

Details

Platform used: SWEGENE H_v2.1.1_27k.

loess normalization was used. Processed data here are those provided by the original study authors.

assayData: 12633 features, 54 samples.

GEO_platform_accession : GSE12418.

platform_summary: GPL5886.

biomart_ID: Swegene

Binary overall survival summary (definitions of long and short provided by study authors):

long = 20. short = 34.

Available sample meta-data:

---

alt_sample_name: Length = 54. Class = character. Mode = character.

sample_type: tumor = 54.

subtype: ser = 54.

primarysite: ov = 54.

summarystage: late = 54.

tumorstage: 3 = 54.

substage: b = 19. c = 35.

age_at_initial_pathologic_diagnosis: Min. = 35.00 1st Qu. = 51.25 Median = 59.50 Mean = 59.56 3rd Qu. = 69.75 Max = 84.00

pltx: y = 54.

os_binary: long = 20. short = 34.

debulking: optimal = 13. suboptimal = 41.

Source

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE12418

Examples

```
data(GSE12418_eset)
## maybe str(GSE12418_eset) ; plot(GSE12418_eset) ...
if(require(affy)){
  summary(GSE12418_eset$os_binary)
}
```

GSE12470_eset *Gene expression profiling of advanced-stage serous ovarian cancers distinguishes novel subclasses and implicates ZEB2 in tumor progression and prognosis.*

Description

Gene expression profiles from 43 ovarian cancer tissues comprising 8 early stage and 35 advanced stage tissues were determined using oligonucleotide microarrays of 18,716 genes. 5 genes (ZEB2, CDH1, LTBP2, COL16A1 and ACTA2) were extracted as candidates for prognostic factors associated with progression-free survival. The findings suggest that the expressions of epithelial-mesenchymal transition-related genes such as ZEB2 and CDH1 may play important roles in the invasion process of advanced-stage serous ovarian cancer.

Usage

```
data(GSE12470_eset)
```

Format

Authors: Yoshihara K, Tajima A, Komata D, Yamamoto T, Kodama S, Fujiwara H, Suzuki M, Onishi Y, Hatae M, Sueyoshi K, Fujiwara H, Kudo Y, Inoue I, Tanaka K.

Lab: Department of Obstetrics and Gynecology, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan.

Contact: tanaken@med.niigata-u.ac.jp

Title: Gene expression profiling of advanced-stage serous ovarian cancers distinguishes novel subclasses and implicates ZEB2 in tumor progression and prognosis.

Url: http://pubget.com/search?q=19486012

PubMedID: 19486012

Details

Platform used: Agilent-012097 Human 1A Microarray (V2) G4110B (Feature Number version).

The data normalization was carried out using GeneSpring GX 7.3 (Agilent Technologies). Processed data here are those provided by the original study authors.

assayData: 16788 features, 53 samples.

GEO_platform_accession: GSE12470. platform_summary: GPL887. biomart_ID: Agilent G4110B agilent_wholegenome.

Available sample meta-data:

---

alt_sample_name: Length = 53. Class = character. Mode = character.

sample_type: healthy = 10. tumor = 43.

subtype: ser = 43. NA's = 10.

primarysite: ov = 53.

summarystage: early = 8. late = 35. NA's = 10.

tumorstage: 2 = 8. 4 = 35. NA's = 10.

debulking: unknown = 53.

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE12470

Examples

```
data(GSE12470_eset)
## maybe str(GSE12470_eset) ; plot(GSE12470_eset) ...
```

GSE13876_eset *Survival-related profile, pathways, and transcription factors in ovarian cancer.*

Description

Ovarian cancer has a poor prognosis due to advanced stage at presentation and either intrinsic or acquired resistance to classic cytotoxic drugs such as platinum and taxoids. Recent large clinical trials with different combinations and sequences of classic cytotoxic drugs indicate that further significant improvement in prognosis by this type of drugs is not to be expected. Currently a large number of drugs, targeting dysregulated molecular pathways in cancer cells have been developed and are introduced in the clinic. A major challenge is to identify those patients who will benefit from drugs targeting these specific dysregulated pathways.The aims of our study were (1) to develop a gene expression profile associated with overall survival in advanced stage serous ovarian cancer, (2) to assess the association of pathways and transcription factors with overall survival, and (3) to validate our identified profile and pathways/transcription factors in an independent set of ovarian cancers.

Usage

```
data(GSE13876_eset)
```

Format

Authors: Crijns AP, Fehrmann RS, de Jong S, Gerbens F, Meersma GJ, Klip HG, Hollema H, Hofstra RM, te Meerman GJ, de Vries EG, van der Zee AG.

Lab: Department of Gynecologic Oncology, University Medical Centre Groningen, University of Groningen, Groningen, The Netherlands.

Contact: A.G.J.van.der.Zee@og.umcg.nl

Title: Survival-Related Profile, Pathways, and Transcription Factors in Ovarian Cancer.

Url: http://pubget.com/search?q=19192944

PubMedID: 19192944

Details

Platform used: Operon human v3 ~35K 70-mer two-color oligonucleotide microarrays.

Quantile normalization was used. Processed data here are those provided by the original study authors.

assayData: 20112 features, 157 samples.

GEO_platform_accession: GSE13876. platform_summary : GPL7759. biomart_ID: Operon two-color oligo

Overall survival time-to-event summary (in years): Call: survfit(formula = Surv(time, cens) ~ -1)

records: 157.00 n.max = 157.00 n.start = 157.00 events = 113.00 median = 2.05 0.95LCL = 1.56 0.95UCL = 2.71

Available sample meta-data:

---------- alt_sample_name: Length = 157. Class = character. Mode = character.

unique_patient_ID: Min. = 1. 1st Qu. = 40. Median = 79. Mean = 79. 3rd Qu. = 118. Max. = 157.

sample_type: tumor = 157.

subtype: ser = 157. 157

primarysite: ov = 157.

summarystage: late = 157.

tumorstage 4 = 157.

age_at_initial_pathologic_diagnosis: Min. = 21.00 1st Qu. = 50.00 Median = 60.00 Mean = 57.95 3rd Qu. = 67.00 Max = 84.00

recurrence_status: norecurrence = 44. recurrence = 113.

days_to_death: Min. = 30. 1st Qu. = 360. Median = 630. Mean = 1100. 3rd Qu. = 1470. Max. = 7020.

vital_status: deceased = 113. living = 44.

debulking: unknown = 157. 157

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE13876

Examples

```
data(GSE13876_eset)
## maybe str(GSE13876_eset) ; plot(GSE13876_eset) ...
if(require(affy)){
  summary(GSE13876_eset$recurrence_status)
  time <- GSE13876_eset$days_to_death / 365
  cens <- ifelse(GSE13876_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  fit          #for summary of survival
```

```
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
}
```

GSE14764_eset *A prognostic gene expression index in ovarian cancer - validation across different independent data sets.*

Description

Ovarian carcinoma has the highest mortality rate among gynaecological malignancies. In this project, we investigated the hypothesis that molecular markers are able to predict outcome of ovarian cancer independently of classical clinical predictors, and that these molecular markers can be validated using independent data sets. We applied a semi-supervised method for prediction of patient survival. Microarrays from a cohort of 80 ovarian carcinomas (TOC cohort) were used for the development of a predictive model, which was then evaluated in an entirely independent cohort of 118 carcinomas (Duke cohort). A 300-gene ovarian prognostic index (OPI) was generated and validated in a leave one out approach in the TOC cohort (Kaplan-Meier analysis, p = 0.0087). In a second validation step, the prognostic power of the OPI was confirmed in an independent data set (Duke cohort, p = 0.0063). In multivariate analysis, the OPI was independent of the post-operative residual tumour, the main clinico-pathological prognostic parameter with an adjusted hazard ratio of 6.4 (TOC cohort, CI 1.8-23.5, p = 0.0049) and 1.9 (Duke cohort, CI 1.2-3.0, p = 0.0068). We constructed a combined score of molecular data (OPI) and clinical parameters (residual tumour), which was able to define patient groups with highly significant differences in survival. The integrated analysis of gene expression data as well as residual tumour can be used for optimized assessment of the prognosis of platinum taxol treated ovarian cancer. As traditional treatment options are limited, this analysis may be able to optimize clinical management and to identify those patients who would be candidates for new therapeutic strategies.

Usage

```
data(GSE14764_eset)
```

Format

Authors: Denkert C, Budczies J, Darb-Esfahani S, Gyorffy B et al.

Lab: Institute of Pathology, Charite University Hospital, Berlin, Germany.

Contact: carsten.denkert@charite.de

Title: A prognostic gene expression index in ovarian cancer: Validation across different independent data sets.

Url: http://pubget.com/search?q=19294737

PubMedID: 19294737

Details

Platform used: [HG-U133A] Affymetrix Human Genome U133A Array.

frma normalization used from the frma bioconductor package.

assayData: 12858 features, 80 samples.

GEO_platform_accession: GSE14764. platform_summary: GPL96. biomart_ID: HG-U133A affy_hg_u133a.

Overall survival time-to-event summary (in years):

Call: survfit(formula = Surv(time, cens) ~ -1)

records = 80.00 n.max = 80.00 n.start = 80.00 events = 21.00 median = 4.52 0.95LCL = 4.19 0.95UCL = NA

Available sample meta-data:

---------- alt_sample_name: Min. = 1.00 1st Qu. = 20.75 Median = 40.50 Mean = 40.50 3rd Qu. = 60.25 Max. = 80.00

sample_type: tumor = 80.

subtype: Length = 80. Class = character. Mode = character.

primarysite: ov = 80.

summarygrade: high = 54. low = 26.

summarystage: early = 9. late = 71.

tumorstage: 1 = 8. 2 = 1. 3 = 69. 4 = 2.

substage: a = 4. b = 6. c = 32. NA's = 38.

G: 1 = 3. 2 = 23. 3 = 54.

recurrence_status: norecurrence = 50. recurrence = 26. NA's = 4.

days_to_death: Min. = 210. 1st Qu. = 660. Median = 1050. Mean = 1011. 3rd Qu. = 1328. Max = 2190.

vital_status: deceased = 21. living = 59.

debulking: unknown = 80.

batch:

2004-09-29 2004-09-30 2004-10-01 2005-01-21 2005-01-25 2005-01-26 2005-01-28

1 2 6 4 7 8 10

2005-03-02 2006-07-26 2006-07-27 2006-07-28 2006-08-11 2006-08-18 2006-08-19

6 4 6 4 10 3 4

2006-08-21

5

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE14764

Examples

```
data(GSE14764_eset)
## maybe str(GSE14764_eset) ; plot(GSE14764_eset) ...
if(require(affy)){
  summary(GSE14764_eset$recurrence_status)
  time <- GSE14764_eset$days_to_death / 365
  cens <- ifelse(GSE14764_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  fit          #for summary of survival
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
}
```

GSE17260_eset *Gene expression profile for predicting survival in advanced-stage serous ovarian cancer across two independent datasets.*

Description

Advanced-stage ovarian cancer patients are generally treated with platinum/taxane-based chemotherapy after primary debulking surgery. However, there is a wide range of outcomes for individual patients. Therefore, the clinicopathological factors alone are insufficient for predicting prognosis. Our aim is to identify a progression-free survival (PFS)-related molecular profile for predicting survival of patients with advanced-stage serous ovarian cancer.

Usage

```
data(GSE17260_eset)
```

Format

Authors: Yoshihara K, Tajima A, Yahata T, Kodama S, Fujiwara H, Suzuki M, Onishi Y, Hatae M, Sueyoshi K, Fujiwara H, Kudo Y, Kotera K, Masuzaki H, Tashiro H, Katabuchi H, Inoue I, Tanaka K.

Lab : Department of Obstetrics and Gynecology, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan.

Contact: tanaken@med.niigata-u.ac.jp

Title: Gene expression profile for predicting survival in advanced-stage serous ovarian cancer across two independent datasets.

Url: http://pubget.com/search?q=20300634

PubMedID: 20300634

Details

Platform used: Agilent-012391 Whole Human Genome Oligo Microarray G4112A.

Data normalization was performed in GeneSpring GX 10 (Agilent Technologies). Processed data here are those provided by the original study authors.

assayData: 18844 features, 110 samples.

GEO_platform_accession: GSE17260. platform_summary = GPL6848. biomart_ID = Agilent G4112A agilent_wholegenome.

Overall survival time-to-event summary (in years): Call: survfit(formula = Surv(time, cens) ~ -1)

records = 110.00 n.max = 110.00 n.start = 110.00 events = 46.00 median = 4.44 0.95LCL = 4.03 0.95UCL = NA

Available sample meta-data:

---

alt_sample_name: Length = 110. Class = character. Mode = character.

sample_type: tumor = 110.

subtype: ser = 110.

primarysite: ov = 110.

summarystage: late = 110.

tumorstage: 3 = 93. 4 = 17.

substage: a = 6. b = 18. c = 69. NA's = 17.

pltx: y = 110.

tax: y = 110.

days_to_tumor_recurrence: Min. = 30.0 1st Qu. = 285.0 Median = 510.0 Mean = 673.9 3rd Qu. = 870.0 Max. = 2250.0

recurrence_status: norecurrence = 34. recurrence = 76.

days_to_death: Min. = 30 1st Qu. = 660 Median = 915 Mean = 1086 3rd Qu. = 1530 Max. = 2430

vital_status: deceased = 46. living = 64.

debulking: optimal = 57. suboptimal = 53.

Source

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE17260

Examples

```
data(GSE17260_eset)
## maybe str(GSE17260_eset) ; plot(GSE17260_eset) ...
if(require(affy)){
  summary(GSE17260_eset$recurrence_status)
  time <- GSE17260_eset$days_to_death / 365
  cens <- ifelse(GSE17260_eset$vital_status=="deceased",1,0)
  library(survival)
```

```
  fit <- survfit(Surv(time,cens)~-1)
  fit          #for summary of survival
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
}
```

GSE18520_eset *A gene signature predictive for outcome in advanced ovarian cancer identifies a survival factor: microfibril-associated glycoprotein 2.*

Description

Advanced stage papillary serous tumors of the ovary are responsible for the majority of ovarian cancer deaths, yet the molecular determinants modulating patient survival are poorly characterized. Here, we identify and validate a prognostic gene expression signature correlating with survival in a series of microdissected serous ovarian tumors. Independent evaluation confirmed the association of a prognostic gene microfibril-associated glycoprotein 2 (MAGP2) with poor prognosis, whereas in vitro mechanistic analyses demonstrated its ability to prolong tumor cell survival and stimulate endothelial cell motility and survival via the alpha(V)beta(3) integrin receptor. Increased MAGP2 expression correlated with microvessel density suggesting a proangiogenic role in vivo. Thus, MAGP2 may serve as a survival-associated target.

Usage

```
data(GSE18520_eset)
```

Format

Authors: Mok SC, Bonome T, Vathipadiekal V, Bell A, Johnson ME, Wong KK, Park DC, Hao K, Yip DK, Donninger H, Ozbun L, Samimi G, Brady J, Randonovich M, Pise-Masison CA, Barrett JC, Wong WH, Welch WR, Berkowitz RS, Birrer MJ.

Lab: Department of Gynecologic Oncology, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA.

Contact:mbirrer@partners.org

Title: A gene signature predictive for outcome in advanced ovarian cancer identifies a survival factor: microfibril-associated glycoprotein 2.

Url: http://pubget.com/search?q=19962670

PubMedID: 19962670

Details

Platform used: [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array.

frma normalization used from the frma bioconductor package.

assayData: 18708 features, 63 samples.

GEO_platform_accession: GSE18520 platform_summary: GPL570 biomart_ID: HG-U133_Plus_2 affy_hg_u133_plus_2

Overall survival time-to-event summary (in years): Call: survfit(formula = Surv(time, cens) ~ -1)

10 observations deleted due to missingness

records = 53.00 n.max = 53.00 n.start = 53.00 events = 41.00 median = 2.05 0.95LCL = 1.48 0.95UCL = 3.70

Available sample meta-data:

---

alt_sample_name: Min. = 312.0 1st Qu. = 395.0 Median = 694.0 Mean = 893.3 3rd Qu. = 1040.0 Max. = 2237.0

sample_type: healthy = 10. tumor = 53.

subtype: ser = 53. NA's = 10.

primarysite: ov = 63.

summarygrade: high = 53. NA's = 10.

summarystage: late = 53. NA's = 10.

tumorstage: 4 = 53. NA's = 10.

G: 4 = 53. NA's = 10.

days_to_death: Min. = 150 1st Qu. = 450 Median = 630 Mean = 1212 3rd Qu. = 1440 Max. = 4500 NA's = 10

vital_status: deceased = 41. living = 12. NA's = 10.

debulking: unknown = 63.

batch:

2004-03-12 2004-04-08 2004-04-09 2004-07-20 2004-08-12 2004-08-13 2004-09-30

20 6 9 11 10 1 6

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE18520

Examples

```
data(GSE18520_eset)
## maybe str(GSE18520_eset) ; plot(GSE18520_eset) ...
if(require(affy)){
  summary(GSE18520_eset$vital_status)
  time <- GSE18520_eset$days_to_death / 365
```

```
  cens <- ifelse(GSE18520_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  fit          #for summary of survival
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
}
```

`GSE19829.GPL570_eset` *Gene Expression Profile of BRCAness That Correlates With Responsiveness to Chemotherapy and With Outcome in Patients With Epithelial Ovarian Cancer.*

Description

To define a gene expression profile of BRCAness that correlates with chemotherapy response and outcome in epithelial ovarian cancer (EOC).

Usage

```
data(GSE19829.GPL570_eset)
```

Format

Authors: Konstantinopoulos PA, Spentzos D, Karlan BY, Taniguchi T et al.

Lab: Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA.

Contact: scannist@bidmc.harvard.edu

Title: Gene Expression Profile of BRCAness That Correlates With Responsiveness to Chemotherapy and With Outcome in Patients With Epithelial Ovarian Cancer.

Url: http://pubget.com/search?q=20547991

PubMedID: 20547991

Details

Platform used: [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array.

frma normalization used from the frma bioconductor package.

assayData: 18708 features, 28 samples.

GEO_platform_accession: GSE19829-GPL570. platform_summary: GPL570 biomart_ID: HG-U133_Plus_2 affy_hg_u133_plus_2

Overall survival time-to-event summary (in years): Call: survfit(formula = Surv(time, cens) ~ -1)

records = 28.00 n.max = 28.00 n.start = 28.00 events = 17.00 median = 3.95 0.95LCL = 2.71 0.95UCL = NA

Available sample meta-data:

---

alt_sample_name: Length = 28. Class = character. Mode = character.

sample_type: tumor = 28.

primarysite: ov = 28.

days_to_death: Min. = 150 1st Qu.= 540 Median = 1050 Mean = 1291 3rd. Qu = 1688 Max. = 3450

vital_status: deceased = 17. living = 11.

debulking: unknown = 28.

batch:

2009-08-14

28

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE19829

Examples

```
data(GSE19829.GPL570_eset)
## maybe str(GSE19829-GPL570_eset) ; plot(GSE19829-GPL570_eset) ...

if(require(affy)){
  summary(GSE19829.GPL570_eset$vital_status)
  time <- GSE19829.GPL570_eset$days_to_death / 365
  cens <- ifelse(GSE19829.GPL570_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  ##fit          #for summary of survival
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
 }
```

GSE19829.GPL8300_eset *Gene Expression Profile of BRCAness That Correlates With Responsiveness to Chemotherapy and With Outcome in Patients With Epithelial Ovarian Cancer.*

Description

To define a gene expression profile of BRCAness that correlates with chemotherapy response and outcome in epithelial ovarian cancer (EOC).

Usage

```
data(GSE19829.GPL8300_eset)
```

Format

Authors: Konstantinopoulos PA, Spentzos D, Karlan BY, Taniguchi T et al.

Lab: Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA.

Contact: scannist@bidmc.harvard.edu

Title: Gene Expression Profile of BRCAness That Correlates With Responsiveness to Chemotherapy and With Outcome in Patients With Epithelial Ovarian Cancer.

Url: http://pubget.com/search?q=20547991

PubMedID: 20547991

Details

Platform used: [HG_U95Av2] Affymetrix Human Genome U95 Version 2 Array.

rma normalization used from the rma bioconductor package.

assayData: 8913 features, 42 samples.

GEO_platform_accession: GSE19829-GPL8300 platform_summary: GPL8300 biomart_ID: HG_U95Av2 affy_hg_u95av2

Overall survival time-to-event summary (in years): Call: survfit(formula = Surv(time, cens) ~ -1)

records = 42.00 n.max = 42.00 n.start = 42.00 events = 23.00 median = 3.78 0.95LCL = 2.79 0.95UCL = NA

Available sample meta-data:

---

alt_sample_name: Length = 42. Class = character. Mode = character.

sample_type: tumor = 42.

primarysite: ov = 42.

days_to_death: Min. = 30.0 1st Qu. = 727.5 Median = 1155.0 Mean = 1089.0 3rd Qu. = 1485.0 Max. = 2040.0

vital_status: deceased = 23. living = 19.

debulking: unknown = 42.

batch:

2001-09-14 2001-12-14 2002-08-20 2003-09-09 2003-09-18

7 4 14 13 4

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE19829

Examples

```
data(GSE19829.GPL8300_eset)
## maybe str(GSE19829-GPL8300_eset) ; plot(GSE19829-GPL8300_eset) ...
if(require(affy)){
  summary(GSE19829.GPL8300_eset$vital_status)
  time <- GSE19829.GPL8300_eset$days_to_death / 365
  cens <- ifelse(GSE19829.GPL8300_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  ##fit          #for summary of survival
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
 }
```

`GSE20565_eset` *A genomic and transcriptomic approach for a differential diagnosis between primary and secondary ovarian carcinomas in patients with a previous history of breast cancer.*

Description

The distinction between primary and secondary ovarian tumors may be challenging for pathologists. The purpose of the present work was to develop genomic and transcriptomic tools to further refine the pathological diagnosis of ovarian tumors after a previous history of breast cancer.

Usage

```
data(GSE20565_eset)
```

Format

Authors: Meyniel JP, Cottu PH, Decraene C, Stern MH, Couturier J, Lebigot I, Nicolas A, Weber N, Fourchotte V, Alran S, Rapinat A, Gentien D, Roman-Roman S, Mignot L, Sastre-Garau X.

Lab: Department of Translational Research, Institut Curie, 26 rue d'Ulm, 75248 Paris, Cedex 05, France.

Contact: jean-philippe.meyniel@curie.fr

Title: A genomic and transcriptomic approach for a differential diagnosis between primary and secondary ovarian carcinomas in patients with a previous history of breast cancer.

Url: http://pubget.com/search?q=20492709.

PubMedID: 20492709.

Details

Platforms used: [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0.

In the study but not included in curatedOvarianData:

**Array|[Mapping50K_Xba240] Affymetrix Human Mapping 50K Xba240 SNP.

**Array|[GenomeWideSNP_6] Affymetrix Genome-Wide Human SNP 6.0 Array.

frma normalization used from frma bioconductor package.

assayData: 18708 features, 140 samples.

Available sample meta-data:

---

alt_sample_name: Length = 140. Class = character. Mode = character.

sample_type: tumor = 140.

subtype: Length = 140. Class = character. Mode = character.

primarysite: other = 44. ov = 96.

summarygrade: high = 63. low = 33. NA's = 44.

summarystage: early = 27. late = 67. NA's = 46.

tumorstage: 1 = 18. 2 = 9. 3 = 52. 4 = 15. NA's = 46.

substage: a = 14. b = 10. c = 55. NA's = 61.

G: 1 = 6. 2 = 27. 3 = 63. NA's = 44.

M: 0 = 96. 1 = 44.

debulking: unknown = 140.

batch:

2006-06-01 2006-06-27 2006-06-28 2006-06-29 2006-06-30 2006-07-20 2008-03-06

21 18 37 20 36 7 1

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE20565

Examples

```
data(GSE20565_eset)
## maybe str(GSE20565_eset) ; plot(GSE20565_eset) ...
```

GSE2109_eset *Expression Project for Oncology (expO)*

Description

This project seeks to integrate longitudinal clinical annotation with gene expression data for a unique and powerful portrait of human malignancies, providing critical perspective on diagnositic markers, prognostic indicators and therapeutic targets. Tissue samples were procured under standard conditions and gene expression analyses were performed on a clinically annotated set of tumor samples.

Usage

```
data(GSE2109_eset)
```

Format

Title: The Expression Project for Oncology (EXPO)

Contact: ecurley@intgen.org

Url: http://www.intgen.org/expo/

Details

Platform used: [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array.

frma normalization used from the frma bioconductor package.

assayData: 18708 features, 204 samples

GEO_platform_accession: GSE2109

platform_summary: GPL570

biomart_ID: HG-U133_Plus_2 affy_hg_u133_plus_2

Available sample meta-data:

---

alt_sample_name: Length= 204. Class= character. Mode= character.

sample_type: tumor = 204.

subtype: Length = 204. Class = character. Mode = character.

primarysite: other= 23. ov= 178. unknown= 3.

summarygrade: high= 91. low= 31. NA's= 82.

summarystage: early= 37. late= 85. NA's= 82.

T: 1= 21. 2= 16. 3= 85. NA's= 82.

substage: a = 17. b= 22. c= 79. NA's = 86.

G: 1 = 11. 2= 20. 3= 83. 4= 8. NA's= 82.

N: 0 = 60. 1 = 43. 2= 2. NA's= 99.

M: 0 = 92. 1 = 18. NA's = 94.

age_at_initial_pathologic_diagnosis: Min.= 25.00 1st Qu.= 45.00 Median= 55.00 Mean= 58.82 3rd Qu.= 65.00 Max.= 85.00

debulking: unknown = 204

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE2109

Examples

```
data(GSE2109_eset)
## maybe str(GSE2109_eset) ; plot(GSE2109_eset) ...
```

`GSE26712_eset` *A gene signature predicting for survival in suboptimally debulked patients with ovarian cancer.*

Description

Despite the existence of morphologically indistinguishable disease, patients with advanced ovarian tumors display a broad range of survival end points. We hypothesize that gene expression profiling can identify a prognostic signature accounting for these distinct clinical outcomes. To resolve survival-associated loci, gene expression profiling was completed for an extensive set of 185 (90 optimal/95 suboptimal) primary ovarian tumors using the Affymetrix human U133A microarray. Cox regression analysis identified probe sets associated with survival in optimally and suboptimally debulked tumor sets at a P value of <0.01. Leave-one-out cross-validation was applied to each tumor cohort and confirmed by a permutation test. External validation was conducted by applying the gene signature to a publicly available array database of expression profiles of advanced stage suboptimally debulked tumors. The prognostic signature successfully classified the tumors according to survival for suboptimally (P = 0.0179) but not optimally debulked (P = 0.144) patients. The suboptimal gene signature was validated using the independent set of tumors (odds ratio, 8.75; P = 0.0146). To elucidate signaling events amenable to therapeutic intervention in suboptimally debulked patients, pathway analysis was completed for the top 57 survival-associated probe sets. For suboptimally debulked patients, confirmation of the predictive gene signature supports the existence of a clinically relevant predictor, as well as the possibility of novel therapeutic opportunities. Ultimately, the prognostic classifier defined for suboptimally debulked tumors may aid in the classification and enhancement of patient outcome for this high-risk population.

Usage

```
data(GSE26712_eset)
```

Format

Authors: Bonome T, Levine DA, Shih J, Randonovich M et al

Lab: Cell and Cancer Biology Branch, National Cancer Institute, NIH, Rockville, Maryland 20892, USA.

Contact: birrerm@mail.nih.gov

Title: A gene signature predicting for survival in suboptimally debulked patients with ovarian cancer.

Url: http://pubget.com/search?q=18593951

PubMedID: 18593951

Details

Platform used: [HG-U133A] Affymetrix Human Genome U133A Array.

frma normalization used from the frma bioconductor package.

assayData: 12858 features, 195 samples.

GEO_platform_accession: GSE26712

platform_summary: GPL96

biomart_ID: HG-U133A affy_hg_u133a

Overall survival time-to-event summary (in years):

Call: survfit(formula = Surv(time, cens) ~ -1)

10 observations deleted due to missingness

records = 185.00 n.max = 185.00 n.start = 185.00 events = 129.00 median = 3.83 0.95LCL = 3.24 0.95UCL = 4.83

Available sample meta-data:

---

alt_sample_name: Length = 195. Class = character. Mode = character.

sample_type: healthy = 10. tumor = 185.

subtype: ser = 185. NA's = 10.

primarysite: ov = 195.

summarygrade: high = 185. NA's = 10.

summarystage: late = 185. NA's = 10.

tumorstage: 4 = 185. NA's = 10.

G: 4 = 185. NA's = 10.

recurrence_status: norecurrence = 42. recurrence = 153.

days_to_death: Min. = 21.9 1st Qu. = 660.6 Median = 1164.0 Mean = 1429.0 3rd Qu. = 1880.0 Max. = 4982.0 NA's = 10.0

vital_status: deceased = 129. living = 56. NA's = 10.

debulking: optimal = 90. suboptimal = 95. unknown = 10.

batch:

2003-11-04 2003-11-05 2003-11-06 2003-11-07 2003-11-20 2003-11-21 2003-12-16

14 16 9 6 10 15 17

2003-12-23 2003-12-24 2004-04-20 2004-04-21 2004-04-27 2004-09-28 2005-07-27

12 11 20 17 9 14 15

2006-11-09

10

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE26712

Examples

```
data(GSE26712_eset)
## maybe str(GSE26712_eset) ; plot(GSE26712_eset) ...

if(require(affy)){
  summary(GSE26712_eset$recurrence_status)
  time <- GSE26712_eset$days_to_death / 365
  cens <- ifelse(GSE26712_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  fit          #for summary of survival
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
}
```

GSE30009_eset *Multidrug Resistance-Linked Gene Signature Predicts Overall Survival of Patients with Primary Ovarian Serous Carcinoma*

Description

qRT-PCR dataset. This study assesses the ability of multidrug resistance (MDR)-associated gene expression patterns to predict survival in patients with newly diagnosed carcinoma of the ovary. The scope of this research differs substantially from that of previous reports, as a very large set of genes was evaluated whose expression has been shown to affect response to chemotherapy.

Usage

```
data(GSE30009_eset)
```

Format

Note this is a qRT-PCR dataset, not microarray.

Authors: Gillet J, Calcagno AM, Varma S, Davidson B, Elstrand MB, Ganapathi R, Kamat A, Sood A, Ambudkar SV, Seiden M, Rueda B, Gottesman MM

Contact: Jean-Pierre Gillet <gilletjp@mail.nih.gov>

Title: Multidrug resistance-linked gene signature predicts overall survival of patients with primary ovarian serous carcinoma.

Url: http://pubget.com/search?q=22492981

PubMedID: 22492981

Details

```
Platform used: TaqMan qRT-PCR Homo sapiens Low-Density Array 380

Processed data here are those provided by the original study authors.

assayData: 380 features, 103 samples.

GEO_platform_accession: GPL13728


-------------------------------------------------------------------------------

Available sample meta-data:

-------------------------------------------------------------------------------

sample_type
tumor
  103

subtype
clearcell       ser
        1       102

summarygrade
high  low
  92    9

summarystage
late
 103

T   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  3.000   3.000   3.000   3.204   3.000   4.000

substage
```

```
 b  c
 2 60

G   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
  1.000   3.000   3.000   2.871   3.000   3.000       2

age_at_initial_pathologic_diagnosis   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  30.00   56.00   61.00   62.45   71.50   87.00

days_to_death   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
     24     598    1053    1156    1568    4748

vital_status
deceased   living
      57       46

debulking
   optimal suboptimal
   81         22
```

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE30009

Examples

```
data(GSE30009_eset)
meta.data <- pData(GSE30009_eset)
for (i in 1:ncol(meta.data)){
   if (!all(is.na(meta.data[[i]])) & length(unique(meta.data[[i]])) < length(meta.data[[i]])){
        cat(colnames(meta.data)[i])
        if(class(meta.data[[i]]) == "numeric" || class(meta.data[[i]]) == "integer"){
            print(summary(meta.data[[i]]))
     }else if(class(meta.data[[i]]) == "character" || class(meta.data[[i]]) == "factor" || class(meta.data[[i]]) =
            print(table(meta.data[[i]]))
        }
        cat("\n")
    }
}
if(require(affy)){
  summary(GSE30009_eset$vital_status)
  time <- GSE30009_eset$days_to_death / 365
  cens <- ifelse(GSE30009_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  fit          #for summary of survival
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
```

```
}
```

`GSE30161_eset` *Multi-gene expression predictors of single drug responses to adjuvant chemotherapy in ovarian carcinoma: predicting platinum resistance*

Description

We adapted and applied the previously-published COXEN algorithm to develop molecular predictors for therapeutic responses of patients' tumors based on expression signatures derived from the NCI-60 in vitro drug activities and genomic expression data. Genome-wide candidate biomarkers were first triaged by examining expression patterns of frozen and formalin-fixed paraffin embedded (FFPE) tissue samples. We then identify initial drug sensitivity biomarkers for carboplatin and paclitaxel, respectively. These biomarkers were further narrowed by examining concordant expression patterns between cell lines and a historical set of ovarian cancer patients. Multivariate predictors were obtained from the NCI-60 cell lines and refined using historical patient cohorts. To independent validate these molecular predictors, we performed genome-wide profiling on FFPE samples of 58 ovarian cancer patients obtained prior to adjuvant chemotherapy.

Usage

```
data(GSE30161_eset)
```

Format

Authors: Ferriss JS, Kim Y, Duska L, Birrer M, Levine DA, Moskaluk C, Theodorescu D, Lee JK.

Lab: University of Virginia, Department of Public Health Sciences

Contact: Youngchul Kim

Title: Multi-gene expression predictors of single drug responses to adjuvant chemotherapy in ovarian carcinoma: predicting platinum resistance.

PubMedID: 22348014

Details

```
Platform used: [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array.

frma normalization used from the frma bioconductor package.

assayData: 18708 features, 63 samples.

GEO_platform_accession: GSE30161

platform_summary: GPL570

biomart_ID: HG-U133_Plus_2 affy_hg_u133_plus_2
```

```
assayData: 19093 features, 58 samples

         GEO_platform_accession platform_summary          biomart_ID
GSE30161                 GPL570   HG-U133_Plus_2 affy_hg_u133_plus_2

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
  58.00   58.00   58.00   36.00    4.19    2.70    6.17

----------------------------
Available sample meta-data:
----------------------------

alt_sample_name:
   Length     Class      Mode
       58 character character

sample_type:
tumor
   58

subtype:
   Length     Class      Mode
       58 character character

summarygrade:
high  low NA's
  33   21    4

summarystage:
late
  58

tumorstage:
 3  4
53  5

substage:
 a  b  c
 9 11 38

G:
   1    2    3 NA's
   2   19   33    4

age_at_initial_pathologic_diagnosis:
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  38.00   53.50   62.00   62.57   72.00   85.00

pltx:
 y
58

tax:
 n  y
 4 54

neo:
 n
58

days_to_tumor_recurrence:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   12.0   255.2   386.0   742.1   768.2  4208.0

recurrence_status:
norecurrence   recurrence         NA's
           6           48            4

days_to_death:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   49.0   585.2  1010.0  1375.0  2131.0  4208.0

vital_status:
deceased   living
      36       22

debulking:
   optimal suboptimal       NA's
        26         30          2
```

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE30161

Examples

```
data(GSE30161_eset)
## maybe str(GSE30161_eset) ; plot(GSE30161_eset) ...
if(require(affy)){
  summary(GSE30161_eset$vital_status)
  time <- GSE30161_eset$days_to_death / 365
  cens <- ifelse(GSE30161_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  fit          #for summary of survival
```

```
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
}
```

GSE32062.GPL6480_eset *Japanese dataset A: Immune-activation as a therapeutic direction for patients with high-risk ovarian cancer based on gene expression signature*

Description

Two hundred sixty patients who were diagnosed as advanced-stage high-grade serous ovarian cancer were analyzed in this study.

Usage

```
data(GSE32062.GPL6480_eset)
```

Format

Authors: Yoshihara K, Tsunoda T, Shigemizu D, Fujiwara H, Hatae M, Fujiwara H, Masuzaki H, Katabuchi H, Kawakami Y, Okamoto A, Nogawa T, Matsumura N, Udagawa Y, Saito T, Itamochi H, Takano M, Miyagi E, Sudo T, Ushijima K, Iwase H, Seki H, Terao Y, Enomoto T, Mikami M, Akazawa K, Tsuda H, Moriya T, Tajima A, Inoue I, Tanaka K

Lab: The Japanese Serous Ovarian Cancer Study G

Contact: Kosuke Yoshihara <yoshikou@med.niigata-u.ac.jp>

Title: High-risk ovarian cancer based on 126-gene expression signature is uniquely characterized by downregulation of antigen presentation pathway.

Url: http://pubget.com/search?q=22241791

PubMedID: 22241791

Details

Platform used: Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version)

Processed data here are those provided by the original study authors.

assayData: 41108 features, 260 samples.

GEO_platform_accession: GPL6480

Available sample meta-data:

---

sample_type tumor 260

subtype ser 260

summarygrade high low 129 131

summarystage late 260

T 3 4 204 56

substage a b c 4 20 180

G 2 3 131 129

pltx y 260

tax y 260

days_to_death Min. 1st Qu. Median Mean 3rd Qu. Max. 30 810 1245 1344 1710 3840

vital_status deceased living 121 139

debulking optimal suboptimal 103 157

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE32062

Examples

```
data(GSE32062.GPL6480_eset)
if(require(affy)){
  summary(GSE32062.GPL6480_eset$vital_status)
  time <- GSE32062.GPL6480_eset$days_to_death / 365
  cens <- ifelse(GSE32062.GPL6480_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  fit          #for summary of survival
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
}
```

GSE32063_eset *Japanese dataset B: Immune-activation as a therapeutic direction for patients with high-risk ovarian cancer based on gene expression signature*

Description

Two hundred sixty patients who were diagnosed as advanced-stage high-grade serous ovarian cancer were analyzed in this study.

Usage

```
data(GSE32063_eset)
```

Format

Authors: Yoshihara K, Tsunoda T, Shigemizu D, Fujiwara H, Hatae M, Fujiwara H, Masuzaki H, Katabuchi H, Kawakami Y, Okamoto A, Nogawa T, Matsumura N, Udagawa Y, Saito T, Itamochi H, Takano M, Miyagi E, Sudo T, Ushijima K, Iwase H, Seki H, Terao Y, Enomoto T, Mikami M, Akazawa K, Tsuda H, Moriya T, Tajima A, Inoue I, Tanaka K

Lab: The Japanese Serous Ovarian Cancer Study G

Contact: Kosuke Yoshihara <yoshikou@med.niigata-u.ac.jp>

Title: High-risk ovarian cancer based on 126-gene expression signature is uniquely characterized by downregulation of antigen presentation pathway.

Url: http://pubget.com/search?q=22241791

PubMedID: 22241791

Details

Platform used: Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version)

Processed data here are those provided by the original study authors.

assayData: 41108 features, 40 samples.

GEO_platform_accession: GPL6480

Available sample meta-data:

---

sample_type tumor 40

subtype ser 40

summarygrade high low 17 23

summarystage late 40

T 3 4 31 9

substage b c 3 28

G 2 3 23 17

pltx y 40

tax y 40

days_to_death Min. 1st Qu. Median Mean 3rd Qu. Max. 210 705 1155 1346 1792 3330

vital_status deceased living 22 18

debulking optimal suboptimal 19 21

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE32063

Examples

```
data(GSE32063_eset)
if(require(affy)){
  summary(GSE32063_eset$vital_status)
  time <- GSE32063_eset$days_to_death / 365
  cens <- ifelse(GSE32063_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  fit          #for summary of survival
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
}
```

GSE6008_eset *Human Ovarian Tumors and Normal Ovaries*

Description

Lysophosphatidic acid (LPA) governs a number of physiologic and pathophysiological processes. Malignant ascites fluid is rich in LPA, and LPA receptors are aberrantly expressed by ovarian cancer cells, implicating LPA in the initiation and progression of ovarian cancer. However, there is an absence of systematic data critically analyzing the transcriptional changes induced by LPA in ovarian cancer.

Usage

```
data(GSE6008_eset)
```

Format

Authors: Murph MM, Liu W, Yu S, Lu Y, Hall H, Hennessy BT, Lahad J, Schaner M, Helland A, Kristensen G, Mills GB

Lab: Department of Systems Biology, University of Texas MD Anderson Cancer Center, Houston, Texas, United States of America.

Contact: mmurph@rx.uga.edu

Title: Lysophosphatidic acid-induced transcriptional profile represents serous epithelial ovarian carcinoma and worsened prognosis.

Url: http://pubget.com/search?q=19440550

PubMedID: 19440550

Details

Platform used: [HG-U133A] Affymetrix Human Genome U133A Array.

frma normalization used from the frma bioconductor package.

assayData: 12858 features, 99 samples.

GEO_platform_accession: GSE6008

platform_summary: GPL96

biomart_ID: HG-U133A affy_hg_u133a

Available sample meta-data:

---

alt_sample_name:

Length = 99. Class = character. Mode = character.

sample_type: tumor = 99.

subtype: clearcell = 8. endo = 37. mucinous = 13. ser = 41.

primarysite: ov = 99.

summarygrade: high = 38. low = 36. NA's = 25.

summarystage: early = 42. late = 53. NA's = 4.

tumorstage 1 = 31. 2 = 11. 3 = 44. 4 = 9. NA's = 4.

substage: a = 19. b = 2. c = 54. d = 1. NA's = 23.

G: 1 = 19. 2 = 17. 3 = 38. NA's = 25.

debulking: unknown = 99.

batch:

2002-04-03 2002-04-04 2002-04-09 2002-04-10 2002-04-12 2002-08-13 2002-08-15

3 8 9 2 3 4 4

2002-08-22 2002-08-23 2002-08-27 2002-08-28 2002-08-29 2002-08-30 2002-09-11

8 8 5 6 16 14 9

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE6008

Examples

```
data(GSE6008_eset)
## maybe str(GSE6008_eset) ; plot(GSE6008_eset) ...
```

GSE6822_eset *Classification of ovarian tumor samples*

Description

We have previously described differences in the expression profiles of tumors by comparing tumors of two known classes: low malignant potential versus highly malignant (invasive). This report presents an effort to discover unknown classes within a large heterogeneous set of ovarian tumors, using unsupervised learning approaches. We were able to define four classes in a set of 74 ovarian cancer tumors. Groups identified as A1 and A2 were closely related and correlated with the invasive pathology class while the group identified as B2 was correlated with low malignant potential tumors; group B1 consisted of a mixture of low potential and invasive samples. We selected characteristic candidate genes, which were validated by quantitative PCR and by comparison to other published studies.

Usage

```
data(GSE6822_eset)
```

Format

Authors: Novak JP, Ouellet V, Le Page C, Martinu K, Ponton A, Bachvarow DD, Filali-Mouhim A, Provencher DM, Tonin PN, Hudson TJ, Mes-Masson A.

Lab: McGill University and Genome Quebec Innovation Centre.

Contact: jaroslav.novak@mail.mcgill.ca, jaroslav.novak@gmail.com

Title: Classification of ovarian tumor samples.

Url: http://www.genomequebec.mcgill.ca/ovarian

Details

Platform used: [Hu6800] Affymetrix Human Full Length HuGeneFL Array.

rma normalization used from the rma bioconductor package.

assayData: 5528 features, 66 samples.

GEO_platform_accession: GSE6822

platform_summary: GPL80

biomart_ID: Hu6800 affy_hugenefl

Available sample meta-data:

---

alt_sample_name: Length = 66. Class = character. Mode = character.

sample_type: tumor = 66.

subtype: Length = 66. Class = character. Mode = character.

primarysite: ov = 66.

summarygrade: high = 40. low = 15. NA's = 11.

G: 1 = 1. 2 = 14. 3 = 40. NA's = 11.

debulking: unknown = 66.

batch: 2000-12-21 2001-05-03 2001-05-29 2001-06-12 2001-09-25 2001-09-26 2001-09-27

1 1 3 3 1 5 8

2002-02-14 2002-04-17 2002-04-18 2002-07-18 2002-07-24 2002-10-20 2002-10-30

4 1 9 7 4 10 5

2002-11-01 2002-11-13 2 2

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE6822

Examples

```
data(GSE6822_eset)
## maybe str(GSE6822_eset) ; plot(GSE6822_eset) ...
```

GSE9891_eset *Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome.*

Description

The study aim to identify novel molecular subtypes of ovarian cancer by gene expression profiling with linkage to clinical and pathologic features.

Usage

```
data(GSE9891_eset)
```

Format

Authors: Tothill RW, Tinker AV, George J, Brown R, Fox SB, Lade S, Johnson DS, Trivett MK, Etemadmoghadam D, Locandro B, Traficante N, Fereday S, Hung JA, Chiew YE, Haviv I, Gertig D, DeFazio A, Bowtell DD.

Lab: Peter MacCallum Cancer Center, University of Melbourne, Melbourne, Australia.

Contact: d.bowtell@petermac.org

Title: Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome.

Url: http://pubget.com/search?q=18698038

PubMedID: 18698038

Details

Platform used: [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array.

frma normalization used from the frma bioconductor package.

assayData: 18708 features, 285 samples.

GEO_platform_accession: GSE9891

platform_summary: GPL570

biomart_ID: HG-U133_Plus_2 affy_hg_u133_plus_2

Overall survival time-to-event summary (in years):

Call: survfit(formula = Surv(time, cens) ~ -1)

10 observations deleted due to missingness

records = 275.00 n.max = 275.00 n.start = 275.00 events = 113.00 median = 3.95 0.95LCL = 3.53 0.95UCL = 5.01

Available sample meta-data:

---

alt_sample_name: Length = 285. Class = character. Mode = character.

sample_type: tumor = 285.

subtype: endo = 20. other = 1. ser = 264.

primarysite: ft = 8. other = 34. ov = 243.

arrayedsite: ft = 2. other = 83. ov = 200.

summarygrade: high = 163. low = 116. NA's = 6.

summarystage: early = 42. late = 240. NA's = 3.

tumorstage: 1 = 24. 2 = 18. 3 = 218. 4 = 22. NA's = 3.

substage: a = 26. b = 19. c = 212. NA's = 28.

G: 1 = 19. 2 = 97. 3 = 163. NA's = 6.

age_at_initial_pathologic_diagnosis: Min. = 22.00 1st Qu. = 53.00 Median = 59.00 Mean = 59.62 3rd Qu. = 68.00 Max. = 80.00 NA's = 3.00

pltx: n = 39. y = 243. NA's = 3.

tax: n = 87. y = 195. NA's = 3.

neo: n = 264. y = 18. NA's = 3.

days_to_tumor_recurrence: Min. = 0.0 1st Qu. = 300.0 Median = 450.0 Mean = 621.6 3rd Qu. = 810.0 Max. = 4980.0 NA's = 7.0

recurrence_status: norecurrence = 94. recurrence = 188. NA's = 3.

days_to_death: Min. = 0.0 1st Qu. = 555.0 Median = 870.0 Mean = 956.1 3rd Qu. = 1245.0 Max. = 6420.0 NA's = 10.0

vital_status: deceased = 113. living = 169. NA's = 3.

debulking: optimal = 160. suboptimal = 88. unknown = 37.

batch:

2004-12-03 2004-12-23 2005-01-12 2005-01-17 2005-01-24 2005-01-31 2005-02-21

3 4 7 7 8 10 10

2005-03-17 2005-05-05 2005-05-09 2005-05-25 2005-05-27 2005-05-30 2005-06-02

2 1 1 2 3 3 6

2005-06-06 2005-06-08 2005-06-16 2005-06-17 2005-06-24 2005-07-06 2005-07-15

4 5 3 5 6 2 9

2005-07-20 2005-07-29 2005-08-03 2005-08-05 2005-08-18 2005-08-24 2005-08-26

7 5 6 3 4 8 4

2005-09-09 2005-09-14 2005-09-16 2005-09-21 2005-10-05 2005-10-26 2005-10-28

4 6 6 4 5 2 4

2005-11-04 2005-11-09 2005-11-11 2005-11-23 2005-12-15 2005-12-21 2006-01-20

6 3 7 4 7 8 3

2006-01-31 2006-02-08 2006-02-28 2006-04-05 2006-04-06 2006-04-12 2006-04-13

7 3 3 7 3 7 4

2006-04-28 2006-05-03 2006-06-06 2006-06-07 2006-06-22 2006-07-07 2006-07-19

6 9 6 3 9 4 7

Source

http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE9891

Examples

```
data(GSE9891_eset)
if(require(affy)){
  summary(GSE9891_eset$recurrence_status)
  time <- GSE9891_eset$days_to_death / 365
  cens <- ifelse(GSE9891_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  fit          #for summary of survival
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
}
```

PMID15897565_eset *Patterns of gene expression that characterize long-term survival in advanced stage serous ovarian cancers.*

Description

A better understanding of the underlying biology of invasive serous ovarian cancer is critical for the development of early detection strategies and new therapeutics. The objective of this study was to define gene expression patterns associated with favorable survival.

Usage

```
data(PMID15897565_eset)
```

Format

Authors: Berchuck A, Iversen ES, Lancaster JM, Pittman J, Luo J, Lee P, Murphy S, Dressman HK, Febbo PG, West M, Nevins JR, Marks JR.

Lab: Department of Obstetrics and Gynecology/Division of Gynecologic Oncology, Institute of Statistics and Decision Sciences, Center for Applied Genomics and Technology, Duke University Medical Center, Durham, North Carolina, USA.

Contact: berch001@mc.duke.edu

Title: Patterns of gene expression that characterize long-term survival in advanced stage serous ovarian cancers.

Url: http://pubget.com/search?q=15897565

PubMedID: 15897565

Details

Platform used: HG-U133A.

frma normalization used from the frma bioconductor package.

assayData: 12858 features, 63 samples.

GEO_platform_accession: PMID15897565

platform_summary: GPL96

biomart_ID: HG-U133A affy_hg_u133a

Binary overall survival summary (definitions of long and short provided by study authors): (Long is when survival > 7 yrs); (Short is when survival < 3 years)

long = 24. short = 28. NA's = 11.

Available sample meta-data:

---

alt_sample_name: Min. = 1761 1st Qu. = 1828 Median = 1907 Mean = 2001 3rd Qu. = 2032 Max. = 2536

primarysite: ov = 63.

summarygrade: high = 25. low = 37. NA's = 1.

summarystage: early = 11. late = 52.

tumorstage: 1 = 7. 2 = 4. 3 = 48. 4 = 4.

G: 1 = 2. 2 = 35. 3 = 24. 4 = 1. NA's = 1.

age_at_initial_pathologic_diagnosis: Min. = 33.00 1st Qu. = 52.50 Median = 59.00 Mean = 59.21 3rd Qu. = 67.00 Max. = 79.00

days_to_death: 1092 = 28. 2555 = 24. NA's = 11.

os_binary: long = 24. short = 28. NA's = 11.

debulking: optimal = 24. suboptimal = 28. unknown = 11.

batch:

2002-09-20 2002-10-23 2002-11-12 2002-12-16 2002-12-21 2003-01-03 2003-05-30

15 9 10 1 3 11 13

2003-07-02

1

Examples

```
data(PMID15897565_eset)
## maybe str(PMID15897565_eset) ; plot(PMID15897565_eset) ...
#ISSUE: NO vital_status...

if(require(affy)){
  summary(PMID15897565_eset$os_binary)
}
```

`PMID17290060_eset` *An integrated genomic-based approach to individualized treatment of patients with advanced-stage ovarian cancer.*

Description

The purpose of this study was to develop an integrated genomic-based approach to personalized treatment of patients with advanced-stage ovarian cancer. We have used gene expression profiles to identify patients likely to be resistant to primary platinum-based chemotherapy and also to identify alternate targeted therapeutic options for patients with de novo platinum-resistant disease.

Usage

```
data(PMID17290060_eset)
```

Format

Authors: Dressman HK, Berchuck A, Chan G, Zhai J, Bild A, Sayer R, Cragun J, Clarke J, Whitaker RS, Li L, Gray J, Marks J, Ginsburg GS, Potti A, West M, Nevins JR, Lancaster JM.

Lab: Division of Gynecologic Surgical Oncology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL 33612, USA.

Contact: lancasjm@moffitt.usf.edu

Title: An integrated genomic-based approach to individualized treatment of patients with advanced-stage ovarian cancer.

Url: http://pubget.com/search?q=17290060

PubMedID: 17290060

Details

Platform used: HG-U133A.

frma normalization used from the frma bioconductor package.

assayData: 12858 features, 117 samples.

GEO_platform_accession: PMID17290060 platform_summary: GPL96 biomart_ID: HG-U133A affy_hg_u133a

Overall survival time-to-event summary (in years):

Call: survfit(formula = Surv(time, cens) ~ -1)

records = 117.00 n.max = 117.00 n.start = 117.00 events = 67.00 median = 5.26 0.95LCL = 2.79 0.95UCL = 7.48

Available sample meta-data:

---

alt_sample_name: Length = 117. Class = character. Mode = character.

primarysite: ov = 117.

summarygrade: high = 57. low = 57. NA's = 3.

summarystage: early = 1. late = 115. NA's = 1.

tumorstage: 2 = 1. 3 = 98. 4 = 17. NA's = 1.

G: 1 = 4. 2 = 53. 3 = 56. 4 = 1. NA's = 3.

days_to_death: Min. = 30. 1st Qu. = 510. Median = 1020. Mean = 1496. 3rd Qu. = 2220. Max. = 5550.

vital_status: deceased = 67. living = 50.

primary_therapy_outcome_success: completeresponse = 85. progressivedisease = 32.

debulking: optimal = 63. suboptimal = 54.

Examples

```
data(PMID17290060_eset)
## maybe str(PMID17290060_eset) ; plot(PMID17290060_eset) ...

##ISSUE: only "primary_therapy_outcome_success"

if(require(affy)){
  summary(PMID17290060_eset$primary_therapy_outcome_success)
  time <- PMID17290060_eset$days_to_death / 365
  cens <- ifelse(PMID17290060_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  fit          #for summary of survival
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
}
```

PMID19318476_eset *Microarray analysis of early stage serous ovarian cancers shows profiles predictive of favorable outcome.*

Description

Although few women with advanced serous ovarian cancer are cured, detection of the disease at an early stage is associated with a much higher likelihood of survival. We previously used gene expression array analysis to distinguish subsets of advanced cancers based on disease outcome. In the present study, we report on gene expression of early-stage cancers and validate our prognostic model for advanced-stage cancers.

Usage

```
data(PMID19318476_eset)
```

Format

Authors: Berchuck A, Iversen ES, Luo J, Clarke JP, Horne H, Levine DA, Boyd J, Alonso MA, Secord AA, Bernardini MQ, Barnett JC, Boren T, Murphy SK, Dressman HK, Marks JR, Lancaster JM.

Lab: Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, Institute for Genome Sciences and Policy, Duke University Medical Center, Durham, North Carolina 27710, USA.

Contact: berch001@mc.duke.edu

Title: Microarray analysis of early stage serous ovarian cancers shows profiles predictive of favorable outcome.

Url: http://pubget.com/search?q=19318476

PubMedID: 19318476

Details

Platform used: HG-U133A.

frma normalization used from the frma bioconductor package.

assayData: 12858 features, 42 samples.

GEO_platform_accession: PMID19318476

platform_summary: GPL96

biomart_ID: HG-U133A affy_hg_u133a

Overall survival time-to-event summary (in years):

Call: survfit(formula = Surv(time, cens) ~ -1)

records = 42.00 n.max = 42.00 n.start = 42.00 events = 22.00 median = 2.79 0.95LCL = 2.30 0.95UCL = NA

Available sample meta-data:

---

alt_sample_name: Length = 42. Class = character. Mode = character.

summarygrade: high = 24. low = 17. NA's = 1.

summarystage: early = 2. late = 39. NA's = 1.

tumorstage: 1 = 1. 2 = 1. 3 = 29. 4 = 10. NA's = 1.

substage: a = 1. b = 1. c = 29. NA's = 11.

G: 1 = 2. 2 = 15. 3 = 24. NA's = 1.

age_at_initial_pathologic_diagnosis: Min. = 33.00 1st Qu. = 55.00 Median = 62.00 Mean = 61.46 3rd Qu. = 70.00 Max. = 81.00 NA's: 1.00

recurrence_status: norecurrence = 6. recurrence = 36.

days_to_death: Min. = 30.0 1st Qu. = 367.5 Median = 825.0 Mean = 1105.0 3rd Qu. = 1050.0 Max. = 3420.0

vital_status: deceased = 22. living = 20.

debulking: optimal = 20. suboptimal = 21. NA's = 1.

batch:

2004-03-09 2004-03-16 2004-04-20 2004-05-18 2004-05-21 2004-05-27 2004-06-22

14 3 4 8 6 5 1

2004-06-23

1

Examples

```
data(PMID19318476_eset)
## maybe str(PMID19318476_eset) ; plot(PMID19318476_eset) ...

if(require(affy)){
  summary(PMID19318476_eset$recurrence_status)
```

```
  time <- PMID19318476_eset$days_to_death / 365
  cens <- ifelse(PMID19318476_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  fit          #for summary of survival
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
}
```

TCGA.mirna.8x15kv2_eset

Ovarian serous cystadenocarcinoma: Sample Counts and Findings

---

Description

The Cancer Genome Atlas

Usage

```
data(TCGA.mirna.8x15kv2_eset)
```

Format

Authors: The Cancer Genome Atlas.

Details

```
Level 3 (fully-processed) microRNA data downloaded from TCGA.

assayData: 799 features, 554 samples

Overall survival time-to-event summary (in years):
Call: survfit(formula = Surv(time, cens) ~ -1)

   10 observations deleted due to missingness
records   n.max n.start  events  median 0.95LCL 0.95UCL
 544.00  544.00  544.00  286.00    3.71    3.42    4.03

---------------------------
Available sample meta-data:
---------------------------

alt_sample_name:
   Length     Class      Mode
      554 character character
```

```
sample_type:
tumor
  554

subtype:
ser
554

primarysite:
other    ov
    4   550

summarygrade:
high  low NA's
 474   68   12

summarystage:
early  late  NA's
   39   511     4

tumorstage:
   1    2    3    4 NA's
  16   23  426   85    4

substage:
   b    c NA's
  31  434   89

G:
   1    2    3    4 NA's
   6   62  473    1   12

age_at_initial_pathologic_diagnosis:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  26.00   51.00   59.00   59.81   69.00   89.00

days_to_tumor_recurrence:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
    8.0   235.2   436.0   618.7   797.0  5480.0      44

recurrence_status:
norecurrence   recurrence
         268          286

days_to_death:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
    8.0   346.0   867.5   997.7  1446.0  5480.0      10
```

```
vital_status:
deceased   living     NA's
     286      261        7

site_of_tumor_first_recurrence:
               locoregional locoregional_plus_metastatic
                        145                           3
                 metastasis                       NA's
                        138                         268

primary_therapy_outcome_success:
  completeresponse    partialresponse progressivedisease      stabledisease
               308                 63                 41                 30
              NA's
               112

debulking:
   optimal suboptimal       NA's
       112        384         58
```

Examples

```
data(TCGA.mirna.8x15kv2_eset)
```

TCGA_eset *Ovarian serous cystadenocarcinoma: Sample Counts and Findings*

Description

The Cancer Genome Atlas

Usage

```
data(TCGA_eset)
```

Format

Authors: The Cancer Genome Atlas.

Title: Ovarian serous cystadenocarcinoma: Sample Counts and Findings.

Url: http://tcga-data.nci.nih.gov/tcga/tcgaCancerDetails.jsp?diseaseType=OV&diseaseName=Ovarian%20serous%20cystad

Details

Note that the TCGA dataset contains some samples replicated across batches, but there are no clinical metadata associated with these samples. The following provides these sample pairs and the Pearson correlation of their RMA-normalized expression profiles:

name1	name2	pearson correlation
TCGA.24.2290	TCGA.13.2071	0.9969045
TCGA.25.2392	TCGA.29.2432	0.9976206
TCGA.25.2404	TCGA.36.2537	0.9978256
TCGA.24.2262	TCGA.42.2591	0.9967085
TCGA.24.2293	TCGA.23.2641	0.997569
TCGA.25.2393	TCGA.29.2434	0.996287
TCGA.59.2350	TCGA.36.2548	0.9968095
TCGA.24.2295	TCGA.23.2643	0.9952585
TCGA.09.2048	TCGA.04.1644	0.9967052
TCGA.24.2271	TCGA.13.2057	0.9969408
TCGA.25.2397	TCGA.61.2612	0.9957279
TCGA.09.2049	TCGA.42.2582	0.9932573
TCGA.24.2280	TCGA.13.2059	0.9937523
TCGA.24.2298	TCGA.23.2647	0.9970841
TCGA.25.2398	TCGA.36.2530	0.996212
TCGA.09.2051	TCGA.42.2588	0.9971348
TCGA.24.2288	TCGA.13.2065	0.9954178
TCGA.29.2428	TCGA.36.2544	0.9964141
TCGA.24.2289	TCGA.13.2066	0.9968514
TCGA.36.2529	TCGA.29.1704	0.9976284
TCGA.36.2532	TCGA.24.1852	0.997482

rma normalization used from the rma bioconductor package.

assayData: 12858 features, 570 samples.

GEO_platform_accession: TCGA

platform_summary: GPL4685

biomart_ID: HT-HG_U133A affy_hg_u133a

Overall survival time-to-event summary (in years):

Call: survfit(formula = Surv(time, cens) ~ -1)

records = 576. n.max = 576. n.start = 576. events = 0. median = NA. 0.95LCL = NA. 0.95UCL = NA.

Available sample meta-data:

---

alt_sample_name: Length = 576. Class = character. Mode = character.

unique_patient_ID: Length = 576. Class = character. Mode = character.

sample_type: tumor = 576.

subtype: ser = 576.

primarysite: ov = 576.

G: 1 = 6. 2 = 68. 3 = 485. 4 = 1.

days_to_death: Min. = 8 Median = 1024 Mean = 1083.22 Max. = 4623

batch: Length = 570. Class = character. Mode = character.

Examples

```
data(TCGA_eset)
## maybe str(TCGA_eset) ; plot(TCGA_eset) ...

if(require(affy)){
  summary(TCGA_eset$recurrence_status)
  time <- TCGA_eset$days_to_death / 365
  cens <- ifelse(TCGA_eset$vital_status=="deceased",1,0)
  library(survival)
  fit <- survfit(Surv(time,cens)~-1)
  fit          #for summary of survival
  summary(fit)
  plot(fit,xlab="Time (years)",ylab="Survivor function")
  inverse.fit <- survfit(Surv(time,1-cens) ~ -1)
  inverse.fit  #for summary of follow-up time
}
```

Index