

# Package ‘curatedOvarianData’

October 8, 2013

**Type** Package

**Title** Clinically Annotated Data for the Ovarian Cancer Transcriptome

**Version** 1.0.3

**Date** 2013-02-22

**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, futile.logger, sva,xtable

**License** Artistic-2.0

**URL** <http://bcb.dfci.harvard.edu/ovariancancer>

**biocViews** ExperimentData, Cancer, Ovarian, RNAExpressionData

## R topics documented:

curatedOvarianData-package . . . . .	2
E.MTAB.386_eset . . . . .	4
GSE12418_eset . . . . .	7
GSE12470_eset . . . . .	9
GSE13876_eset . . . . .	11
GSE14764_eset . . . . .	14
GSE17260_eset . . . . .	17
GSE18520_eset . . . . .	20

GSE19829.GPL570_eset . . . . .	23
GSE19829.GPL8300_eset . . . . .	25
GSE20565_eset . . . . .	27
GSE2109_eset . . . . .	30
GSE26712_eset . . . . .	32
GSE30009_eset . . . . .	36
GSE30161_eset . . . . .	38
GSE32062.GPL6480_eset . . . . .	42
GSE32063_eset . . . . .	44
GSE6008_eset . . . . .	47
GSE6822_eset . . . . .	50
GSE9891_eset . . . . .	52
PMID15897565_eset . . . . .	55
PMID17290060_eset . . . . .	58
PMID19318476_eset . . . . .	61
TCGA.mirna.8x15kv2_eset . . . . .	64
TCGA_eset . . . . .	68

**Index** **73**

curatedOvarianData-package

*Clinically Annotated Data for the Ovarian Cancer Transcriptome*

**Description**

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

**Details**

Package: curatedOvarianData  
 Type: Package  
 Version: 1.0.3  
 Date: 2013-02-22  
 License: Artistic-2.0  
 Depends: R (>= 2.10.0), affy

Please see <http://bcf.dfci.harvard.edu/ovariancancer> for alternative versions of this package, differing in how redundant probe sets are dealt with. In the curatedOvarianData version, each gene is represented by the gene with maximum mean. In NormalizerVcuratedOvarianData, each gene is represented by the mean of the probesets after removing "noisy" probesets (see the Normalizer function of the Sleipnir library for computational biology), and in FULLVcuratedOvarianData, no collapsing of probe sets is done, but a map is provided to allow the user to do so by their method of choice through featureData(eset).

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

For "sample\_type": tumor / metastatic / adjacentnormal / healthy / cellline: "healthy" should be only from individuals without cancer, "adjacentnormal" from individuals with cancer, "metastatic" for non-primary tumors.

For "histological\_type": ser=serous / endo=endometrioid / clearcell / mucinous, undifferentiated / other / mix. Other includes sarcomatoid, adenocarcinoma, dysgerminoma.

For "primarysite" and for "arrayedsite": ovlftlother. 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 "substage": substage (abcd). For cases like ab, bc, use highest given.

For "grade": Grade (1-3): If multiple given, ie 12, 23, use highest given. Most ovarian cancer studies use FIGO grading, with a couple exceptions in this package (Yoshihara and Tothill).

For "pltx": (y/n): patient treated with platin.

For "tax": (y/n): patient treated with taxol.

For "neo": (y/n): patient treated with neoadjuvant treatment.

For "primary\_therapy\_outcome\_success": completeresponse|partialresponse|progressivedisease|stabledisease: response to any kind of therapy (including radiation only).

For "days\_to\_tumor\_recurrence": time to recurrence or last follow-up in days

For "recurrence\_status": recurrence censoring variable (recurrence / norecurrence)

For "days\_to\_death": time to death or last follow-up in days

For "vital\_status": Overall survival censoring variable (living / deceased)

For "os\_binary": dichotomized overall survival variable as defined by study authors (short / long).

For "relapse\_binary": dichotomized relapse variable as defined by study authors (short / long)

For "site\_of\_tumor\_first\_recurrence": (metastasis / locoregional / none / locoregional\_plus\_metastatic). none for no recurrence, na for unknown

For "primary\_therapy\_outcome\_success": (completeresponse / partialresponse / progressivedisease / stabledisease) Response to any kind of therapy (including radiation only).

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

For "percent\_normal\_cells": Estimated percentage of normal cells. An integer 0-100, or can be >70, <70, etc.

For "percent\_stromal\_cells": Estimated percentage of stromal cells. An integer 0-100, or can be >70, <70, etc.

For "percent\_tumor\_cells": Estimated percentage of tumor cells. An integer 0-100, or can be >70, <70, etc.

For "batch": batch variable when available. Hybridization date when Affymetrix CEL files are available.

For "uncurated\_author\_metadata": Original uncurated data, with each field separated by ///.

**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")
##
##See the actual template used for syntax checking of clinical metadata:
template.file <- system.file("extdata/template_ov.csv", package = "curatedOvarianData")
template <- read.csv(template.file, as.is=TRUE)
head(template)
```

---

E.MTAB.386\_eset

*Angiogenic mRNA and microRNA gene expression signature predicts a novel subtype of serous ovarian cancer.*

---

**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 patients eventually develop recurrent cancer and succumb to increasingly platinum-resistant disease. Modern, targeted cancer drugs intervene in cell signaling, and identifying key disease mechanisms and pathways would greatly advance our treatment abilities. In order to shed light on the molecular diversity of ovarian cancer, we performed comprehensive transcriptional profiling on 129 advanced stage, high grade serous ovarian cancers. We implemented a, 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 previously undescribed patient stratification, further supported by micro-RNA expression profiles, and gene set enrichment analysis found strong biological support for the stratification by extracellular matrix, cell adhesion, and angiogenesis genes. The corresponding "angiogenesis signature" was validated in ten published independent ovarian cancer gene expression datasets and is significantly associated with overall survival. The subtypes we have defined are of potential translational interest as they may be relevant for identifying patients who may benefit from the addition of anti-angiogenic therapies that are now being tested in clinical trials.

**Usage**

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

**Format**

```

experimentData(eset):
Experiment data
  Experimenter name: Bentink S, Haibe-Kains B, Risch T, Fan J-B, Hirsch MS, Holton K, Rubio R, April C, Ch
  Laboratory: Bentink, Matulonis 2012
  Contact information:
  Title: Angiogenic mRNA and microRNA gene expression signature predicts a novel subtype of serous ovaria
  URL:
  PMIDs: 22348002

  Abstract: A 212 word abstract is available. Use 'abstract' method.
  Information is available on: preprocessing
  notes:
  platform_title:
    Illumina humanRef-8 v2.0 expression beadchip
  platform_shorttitle:
    Illumina humanRef-8 v2.0
  platform_summary:
    illuminaHumanv2
  platform_manufacturer:
    Illumina
  platform_distribution:
    commercial
  platform_accession:
    GPL6104

Preprocessing: default
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A2M A4GALT ... ZZEF1 (10327 total)
  varLabels: probeset gene
  varMetadata: labelDescription

```

**Details**

```

assayData: 10327 features, 129 samples
Platform type: illuminaHumanv2
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
 129.00 129.00 129.00   73.00   3.51   2.68   4.13

```

```

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

```

```

unique_patient_ID:
  Length      Class      Mode
  129 character character

sample_type:
tumor
  129

histological_type:
ser
  129

primarysite:
ov
  129

summarygrade:
high
  129

summarystage:
early late
  1  128

tumorstage:
  2  3  4
  1 109 19

substage:
  a  b  c NA's
  5 12 93 19

age_at_initial_pathologic_diagnosis:
  Min. 1st Qu.  Median    Mean 3rd Qu.  Max.
  21.00  50.00  66.00  60.71  72.00  95.00

days_to_death:
  Min. 1st Qu.  Median    Mean 3rd Qu.  Max.
  3.9  516.9  917.1  1007.0  1401.0  2724.0

vital_status:
deceased  living
  73      56

debulking:
  optimal suboptimal  NA's
  98      28          3

```

```

uncurated_author_metadata:
  Length      Class      Mode
      129 character character

```

---

GSE12418_eset	<i>Expression analysis of stage III serous ovarian adenocarcinoma distinguishes a sub-group of survivors.</i>
---------------	---

---

## Description

It is difficult to predict the clinical outcome for patients with ovarian cancer. However, the use of biomarkers as additional prognostic factors may improve the outcome for these patients. In order to find novel candidate biomarkers, differences in gene expressions were analysed 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. Using hierarchical cluster analysis we detected a sub-group that included 60% of the survivors. The gene expressions in tumours from patients in this sub-group of survivors were compared with the remaining tumours, and 204 genes were found to be differently expressed. We conclude that the sub-group of survivors might represent patients with favourable tumour biology and sensitivity to treatment. A selection of the 204 genes might be used as a predictive model to distinguish patients within and outside of this group. Alternative chemotherapy strategies could then be offered as first-line treatment, which may lead to improvements in the clinical outcome for these patients.

## Usage

```
data( GSE12418_eset )
```

## Format

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Partheen K, Levan K, Osterberg L, Horvath G.Expression analysis of stage III serous
```

```
  Laboratory: Partheen, Horvath 2006
```

```
  Contact information:
```

```
  Title: Expression analysis of stage III serous ovarian adenocarcinoma distinguishes a sub-group of sur
```

```
  URL:
```

```
  PMIDs: 16996261
```

```
  Abstract: A 177 word abstract is available. Use 'abstract' method.
```

```
  Information is available on: preprocessing
```

```
  notes:
```

```
  platform_title:
```

```
    SWEGENE H_v2.1.1_27k
```

```
  platform_shorttitle:
```

```
    SWEGENE H_v2.1.1_27k
```

```

platform_summary:
  PartheenMetaData
platform_manufacturer:
  other
platform_distribution:
  non-commercial
platform_accession:
  GPL5886

```

```

Preprocessing: default
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A1CF A2LD1 ... ZZZ3 (12656 total)
  varLabels: probeset gene
  varMetadata: labelDescription

```

### Details

```

assayData: 12656 features, 54 samples
Platform type: PartheenMetaData
Binary overall survival summary (definitions of long and short provided by study authors):

```

```

  long short
    20    34

```

```

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

```

```

alt_sample_name:
  Length      Class      Mode
    54 character character

```

```

sample_type:
tumor
    54

```

```

histological_type:
ser
    54

```

```

primarysite:
ov
    54

```

```

summarystage:
late
    54

```



```

tumorstage:
  3
  54

substage:
  b c
  19 35

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

pltx:
  y
  54

os_binary:
  long short
  20    34

debulking:
  optimal suboptimal
  13         41

uncurated_author_metadata:
  Length    Class      Mode
  54 character character

```

GSE12470\_eset

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

## Description

To elucidate the mechanisms of rapid progression of serous ovarian cancer, gene expression profiles from 43 ovarian cancer tissues comprising eight early stage and 35 advanced stage tissues were carried out using oligonucleotide microarrays of 18,716 genes. By non-negative matrix factorization analysis using 178 genes, which were extracted as stage-specific genes, 35 advanced stage cases were classified into two subclasses with superior (n = 17) and poor (n = 18) outcome evaluated by progression-free survival (log rank test, P = 0.03). Of the 178 stage-specific genes, 112 genes were identified as showing different expression between the two subclasses. Of the 48 genes selected for biological function by gene ontology analysis or Ingenuity Pathway Analysis, five genes (ZEB2, CDH1, LTBP2, COL16A1, and ACTA2) were extracted as candidates for prognostic factors associated with progression-free survival. The relationship between high ZEB2 or low CDH1

expression and shorter progression-free survival was validated by real-time RT-PCR experiments of 37 independent advanced stage cancer samples. ZEB2 expression was negatively correlated with CDH1 expression in advanced stage samples, whereas ZEB2 knockdown in ovarian adenocarcinoma SKOV3 cells resulted in an increase in CDH1 expression. Multivariate analysis showed that high ZEB2 expression was independently associated with poor prognosis. Furthermore, the prognostic effect of E-cadherin encoded by CDH1 was verified using immunohistochemical analysis of an independent advanced stage cancer samples set (n = 74). These findings suggest that the expression 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

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Yoshihara K, Tajima A, Komata D, Yamamoto T, Kodama S, Fujiwara H, Suzuki M, Onishi Y
```

```
  Laboratory: Yoshihara, Tanaka 2009
```

```
  Contact information:
```

```
  Title: Gene expression profiling of advanced-stage serous ovarian cancers distinguishes novel subclasses
```

```
  URL:
```

```
  PMIDs: 19486012
```

```
  Abstract: A 253 word abstract is available. Use 'abstract' method.
```

```
  Information is available on: preprocessing
```

```
  notes:
```

```
    platform_title:
```

```
      Agilent-012097 Human 1A Microarray (V2) G4110B (Feature Number version)
```

```
    platform_shorttitle:
```

```
      Agilent G4110B
```

```
    platform_summary:
```

```
      hgug4110b
```

```
    platform_manufacturer:
```

```
      Agilent
```

```
    platform_distribution:
```

```
      commercial
```

```
    platform_accession:
```

```
      GPL887
```

```
Preprocessing: default
```

```
featureData(eset):
```

```
An object of class 'AnnotatedDataFrame'
```

```
  featureNames: A1BG A1CF ... ZZZ3 (16880 total)
```

```
  varLabels: probeset gene
```

```
  varMetadata: labelDescription
```

**Details**

assayData: 16880 features, 53 samples

Platform type: hgug4110b

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

alt\_sample\_name:

Length	Class	Mode
53	character	character

sample\_type:

healthy	tumor
10	43

histological\_type:

ser NA's
43 10

primarysite:

ov
53

summarystage:

early	late	NA's
8	35	10

tumorstage:

1 NA's
8 45

uncurated\_author\_metadata:

Length	Class	Mode
53	character	character

---

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. According to a randomized design, profiling of 157 advanced stage serous ovarian cancers was performed in duplicate using approximately 35,000 70-mer oligonucleotide microarrays. A continuous predictor of overall survival was built taking into account well-known issues in microarray analysis, such as multiple testing and overfitting. A functional class scoring analysis was utilized to assess pathways/transcription factors for their association with overall survival. The prognostic value of genes that constitute our overall survival profile was validated on a fully independent, publicly available dataset of 118 well-defined primary serous ovarian cancers. Furthermore, functional class scoring analysis was also performed on this independent dataset to assess the similarities with results from our own dataset. An 86-gene overall survival profile discriminated between patients with unfavorable and favorable prognosis (median survival, 19 versus 41 mo, respectively; permutation p-value of log-rank statistic = 0.015) and maintained its independent prognostic value in multivariate analysis. Genes that composed the overall survival profile were also able to discriminate between the two risk groups in the independent dataset. In our dataset 17/167 pathways and 13/111 transcription factors were associated with overall survival, of which 16 and 12, respectively, were confirmed in the independent dataset. Our study provides new clues to genes, pathways, and transcription factors that contribute to the clinical outcome of serous ovarian cancer and might be exploited in designing new treatment strategies.

### Usage

```
data( GSE13876_eset )
```

### Format

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Crijns AP, Fehrmann RS, de Jong S, Gerbens F, Meersma GJ, Klip HG, Hollema H, Hofstra
```

```
  Laboratory: Crijns, van der Zee 2009
```

```
  Contact information:
```

```
  Title: Survival-related profile, pathways, and transcription factors in ovarian cancer.
```

```
  URL:
```

```
  PMIDs: 19192944
```

```
  Abstract: A 371 word abstract is available. Use 'abstract' method.
```

```
  Information is available on: preprocessing
```

```
  notes:
```

```
    platform_title:
```

```
      Operon human v3 ~35K 70-mer two-color oligonucleotide microarrays
```

```
    platform_shorttitle:
```

```
      Operon v3 two-color
```

```
    platform_summary:
```

```
      OperonHumanV3
```

```
    platform_manufacturer:
```

```

      other
platform_distribution:
  non-commercial
platform_accession:
  GPL7759

Preprocessing: default
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A1BG A1CF ... ZZZ3 (20505 total)
  varLabels: probeset gene
  varMetadata: labelDescription

```

## Details

```

assayData: 20505 features, 157 samples
Platform type: OperonHumanV3
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
 157.00 157.00 157.00 113.00   2.05   1.56   2.71

```

```

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

```

```

alt_sample_name:
  151 NA's
  1 156

```

```

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

```

```

sample_type:
tumor
 157

```

```

histological_type:
ser
157

```

```

primarysite:
ov
157

```

```

summarygrade:

```

```
high low NA's
  85  59  13
```

```
summarystage:
late
  157
```

```
grade:
  1  2  3  4 NA's
 14 45 82  3  13
```

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

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

```
vital_status:
deceased living
   113    44
```

```
uncurated_author_metadata:
  Length Class Mode
   157 character character
```

---

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

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Denkert C, Budczies J, Darb-Esfahani S, Gy??rffy B et al. A prognostic gene expressi
```

```
  Laboratory: Denkert, Lage 2009
```

```
  Contact information:
```

```
  Title: A prognostic gene expression index in ovarian cancer - validation across different independent c
```

```
  URL:
```

```
  PMIDs: 19294737
```

```
  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
```

```
Preprocessing: frma
```

```
featureData(eset):
```

```
An object of class 'AnnotatedDataFrame'
```

```
  featureNames: A1CF A2M ... ZZZ3 (12981 total)
```

```
  varLabels: probeset gene
```

```
  varMetadata: labelDescription
```

## Details

```
assayData: 12981 features, 80 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
80.00	80.00	80.00	21.00	4.52	4.19	NA

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

alt\_sample\_name:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.00	20.75	40.50	40.50	60.25	80.00

sample\_type:

tumor  
80

histological\_type:

Length	Class	Mode
80	character	character

primarysite:

ov  
80

summarygrade:

high	low
54	26

summarystage:

early	late
9	71

tumorstage:

1	2	3	4
8	1	69	2

substage:

a	b	c	NA's
4	6	32	38

grade:

1	2	3
3	23	54

recurrence\_status:

norecurrence	recurrence	NA's
--------------	------------	------



```

                    50          26          4

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

vital_status:
deceased  living
         21     59

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

uncurated_author_metadata:
  Length    Class      Mode
     80 character character

```

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. Advanced-stage serous ovarian cancer tissues from 110 Japanese patients who underwent primary surgery and platinum/taxane-based chemotherapy were profiled using oligonucleotide microarrays. We selected 88 PFS-related genes by a univariate Cox model ( $p < 0.01$ ) and generated the prognostic index based on 88 PFS-related genes after adjustment of regression coefficients of the respective genes by ridge regression Cox model using 10-fold cross-validation. The prognostic index was independently associated with PFS time compared to other clinical factors in multivariate analysis [hazard ratio (HR), 3.72; 95% confidence interval (CI), 2.66-5.43;  $p < 0.0001$ ]. In an external dataset, multivariate analysis revealed that this prognostic index was significantly correlated with PFS time (HR, 1.54; 95% CI, 1.20-1.98;  $p = 0.0008$ ). Furthermore, the correlation between the prognostic index and overall survival time was confirmed in the two independent external datasets (log rank test,  $p = 0.0010$  and  $0.0008$ ). The prognostic ability of our index based on the 88-gene expression profile in ridge regression Cox hazard model was shown to be independent of other clinical factors in predicting cancer prognosis across two distinct datasets. Further study will be necessary to improve predictive accuracy of the

prognostic index toward clinical application for evaluation of the risk of recurrence in patients with advanced-stage serous ovarian cancer.

### Usage

```
data( GSE17260_eset )
```

### Format

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Yoshihara K, Tajima A, Yahata T, Kodama S, Fujiwara H, Suzuki M, Onishi Y, Hatae M, S
```

```
  Laboratory: Yoshihara, Tanaka 2010
```

```
  Contact information:
```

```
  Title: Gene expression profile for predicting survival in advanced-stage serous ovarian cancer across
```

```
  URL:
```

```
  PMIDs: 20300634
```

```
  Abstract: A 257 word abstract is available. Use 'abstract' method.
```

```
  Information is available on: preprocessing
```

```
  notes:
```

```
    platform_title:
```

```
      Agilent-012391 Whole Human Genome Oligo Microarray G4112A
```

```
    platform_shorttitle:
```

```
      Agilent G4112A
```

```
    platform_summary:
```

```
      hgug4112a
```

```
    platform_manufacturer:
```

```
      Agilent
```

```
    platform_distribution:
```

```
      commercial
```

```
    platform_accession:
```

```
      GPL6848
```

```
Preprocessing: default
```

```
featureData(eset):
```

```
An object of class 'AnnotatedDataFrame'
```

```
  featureNames: A1CF A2LD1 ... ZZZ3 (19358 total)
```

```
  varLabels: probeset gene
```

```
  varMetadata: labelDescription
```

### Details

```
assayData: 19358 features, 110 samples
```

```
Platform type: hgug4112a
```

```
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
110.00	110.00	110.00	46.00	4.44	4.03	NA

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

alt\_sample\_name:  
 Length Class Mode  
 110 character character

sample\_type:  
 tumor  
 110

histological\_type:  
 ser  
 110

primarysite:  
 ov  
 110

summarygrade:  
 high low  
 43 67

summarystage:  
 late  
 110

tumorstage:  
 3 4  
 93 17

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

grade:  
 1 2 3  
 26 41 43

pltx:  
 y  
 110

tax:

```

y
110

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

recurrence_status:
norecurrence  recurrence
           34           76

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

vital_status:
deceased  living
        46        64

debulking:
  optimal suboptimal
        57         53

uncurated_author_metadata:
  Length    Class      Mode
    110 character character

```

---

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**

```

experimentData(eset):
Experiment data
  Experimenter name: Mok SC, Bonome T, Vathipadiekal V, Bell A, Johnson ME, Wong KK, Park DC, Hao K, Yip D
  Laboratory: Mok, Birrer 2009
  Contact information:
  Title: A gene signature predictive for outcome in advanced ovarian cancer identifies a survival factor
  URL:
  PMIDs: 19962670

  Abstract: A 110 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|Operon
  platform_distribution:
    commercial|non-commercial
  platform_accession:
    GPL570|GPL9216

Preprocessing: frma
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A1BG A1CF ... ZZZ3 (19093 total)
  varLabels: probeset gene
  varMetadata: labelDescription

```

**Details**

```

assayData: 19093 features, 63 samples
Platform type: hgu133plus2
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
  53.00  53.00  53.00   41.00    2.05   1.48   3.70

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

```

```

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

sample_type:
healthy  tumor
     10    53

histological_type:
ser NA's
  53  10

primarysite:
ov
63

summarygrade:
high NA's
  53  10

summarystage:
late NA's
  53  10

tumorstage:
  3 NA's
  53  10

grade:
  3 NA's
  53  10

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

vital_status:
deceased  living  NA's
     41    12    10

debulking:
optimal
  63

percent_normal_cells:
  0
  63

```

```

percent_stromal_cells:
  0
  63

percent_tumor_cells:
  100
  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

uncurated_author_metadata:
  Length      Class      Mode
    63 character character

```

---

GSE19829.GPL570_eset	<i>Gene expression profile of BRCAness that correlates with responsiveness to chemotherapy and with outcome in patients with epithelial ovarian cancer.</i>
----------------------	---

---

## Description

To define a gene expression profile of BRCAness that correlates with chemotherapy response and outcome in epithelial ovarian cancer (EOC). A publicly available microarray data set including 61 patients with EOC with either sporadic disease or BRCA(1/2) germline mutations was used for development of the BRCAness profile. Correlation with platinum responsiveness was assessed in platinum-sensitive and platinum-resistant tumor biopsy specimens from six patients with BRCA germline mutations. Association with poly-ADP ribose polymerase (PARP) inhibitor responsiveness and with radiation-induced RAD51 foci formation (a surrogate of homologous recombination) was assessed in Capan-1 cell line clones. The BRCAness profile was validated in 70 patients enriched for sporadic disease to assess its association with outcome. The BRCAness profile accurately predicted platinum responsiveness in eight out of 10 patient-derived tumor specimens, and between PARP-inhibitor sensitivity and resistance in four out of four Capan-1 clones. [corrected] When applied to the 70 patients with sporadic disease, patients with the BRCA-like (BL) profile had improved disease-free survival (34 months v 15 months; log-rank P = .013) and overall survival (72 months v 41 months; log-rank P = .006) compared with patients with a non-BRCA-like (NBL) profile, respectively. The BRCAness profile maintained independent prognostic value in multivariate analysis, which controlled for other known clinical prognostic factors. The BRCAness profile correlates with responsiveness to platinum and PARP inhibitors and identifies a subset of sporadic patients with improved outcome. Additional evaluation of this profile as a predictive tool in patients with sporadic EOC is warranted.

## Usage

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

**Format**

```

experimentData(eset):
Experiment data
  Experimenter name: Konstantinopoulos PA, Spentzos D, Karlan BY, Taniguchi T et al. Gene expression profile of BRCAness that correlates with responsiveness to chemotherapy and with overall survival
  Laboratory: Konstantinopoulos, Cannistra 2010 hgu133plus2
  Contact information:
  Title: Gene expression profile of BRCAness that correlates with responsiveness to chemotherapy and with overall survival
  URL:
  PMIDs: 20547991

Abstract: A 241 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|GPL8300

Preprocessing: frma
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A1BG A1CF ... ZZZ3 (19093 total)
  varLabels: probeset gene
  varMetadata: labelDescription

```

**Details**

```

assayData: 19093 features, 28 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
   28.00  28.00  28.00   17.00   3.95    2.71    NA

```

```

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

```



```

alt_sample_name:
  Length      Class      Mode
    28 character character

sample_type:
tumor
  28

primarysite:
ov
  28

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

vital_status:
deceased  living
    17     11

batch:
2009-08-14
  28

uncurated_author_metadata:
  Length      Class      Mode
    28 character character

```

---

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). A publicly available microarray data set including 61 patients with EOC with either sporadic disease or BRCA(1/2) germline mutations was used for development of the BRCAness profile. Correlation with platinum responsiveness was assessed in platinum-sensitive and platinum-resistant tumor biopsy specimens from six patients with BRCA germline mutations. Association with poly-ADP ribose polymerase (PARP) inhibitor responsiveness and with radiation-induced RAD51 foci formation (a surrogate of homologous recombination) was assessed in Capan-1 cell line clones. The BRCAness profile was validated in 70 patients enriched for sporadic disease to assess its association with outcome. The BRCAness profile accurately predicted platinum responsiveness in eight out of 10 patient-derived tumor specimens, and between PARP-inhibitor sensitivity and resistance in four out of four Capan-1 clones. [corrected] When

applied to the 70 patients with sporadic disease, patients with the BRCA-like (BL) profile had improved disease-free survival (34 months v 15 months; log-rank  $P = .013$ ) and overall survival (72 months v 41 months; log-rank  $P = .006$ ) compared with patients with a non-BRCA-like (NBL) profile, respectively. The BRCAness profile maintained independent prognostic value in multivariate analysis, which controlled for other known clinical prognostic factors. The BRCAness profile correlates with responsiveness to platinum and PARP inhibitors and identifies a subset of sporadic patients with improved outcome. Additional evaluation of this profile as a predictive tool in patients with sporadic EOC is warranted.

## Usage

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

## Format

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Konstantinopoulos PA, Spentzos D, Karlan BY, Taniguchi T et al. Gene expression profile of BRCA-like and non-BRCA-like ovarian cancer
```

```
  Laboratory: Konstantinopoulos, Cannistra 2010 hgu95
```

```
  Contact information:
```

```
  Title: Gene expression profile of BRCAness that correlates with responsiveness to chemotherapy and with platinum sensitivity
```

```
  URL:
```

```
  PMIDs: 20547991
```

```
  Abstract: A 241 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 HG_U95Av2
```

```
    platform_summary:
```

```
      hgu95av2
```

```
    platform_manufacturer:
```

```
      Affymetrix
```

```
    platform_distribution:
```

```
      commercial
```

```
    platform_accession:
```

```
      GPL570|GPL8300
```

```
Preprocessing: rma
```

```
featureData(eset):
```

```
An object of class 'AnnotatedDataFrame'
```

```
  featureNames: AADAC AAK1 ... ZZZ3 (8957 total)
```

```
  varLabels: probeset gene
```

```
  varMetadata: labelDescription
```

**Details**

```

assayData: 8957 features, 42 samples
Platform type: hgu95av2
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
   42.00  42.00  42.00   23.00   3.78    2.79    NA

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

alt_sample_name:
  Length      Class      Mode
    42 character character

sample_type:
tumor
  42

primarysite:
ov
42

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

vital_status:
deceased  living
    23      19

batch:
2001-09-14 2001-12-14 2002-08-20 2003-09-09 2003-09-18
           7         4         14         13         4

uncurated_author_metadata:
  Length      Class      Mode
    42 character character

```

## 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. Sixteen paired breast-ovary tumors from patients with a former diagnosis of breast cancer were collected. The genomic profiles of paired tumors were analyzed using the Affymetrix GeneChip Mapping 50 K Xba Array or Genome-Wide Human SNP Array 6.0 (for one pair), and the data were normalized with ITALICS (Iterative and Alternative normalization and Copy number calling for affymetrix Snp arrays) algorithm or Partek Genomic Suite, respectively. The transcriptome of paired samples was analyzed using Affymetrix GeneChip Human Genome U133 Plus 2.0 Arrays, and the data were normalized with gc-Robust Multi-array Average (gcRMA) algorithm. A hierarchical clustering of these samples was performed, combined with a dataset of well-identified primary and secondary ovarian tumors. In 12 of the 16 paired tumors analyzed, the comparison of genomic profiles confirmed the pathological diagnosis of primary ovarian tumor (n = 5) or metastasis of breast cancer (n = 7). Among four cases with uncertain pathological diagnosis, genomic profiles were clearly distinct between the ovarian and breast tumors in two pairs, thus indicating primary ovarian carcinomas, and showed common patterns in the two others, indicating metastases from breast cancer. In all pairs, the result of the transcriptomic analysis was concordant with that of the genomic analysis. In patients with ovarian carcinoma and a previous history of breast cancer, SNP array analysis can be used to distinguish primary and secondary ovarian tumors. Transcriptomic analysis may be used when primary breast tissue specimen is not available.

## Usage

```
data( GSE20565_eset )
```

## Format

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Meyniel JP, Cottu PH, Decraene C, Stern MH, Couturier J, Lebigot I, Nicolas A, Weber
```

```
  Laboratory: Meyniel, Sastre-Garau 2010
```

```
  Contact information:
```

```
  Title: A genomic and transcriptomic approach for a differential diagnosis between primary and secondary
```

```
  URL:
```

```
  PMIDs: 20492709
```

```
  Abstract: A 277 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|GPL2005|GPL6801

Preprocessing: frma
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A1BG A1CF ... ZZZ3 (19093 total)
  varLabels: probeset gene
  varMetadata: labelDescription

```

## Details

```

assayData: 19093 features, 140 samples
Platform type: hgu133plus2
-----

```

```

Available sample meta-data:
-----

```

```

alt_sample_name:
  Length      Class      Mode
    140 character character

```

```

sample_type:
tumor
  140

```

```

histological_type:
  Length      Class      Mode
    140 character character

```

```

primarysite:
other  ov
  44   96

```

```

summarygrade:
high low NA's
  63  33  44

```

```

summarystage:
early late NA's
  27   67  46

```

```

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

```

substage:

a	b	c	NA's
14	10	55	61

grade:

1	2	3	NA's
6	27	63	44

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

uncurated\_author\_metadata:

Length	Class	Mode
140	character	character

---

GSE2109\_eset

*IGC EXpression Project for Oncology*

---

## Description

EXpression Project for Oncology, International Genomics Consortium, [www.intgen.org](http://www.intgen.org)

## Usage

```
data( GSE2109_eset )
```

## Format

experimentData(eset):

Experiment data

Experimenter name: EXpression Project for Oncology, International Genomics Consortium, [www.intgen.org](http://www.intgen.org)

Laboratory: exp0, IGC 2005

Contact information:

Title: IGC EXpression Project for Oncology

URL:

PMIDs: PMID unknown

Abstract: A 8 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

```

```

Preprocessing: frma
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A1BG A1CF ... ZZZ3 (19093 total)
  varLabels: probeset gene
  varMetadata: labelDescription

```

## Details

```

assayData: 19093 features, 204 samples
Platform type: hgu133plus2
-----

```

```

Available sample meta-data:
-----

```

```

alt_sample_name:
  Length      Class      Mode
    204 character character

```

```

sample_type:
tumor
  204

```

```

histological_type:
  Length      Class      Mode
    204 character character

```

```

primarysite:
other   ov  NA's
  23  178   3

```

```

summarygrade:
high low NA's
  91  31  82

```

```

summarystage:
early late NA's
  37   87   80

```

```

tumorstage:

```

```

  1   2   3   4 NA's
20  14  58  18  94

```

substage:

```

  a   b   c NA's
17  22  79  86

```

grade:

```

  1   2   3   4 NA's
11  20  83   8  82

```

age\_at\_initial\_pathologic\_diagnosis:

```

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

```

batch:

```

2004-12-03 2004-12-04 2004-12-07 2005-02-11 2005-03-03 2005-03-10 2005-03-11
           3           3           1           1           1           1           1
2005-03-15 2005-03-16 2005-03-17 2005-03-19 2005-03-22 2005-04-13 2005-04-26
           3           1           2           4           2           1           5
2005-04-29 2005-05-10 2005-06-01 2005-06-03 2005-06-08 2005-06-17 2005-08-05
           2           2           5           3           3           6           3
2005-08-09 2005-08-11 2005-09-07 2005-09-09 2005-09-13 2005-11-02 2005-11-04
           1           6           1           3           3           6           3
2005-11-15 2005-11-18 2005-12-02 2006-01-24 2006-01-26 2006-02-07 2006-02-28
           3           1           4           2           1           1           1
2006-03-06 2006-03-14 2006-04-18 2006-04-20 2006-05-16 2006-06-08 2006-07-26
           2           2           1           2           3           1           2
2006-07-28 2006-09-12 2006-09-14 2006-10-10 2006-10-24 2006-10-31 2006-11-09
           1           2           1           1           9           5           10
2006-11-21 2006-11-30 2006-12-07 2007-01-12 2007-02-09 2007-03-07 2007-03-09
           1           6           3           1           1           8           1
2007-03-15 2007-05-01 2007-05-03 2007-05-15 2007-05-18 2007-05-30 2007-06-12
           4           2           3           4           2           2           1
2007-07-27 2007-09-05 2007-09-07 2007-09-11 2007-09-12 2008-02-15 2008-02-21
           2           3           1           4           4           1           3
2008-02-27 2008-03-04 2008-05-13 2008-05-16 2008-05-23
           2           1           4           4           5

```

uncurated\_author\_metadata:

```

  Length   Class      Mode
    204 character character

```



**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**

```
experimentData(eset):
Experiment data
  Experimenter name: Bonome T, Levine DA, Shih J, Randonovich M, Pise-Masison CA, Bogomolnii F, Ozbun L,
  Laboratory: Bonome, Birrer 2008
  Contact information:
  Title: A gene signature predicting for survival in suboptimally debulked patients with ovarian cancer.
  URL:
  PMIDs: 18593951

Abstract: A 238 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

Preprocessing: frma
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A1CF A2M ... ZZZ3 (12981 total)
  varLabels: probeset gene
  varMetadata: labelDescription

```

### Details

```

assayData: 12981 features, 195 samples
Platform type: hgu133a
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
 185.00 185.00 185.00 129.00   3.83   3.24   4.83

```

```

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

```

```

alt_sample_name:
  Length      Class      Mode
    195 character character

```

```

sample_type:
healthy  tumor
    10    185

```

```

histological_type:
ser NA's
 185   10

```

```

primarysite:
ov
195

```

```

summarygrade:
high NA's
 185   10

```

```

summarystage:
late NA's
 185   10

```

tumorstage:

3	4	NA's
146	36	13

substage:

b	c	NA's
9	137	49

age\_at\_initial\_pathologic\_diagnosis:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
26.00	52.00	63.00	61.54	70.00	84.00	13

recurrence\_status:

norecurrence	recurrence
42	153

days\_to\_death:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
21.9	660.6	1164.0	1429.0	1880.0	4982.0	10

vital\_status:

deceased	living	NA's
129	56	10

debulking:

optimal	suboptimal	NA's
90	95	10

percent\_normal\_cells:

20-  
195

percent\_stromal\_cells:

20-  
195

percent\_tumor\_cells:

80+  
195

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						

```

uncurated_author_metadata:
  Length      Class      Mode
    195 character character

```

---

GSE30009_eset	<i>Multidrug resistance-linked gene signature predicts overall survival of patients with primary ovarian serous carcinoma.</i>
---------------	--

---

## Description

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. We applied a customized TaqMan low density array, a highly sensitive and specific assay, to study the expression profiles of 380 MDR-linked genes in 80 tumor specimens collected at initial surgery to debulk primary serous carcinoma. The RNA expression profiles of these drug resistance genes were correlated with clinical outcomes. Leave-one-out cross-validation was used to estimate the ability of MDR gene expression to predict survival. Although gene expression alone does not predict overall survival (OS;  $P = 0.06$ ), four covariates (age, stage, CA125 level, and surgical debulking) do ( $P = 0.03$ ). When gene expression was added to the covariates, we found an 11-gene signature that provides a major improvement in OS prediction (log-rank statistic  $P < 0.003$ ). The predictive power of this 11-gene signature was confirmed by dividing high- and low-risk patient groups, as defined by their clinical covariates, into four specific risk groups on the basis of expression levels. This study reveals an 11-gene signature that allows a more precise prognosis for patients with serous cancer of the ovary treated with carboplatin- and paclitaxel-based therapy. These 11 new targets offer opportunities for new therapies to improve clinical outcome in ovarian cancer.

## Usage

```
data( GSE30009_eset )
```

## Format

```

experimentData(eset):
Experiment data
  Experimenter name: Gillet JP, Calcagno AM, Varma S, Davidson B et al. Multidrug resistance-linked gene
  Laboratory: Gillet, Gottesman 2012
  Contact information:
  Title: Multidrug resistance-linked gene signature predicts overall survival of patients with primary o
  URL:
  PMIDs: 22492981

  Abstract: A 244 word abstract is available. Use 'abstract' method.
  Information is available on: preprocessing

```

```

notes:
  platform_title:
    TaqMan qRT-PCR Homo sapiens Low-Density Array 380
  platform_shorttitle:
    TaqMan qRT-PCR
  platform_summary:
    NA
  platform_manufacturer:
    TaqMan
  platform_distribution:
    custom
  platform_accession:
    GPL13728

```

```

Preprocessing: default
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: ABCA1 ABCA10 ... XRCC6 (359 total)
  varLabels: probeset gene
  varMetadata: labelDescription

```

## Details

```

assayData: 359 features, 103 samples
Platform type: NA
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
 103.00 103.00 103.00   57.00   3.42   2.92   5.34

```

```

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

```

```

alt_sample_name:
  Length      Class      Mode
    103 character character

```

```

sample_type:
tumor
  103

```

```

histological_type:
clearcell      ser
    1         102

```

```

summarygrade:

```

```

high  low NA's
 92   9   2

summarystage:
late
 103

tumorstage:
 3  4
82 21

substage:
  b  c NA's
 2 60 41

grade:
 1  2  3 NA's
 4  5 92  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

uncurated_author_metadata:
  Length    Class      Mode
   103 character character

```

---

GSE30161\_eset

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

---

### Description

Despite advances in radical surgery and chemotherapy delivery, ovarian cancer is the most lethal gynecologic malignancy. Standard therapy includes treatment with platinum-based combination

chemotherapies yet there is no biomarker model to predict their responses to these agents. We here have developed and independently tested our multi-gene molecular predictors for forecasting patients' responses to individual drugs on a cohort of 55 ovarian cancer patients. To independently validate these molecular predictors, we performed microarray profiling on FFPE tumor samples of 55 ovarian cancer patients (UVA-55) treated with platinum-based adjuvant chemotherapy. Genome-wide chemosensitivity biomarkers were initially discovered from the in vitro drug activities and genomic expression data for carboplatin and paclitaxel, respectively. Multivariate predictors were trained with the cell line data and then evaluated with a historical patient cohort. For the UVA-55 cohort, the carboplatin, taxol, and combination predictors significantly stratified responder patients and non-responder patients ( $p = 0.019, 0.04, 0.014$ ) with sensitivity = 91%, 96%, 93 and NPV = 57%, 67%, 67% in pathologic clinical response. The combination predictor also demonstrated a significant survival difference between predicted responders and non-responders with a median survival of 55.4 months vs. 32.1 months. Thus, COXEN single- and combination-drug predictors successfully stratified platinum resistance and taxane response in an independent cohort of ovarian cancer patients based on their FFPE tumor samples.

### Usage

```
data( GSE30161_eset )
```

### Format

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Ferriss JS, Kim Y, Duska L, Birrer M, Levine DA, Moskaluk C, Theodorescu D, Lee JK  
  Laboratory: Ferriss, Lee 2012
```

```
  Contact information:
```

```
  Title: Multi-gene expression predictors of single drug responses to adjuvant chemotherapy in ovarian c
```

```
  URL:
```

```
  PMIDs: 22348014
```

```
  Abstract: A 215 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
```

```
Preprocessing: frma
```

```
featureData(eset):
```

```
An object of class 'AnnotatedDataFrame'
  featureNames: A1BG A1CF ... ZZZ3 (19093 total)
  varLabels: probeset gene
  varMetadata: labelDescription
```

## Details

```
assayData: 19093 features, 58 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
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
```

```
histological_type:
  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
```

```
grade:
```



```

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

```

```

batch:
  2009-10-07 2009-10-08 2009-10-09 2009-10-20
             28         18         8         4

```

```

uncurated_author_metadata:
  Length Class Mode
     58 character character

```

---

GSE32062.GPL6480\_eset *High-risk ovarian cancer based on 126-gene expression signature is uniquely characterized by downregulation of antigen presentation pathway.*

---

## Description

High-grade serous ovarian cancers are heterogeneous not only in terms of clinical outcome but also at the molecular level. Our aim was to establish a novel risk classification system based on a gene expression signature for predicting overall survival, leading to suggesting novel therapeutic strategies for high-risk patients. In this large-scale cross-platform study of six microarray data sets consisting of 1,054 ovarian cancer patients, we developed a gene expression signature for predicting overall survival by applying elastic net and 10-fold cross-validation to a Japanese data set A (n = 260) and evaluated the signature in five other data sets. Subsequently, we investigated differences in the biological characteristics between high- and low-risk ovarian cancer groups. An elastic net analysis identified a 126-gene expression signature for predicting overall survival in patients with ovarian cancer using the Japanese data set A (multivariate analysis,  $P = 4 \times 10^{-20}$ ). We validated its predictive ability with five other data sets using multivariate analysis (Tohill's data set,  $P = 1 \times 10^{-5}$ ; Bonome's data set,  $P = 0.0033$ ; Dressman's data set,  $P = 0.0016$ ; TCGA data set,  $P = 0.0027$ ; Japanese data set B,  $P = 0.021$ ). Through gene ontology and pathway analyses, we identified a significant reduction in expression of immune-response-related genes, especially on the antigen presentation pathway, in high-risk ovarian cancer patients. This risk classification based on the 126-gene expression signature is an accurate predictor of clinical outcome in patients with advanced stage high-grade serous ovarian cancer and has the potential to develop new therapeutic strategies for high-grade serous ovarian cancer patients.

## Usage

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

## Format

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Yoshihara K, Tsunoda T, Shigemizu D, Fujiwara H et al. High-risk ovarian cancer base
```

```
  Laboratory: Yoshihara, Tanaka 2012
```

```
  Contact information:
```

```
  Title: High-risk ovarian cancer based on 126-gene expression signature is uniquely characterized by do
```

```
  URL:
```

```
  PMIDs: 22241791
```

```
  Abstract: A 255 word abstract is available. Use 'abstract' method.
```

```
  Information is available on: preprocessing
```

```
  notes:
```

```
    platform_title:
```

```
      Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name vers
ion)
```

```

platform_shorttitle:
  Agilent G4112F
platform_summary:
  hgug4112a
platform_manufacturer:
  Agilent
platform_distribution:
  commercial
platform_accession:
  GPL6480

```

```

Preprocessing: default
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A1CF A2LD1 ... ZZZ3 (19358 total)
  varLabels: probeset gene
  varMetadata: labelDescription

```

## Details

```

assayData: 19358 features, 260 samples
Platform type: hgug4112a
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
 260.00 260.00 260.00 121.00   4.93   4.11   6.58

```

```

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

```

```

alt_sample_name:
  Length      Class      Mode
    260 character character

```

```

sample_type:
tumor
  260

```

```

histological_type:
ser
  260

```

```

summarygrade:
high low
 129 131

```

```

summarystage:
late
  260

tumorstage:
  3  4
204 56

substage:
  a  b  c NA's
  4  20 180 56

grade:
  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

uncurated_author_metadata:
  Length  Class  Mode
    260 character character

```

## Description

High-grade serous ovarian cancers are heterogeneous not only in terms of clinical outcome but also at the molecular level. Our aim was to establish a novel risk classification system based on a gene expression signature for predicting overall survival, leading to suggesting novel therapeutic strategies for high-risk patients. In this large-scale cross-platform study of six microarray data sets consisting of 1,054 ovarian cancer patients, we developed a gene expression signature for predicting overall survival by applying elastic net and 10-fold cross-validation to a Japanese data set A (n = 260) and evaluated the signature in five other data sets. Subsequently, we investigated differences in the biological characteristics between high- and low-risk ovarian cancer groups. An elastic net analysis identified a 126-gene expression signature for predicting overall survival in patients with ovarian cancer using the Japanese data set A (multivariate analysis,  $P = 4 \times 10^{-20}$ ). We validated its predictive ability with five other data sets using multivariate analysis (Tohill's data set,  $P = 1 \times 10^{-5}$ ; Bonome's data set,  $P = 0.0033$ ; Dressman's data set,  $P = 0.0016$ ; TCGA data set,  $P = 0.0027$ ; Japanese data set B,  $P = 0.021$ ). Through gene ontology and pathway analyses, we identified a significant reduction in expression of immune-response-related genes, especially on the antigen presentation pathway, in high-risk ovarian cancer patients. This risk classification based on the 126-gene expression signature is an accurate predictor of clinical outcome in patients with advanced stage high-grade serous ovarian cancer and has the potential to develop new therapeutic strategies for high-grade serous ovarian cancer patients.

## Usage

```
data( GSE32063_eset )
```

## Format

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Yoshihara K, Tsunoda T, Shigemizu D, Fujiwara H et al. High-risk ovarian cancer base
```

```
  Laboratory: Yoshihara, Tanaka 2012
```

```
  Contact information:
```

```
  Title: High-risk ovarian cancer based on 126-gene expression signature is uniquely characterized by do
```

```
  URL:
```

```
  PMIDs: 22241791
```

```
  Abstract: A 255 word abstract is available. Use 'abstract' method.
```

```
  Information is available on: preprocessing
```

```
  notes:
```

```
    platform_title:
```

```
      Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name vers  
ion)
```

```
    platform_shorttitle:
```

```
      Agilent G4112F
```

```
    platform_summary:
```

```
      hgug4112a
```

```
    platform_manufacturer:
```

```
      Agilent
```

```
    platform_distribution:
```

```

      commercial
platform_accession:
  GPL6480

Preprocessing: default
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A1CF A2LD1 ... ZZZ3 (19358 total)
  varLabels: probeset gene
  varMetadata: labelDescription

```

### Details

```

assayData: 19358 features, 40 samples
Platform type: hgug4112a
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
   40.00  40.00  40.00   22.00   4.44   3.29    NA

```

```

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

```

```

alt_sample_name:
  Length      Class      Mode
     40 character character

```

```

sample_type:
tumor
  40

```

```

histological_type:
ser
  40

```

```

summarygrade:
high low
  17  23

```

```

summarystage:
late
  40

```

```

tumorstage:
  3  4
31  9

```

```

substage:
  b   c NA's
  3  28   9

grade:
  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

uncurated_author_metadata:
  Length    Class      Mode
    40 character character

```

GSE6008\_eset

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

## 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. In this study, gene expression profiling was used to examine LPA-mediated transcription by exogenously adding LPA to human epithelial ovarian cancer cells for 24 h to mimic long-term stimulation in the tumor microenvironment. The resultant transcriptional profile comprised a 39-gene

signature that closely correlated to serous epithelial ovarian carcinoma. Hierarchical clustering of ovarian cancer patient specimens demonstrated that the signature is associated with worsened prognosis. Patients with LPA-signature-positive ovarian tumors have reduced disease-specific and progression-free survival times. They have a higher frequency of stage IIIc serous carcinoma and a greater proportion is deceased. Among the 39-gene signature, a group of seven genes associated with cell adhesion recapitulated the results. Out of those seven, claudin-1, an adhesion molecule and phenotypic epithelial marker, is the only independent biomarker of serous epithelial ovarian carcinoma. Knockdown of claudin-1 expression in ovarian cancer cells reduces LPA-mediated cellular adhesion, enhances suspended cells and reduces LPA-mediated migration. The data suggest that transcriptional events mediated by LPA in the tumor microenvironment influence tumor progression through modulation of cell adhesion molecules like claudin-1 and, for the first time, report an LPA-mediated expression signature in ovarian cancer that predicts a worse prognosis.

### Usage

```
data( GSE6008_eset )
```

### Format

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Murph MM, Liu W, Yu S, Lu Y, Hall H, Hennessy BT, Lahad J, Schaner M, Helland A, Krist
```

```
  Laboratory: Murph, Mills 2009
```

```
  Contact information:
```

```
  Title: Lysophosphatidic acid-induced transcriptional profile represents serous epithelial ovarian car
```

```
  URL:
```

```
  PMIDs: 19440550
```

```
  Abstract: A 247 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
```

```
Preprocessing: frma
```

```
featureData(eset):
```

```
An object of class 'AnnotatedDataFrame'
```

```
  featureNames: A1CF A2M ... ZZZ3 (12981 total)
```

```
  varLabels: probeset gene
```



varMetadata: labelDescription

## Details

assayData: 12981 features, 103 samples

Platform type: hgu133a

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

alt\_sample\_name:

Length	Class	Mode
103	character	character

sample\_type:

healthy	tumor
4	99

histological\_type:

clearcell	endo	mucinous	ser	NA's
8	37	13	41	4

primarysite:

ov
103

summarygrade:

high	low	NA's
38	36	29

summarystage:

early	late	NA's
42	53	8

tumorstage:

1	2	3	4	NA's
35	11	44	9	4

substage:

a	b	c	d	NA's
19	2	54	1	27

grade:

1	2	3	NA's
19	17	38	29

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
2006-01-27							
	4						

uncurated\_author\_metadata:

Length	Class	Mode
103	character	character

---

GSE6822\_eset

*Classification of ovarian tumor samples*

---

## Description

Ouellet V, Provencher DM, Maugard CM, Le Page C, Ren F, Lussier C, Novak J, Ge B, Hudson TJ, Tonin PN, Mes-Masson A-M: Discrimination between serous low malignant potential and invasive epithelial ovarian tumors using molecular profiling. *Oncogene* 2005, 24:4672-4687.

## Usage

```
data( GSE6822_eset )
```

## Format

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

Experimenter name: Ouellet V, Provencher DM, Maugard CM, Le Page C, Ren F, Lussier C, Novak J, Ge B, Hudson TJ

Laboratory: Ouellet, Mes-Masson 2005

Contact information:

Title: Classification of ovarian tumor samples

URL:

PMIDs: PMID unknown

Abstract: A 40 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 Hu6800

platform\_summary:

hu6800

platform\_manufacturer:

Affymetrix

platform\_distribution:

```

commercial
platform_accession:
  GPL80

Preprocessing: rma
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A2M AADAC ... ZYX (5231 total)
  varLabels: probeset gene
  varMetadata: labelDescription

```

## Details

```

assayData: 5231 features, 66 samples
Platform type: hu6800
-----

```

```

Available sample meta-data:
-----

```

```

alt_sample_name:
  Length      Class      Mode
    66 character character

```

```

sample_type:
tumor
  66

```

```

histological_type:
  Length      Class      Mode
    66 character character

```

```

primarysite:
ov
  66

```

```

summarygrade:
high low NA's
  40  15  11

```

```

grade:
  1  2  3 NA's
  1 14 40  11

```

```

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

uncurated\_author\_metadata:  
Length Class Mode  
66 character character

---

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. Microarray gene expression profiling was done on 285 serous and endometrioid tumors of the ovary, peritoneum, and fallopian tube. K-means clustering was applied to identify robust molecular subtypes. Statistical analysis identified differentially expressed genes, pathways, and gene ontologies. Laser capture microdissection, pathology review, and immunohistochemistry validated the array-based findings. Patient survival within k-means groups was evaluated using Cox proportional hazards models. Class prediction validated k-means groups in an independent dataset. A semisupervised survival analysis of the array data was used to compare against unsupervised clustering results. Optimal clustering of array data identified six molecular subtypes. Two subtypes represented predominantly serous low malignant potential and low-grade endometrioid subtypes, respectively. The remaining four subtypes represented higher grade and advanced stage cancers of serous and endometrioid morphology. A novel subtype of high-grade serous cancers reflected a mesenchymal cell type, characterized by overexpression of N-cadherin and P-cadherin and low expression of differentiation markers, including CA125 and MUC1. A poor prognosis subtype was defined by a reactive stroma gene expression signature, correlating with extensive desmoplasia in such samples. A similar poor prognosis signature could be found using a semisupervised analysis. Each subtype displayed distinct levels and patterns of immune cell infiltration. Class prediction identified similar subtypes in an independent ovarian dataset with similar prognostic trends. Gene expression profiling identified molecular subtypes of ovarian cancer of biological and clinical importance.

## Usage

```
data( GSE9891_eset )
```

## Format

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Tothill RW, Tinker AV, George J, Brown R, Fox SB, Lade S, Johnson DS, Trivett MK, Ete
```

```
  Laboratory: Tothill, Bowtell 2008
```

```
  Contact information:
```

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

URL:  
PMIDs: 18698038

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

Preprocessing: frma  
featureData(eset):  
An object of class 'AnnotatedDataFrame'  
featureNames: A1BG A1CF ... ZZZ3 (19093 total)  
varLabels: probeset gene  
varMetadata: labelDescription

## Details

assayData: 19093 features, 285 samples  
Platform type: hgu133plus2  
Overall survival time-to-event summary (in years):  
Call: survfit(formula = Surv(time, cens) ~ -1)

7 observations deleted due to missingness

records	n.max	n.start	events	median	0.95LCL	0.95UCL
278.00	278.00	278.00	113.00	3.95	3.53	5.01

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

alt\_sample\_name:  

Length	Class	Mode
285	character	character

sample\_type:  
tumor  
285

## histological\_type:

endo	other	ser
20	1	264

## primarysite:

ft	other	ov
8	34	243

## arrayedsite:

ft	other	ov
2	83	200

## summarygrade:

high	low	NA's
163	116	6

## summarystage:

early	late	NA's
42	240	3

## tumorstage:

1	2	3	4	NA's
24	18	218	22	3

## substage:

a	b	c	NA's
26	19	212	28

## grade:

1	2	3	NA's
19	97	163	6

## age\_at\_initial\_pathologic\_diagnosis:

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

## pltx:

n	y	NA's
39	243	3

## tax:

n	y	NA's
87	195	3

## neo:

n	y	NA's
264	18	3

## days\_to\_tumor\_recurrence:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
0.0	300.0	450.0	618.9	810.0	4980.0	10

## recurrence\_status:

norecurrence	recurrence	NA's
94	188	3

## days\_to\_death:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
0.0	547.5	855.0	955.1	1252.0	6420.0	7

## vital\_status:

deceased	living	NA's
113	169	3

## debulking:

optimal	suboptimal	NA's
160	88	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

## uncurated\_author\_metadata:

Length	Class	Mode
285	character	character

**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. RNA from 65 serous ovarian cancers was analyzed using Affymetrix U133A microarrays. This included 54 stage III/IV cases (30 short-term survivors who lived <3 years and 24 long-term survivors who lived >7 years) and 11 stage I/II cases. Genes were screened on the basis of their level of and variability in expression, leaving 7,821 for use in developing a predictive model for survival. A composite predictive model was developed that combines Bayesian classification tree and multivariate discriminant models. Leave-one-out cross-validation was used to select and evaluate models. Patterns of genes were identified that distinguish short-term and long-term ovarian cancer survivors. The expression model developed for advanced stage disease classified all 11 early-stage ovarian cancers as long-term survivors. The MAL gene, which has been shown to confer resistance to cancer therapy, was most highly overexpressed in short-term survivors (3-fold compared with long-term survivors, and 29-fold compared with early-stage cases). These results suggest that gene expression patterns underlie differences in outcome, and an examination of the genes that provide this discrimination reveals that many are implicated in processes that define the malignant phenotype. Differences in survival of advanced ovarian cancers are reflected by distinct patterns of gene expression. This biological distinction is further emphasized by the finding that early-stage cancers share expression patterns with the advanced stage long-term survivors, suggesting a shared favorable biology.

**Usage**

```
data( PMID15897565_eset )
```

**Format**

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

```
  Experimenter name: Berchuck A, Iversen ES, Lancaster JM, Pittman J, Luo J, Lee P, Murphy S, Dressman HK,
```

```
  Laboratory: Berchuck, Marks 2005
```

```
  Contact information:
```

```
  Title: Patterns of gene expression that characterize long-term survival in advanced stage serous ovar
```

```
  URL:
```

```
  PMIDs: 15897565
```

```
  Abstract: A 258 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
warnings:
  These samples are a subset of PMID17290060.

```

```

Preprocessing: frma
featureData(eset):
An object of class 'AnnotatedDataFrame'
  featureNames: A1CF A2M ... ZZZ3 (12981 total)
  varLabels: probeset gene
  varMetadata: labelDescription

```

## Details

```

assayData: 12981 features, 63 samples
Platform type: hgu133a
Binary overall survival summary (definitions of long and short provided by study authors):

```

```

long short NA's
  24   28   11

```

```

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

```

```

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

```

```

sample_type:
tumor
  63

```

```

histological_type:
ser
  63

```

```

primarysite:
ov
  63

```

```

summarygrade:
high low NA's
  25  37   1

```

```

summarystage:
early late

```

```

11    52

tumorstage:
  1  2  3  4
  7  4 48  4

grade:
  1  2  3  4 NA's
  2 35 24  1  1

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

os_binary:
  long short  NA's
  24    28    11

debulking:
  optimal suboptimal    NA's
      24         28         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

uncurated_author_metadata:
  Length    Class    Mode
     63 character character

```

---

 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. A gene expression model that predicts response to platinum-based therapy was developed using a training set of 83 advanced-stage serous ovarian cancers and tested on a 36-sample external validation set. In parallel, expression signatures that define the status of oncogenic signaling pathways were evaluated in 119 primary ovarian cancers and 12 ovarian cancer cell lines. In an effort to increase

chemotherapy sensitivity, pathways shown to be activated in platinum-resistant cancers were subject to targeted therapy in ovarian cancer cell lines. Gene expression profiles identified patients with ovarian cancer likely to be resistant to primary platinum-based chemotherapy with greater than 80% accuracy. In patients with platinum-resistant disease, we identified expression signatures consistent with activation of Src and Rb/E2F pathways, components of which were successfully targeted to increase response in ovarian cancer cell lines. We have defined a strategy for treatment of patients with advanced-stage ovarian cancer that uses therapeutic stratification based on predictions of response to chemotherapy, coupled with prediction of oncogenic pathway deregulation, as a method to direct the use of targeted agents.

## Usage

```
data( PMID17290060_eset )
```

## Format

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Dressman HK, Berchuck A, Chan G, Zhai J, Bild A, Sayer R, Cragun J, Clarke J, Whitake
```

```
  Laboratory: Dressman, Lancaster 2007
```

```
  Contact information:
```

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

```
  URL:
```

```
  PMIDs: 17290060
```

```
  Abstract: A 223 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
```

```
  warnings:
```

```
    This paper has been retracted.
```

```
Preprocessing: frma
```

```
featureData(eset):
```

```
An object of class 'AnnotatedDataFrame'
```

```
  featureNames: A1CF A2M ... ZZZ3 (12981 total)
```

```
  varLabels: probeset gene
```

```
  varMetadata: labelDescription
```

**Details**

assayData: 12981 features, 117 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
117.00	117.00	117.00	67.00	5.26	2.79	7.48

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

alt\_sample\_name:  

Length	Class	Mode
117	character	character

sample\_type:  
 tumor  
 117

histological\_type:  
 ser  
 117

primarysite:  
 ov  
 117

summarygrade:  
 high low NA's  
 57 57 3

summarystage:  
 early late NA's  
 1 115 1

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

grade:  
 1 2 3 4 NA's  
 4 53 56 1 3

days\_to\_death:

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

vital\_status:

deceased	living
67	50

primary\_therapy\_outcome\_success:

completeresponse	progressivedisease
85	32

debulking:

optimal	suboptimal
63	54

batch:

2002-09-20	2002-10-23	2002-11-12	2002-12-16	2002-12-21	2003-01-03	2003-05-30
10	8	9	1	3	11	10
2004-03-09	2004-03-16	2004-04-20	2004-05-18	2004-05-21	2004-05-27	2004-06-22
16	6	5	15	7	7	1
2004-06-23						
8						

uncurated\_author\_metadata:

Length	Class	Mode
117	character	character

---

 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. Frozen specimens from 39 stage I/II, 42 stage III/IV, and 20 low malignant potential cancers were obtained from four different sites. A linear discriminant model was used to predict survival based upon array data. We validated the late-stage survival model and show that three of the most differentially expressed genes continue to be predictive of outcome. Most early-stage cancers (38 of 39 invasive, 15 of 20 low malignant potential) were classified as long-term survivors (median probabilities 0.97 and 0.86). MAL, the most differentially expressed gene, was further validated at the protein level and found to be an independent predictor of poor survival in an unselected group of advanced serous cancers ( $P = 0.0004$ ). These data suggest that serous ovarian cancers detected at an early stage generally have a favorable underlying biology

similar to advanced-stage cases that are long-term survivors. Conversely, most late-stage ovarian cancers seem to have a more virulent biology. This insight suggests that if screening approaches are to succeed it will be necessary to develop approaches that are able to detect these virulent cancers at an early stage.

## Usage

```
data( PMID19318476_eset )
```

## Format

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Berchuck A, Iversen ES, Luo J, Clarke JP, Horne H, Levine DA, Boyd J, Alonso MA, Seco
```

```
  Laboratory: Berchuck, Lancaster 2009
```

```
  Contact information:
```

```
  Title: Microarray analysis of early stage serous ovarian cancers shows profiles predictive of favorabl
```

```
  URL:
```

```
  PMIDs: 19318476
```

```
  Abstract: A 241 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
```

```
  warnings:
```

```
    These samples are a subset of PMID17290060.
```

```
Preprocessing: frma
```

```
featureData(eset):
```

```
An object of class 'AnnotatedDataFrame'
```

```
  featureNames: A1CF A2M ... ZZZ3 (12981 total)
```

```
  varLabels: probeset gene
```

```
  varMetadata: labelDescription
```

## Details

```
assayData: 12981 features, 42 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
42.00	42.00	42.00	22.00	2.79	2.30	NA

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

alt\_sample\_name:

Length	Class	Mode
42	character	character

sample\_type:

tumor
42

histological\_type:

ser
42

summarygrade:

high	low	NA's
24	17	1

summarystage:

early	late	NA's
2	39	1

tumorstage:

1	2	3	4	NA's
1	1	29	10	1

substage:

a	b	c	NA's
1	1	29	11

grade:

1	2	3	NA's
2	15	24	1

age\_at\_initial\_pathologic\_diagnosis:

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

recurrence\_status:

```

norecurrence  recurrence
              6          36

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

vital_status:
deceased  living
         22    20

debulking:
  optimal suboptimal    NA's
         20         21         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

uncurated_author_metadata:
  Length    Class    Mode
     42 character character

```

---

TCGA.mirna.8x15kv2\_eset

*Integrated genomic analyses of ovarian carcinoma.*

---

## Description

A catalogue of molecular aberrations that cause ovarian cancer is critical for developing and deploying therapies that will improve patients' lives. The Cancer Genome Atlas project has analysed messenger RNA expression, microRNA expression, promoter methylation and DNA copy number in 489 high-grade serous ovarian adenocarcinomas and the DNA sequences of exons from coding genes in 316 of these tumours. Here we report that high-grade serous ovarian cancer is characterized by TP53 mutations in almost all tumours (96%); low prevalence but statistically recurrent somatic mutations in nine further genes including NF1, BRCA1, BRCA2, RB1 and CDK12; 113 significant focal DNA copy number aberrations; and promoter methylation events involving 168 genes. Analyses delineated four ovarian cancer transcriptional subtypes, three microRNA subtypes, four promoter methylation subtypes and a transcriptional signature associated with survival duration, and shed new light on the impact that tumours with BRCA1/2 (BRCA1 or BRCA2) and CCNE1 aberrations have on survival. Pathway analyses suggested that homologous recombination is defective in about half of the tumours analysed, and that NOTCH and FOXM1 signalling are involved in serous ovarian cancer pathophysiology.



**Usage**

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

**Format**

```
experimentData(eset):
```

```
Experiment data
```

```
  Experimenter name: Integrated genomic analyses of ovarian carcinoma. Nature 2011, 474:609-615.
```

```
  Laboratory: Cancer Genome Atlas Research Network 2011
```

```
  Contact information:
```

```
  Title: Integrated genomic analyses of ovarian carcinoma.
```

```
  URL:
```

```
  PMIDs: 21720365
```

```
  Abstract: A 179 word abstract is available. Use 'abstract' method.
```

```
  Information is available on: preprocessing
```

```
  notes:
```

```
    platform_title:
```

```
      [miRNA-8x15k2] Agilent Human miRNA G4470B
```

```
    platform_shorttitle:
```

```
      Agilent miRNA-8x15k2 G4470B
```

```
    platform_summary:
```

```
      NA
```

```
    platform_manufacturer:
```

```
      Agilent
```

```
    platform_distribution:
```

```
      commercial
```

```
    platform_accession:
```

```
      NA
```

```
Preprocessing: default
```

```
featureData(eset):
```

```
An object of class 'AnnotatedDataFrame'
```

```
  featureNames: ebv-miR-BART1-3p ebv-miR-BART1-5p ... kshv-miR-K12-9*  
  (799 total)
```

```
  varLabels: probeset gene
```

```
  varMetadata: labelDescription
```

**Details**

```
assayData: 799 features, 554 samples
```

```
Platform type: NA
```

```
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

unique\_patient\_ID:

Length	Class	Mode
554	character	character

sample\_type:

tumor
554

histological\_type:

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

grade:

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

## pltx:

n y NA's  
19 478 57

## tax:

n y NA's  
43 454 57

## neo:

n NA's  
497 57

## 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	metastasis	NA's
145	3	138	268

## primary\_therapy\_outcome\_success:

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

## debulking:

optimal	suboptimal	NA's
359	137	58

## percent\_normal\_cells:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
------	---------	--------	------	---------	------	------

```

0.000  0.000  0.000  2.375  0.000  55.000  10

percent_stromal_cells:
  Min. 1st Qu.  Median    Mean 3rd Qu.  Max.  NA's
  0.00  5.00  10.00  12.78  19.00  70.00  16

percent_tumor_cells:
  Min. 1st Qu.  Median    Mean 3rd Qu.  Max.  NA's
  0.00  75.00  85.00  80.72  90.00  100.00  13

uncurated_author_metadata:
  Length      Class      Mode
  554 character character

```

---

TCGA\_eset

---

*Integrated genomic analyses of ovarian carcinoma.*


---

## Description

A catalogue of molecular aberrations that cause ovarian cancer is critical for developing and deploying therapies that will improve patients' lives. The Cancer Genome Atlas project has analysed messenger RNA expression, microRNA expression, promoter methylation and DNA copy number in 489 high-grade serous ovarian adenocarcinomas and the DNA sequences of exons from coding genes in 316 of these tumours. Here we report that high-grade serous ovarian cancer is characterized by TP53 mutations in almost all tumours (96%); low prevalence but statistically recurrent somatic mutations in nine further genes including NF1, BRCA1, BRCA2, RB1 and CDK12; 113 significant focal DNA copy number aberrations; and promoter methylation events involving 168 genes. Analyses delineated four ovarian cancer transcriptional subtypes, three microRNA subtypes, four promoter methylation subtypes and a transcriptional signature associated with survival duration, and shed new light on the impact that tumours with BRCA1/2 (BRCA1 or BRCA2) and CCNE1 aberrations have on survival. Pathway analyses suggested that homologous recombination is defective in about half of the tumours analysed, and that NOTCH and FOXM1 signalling are involved in serous ovarian cancer pathophysiology.

## Usage

```
data( TCGA_eset )
```

## Format

```

experimentData(eset):
Experiment data
  Experimenter name: Integrated genomic analyses of ovarian carcinoma. Nature 2011, 474:609-615.
  Laboratory: Cancer Genome Atlas Research Network 2011
  Contact information:
  Title: Integrated genomic analyses of ovarian carcinoma.

```

URL:

PMIDs: 21720365

Abstract: A 179 word abstract is available. Use 'abstract' method.

Information is available on: preprocessing

notes:

platform\_title:

[HT\_HG-U133A] Affymetrix HT Human Genome U133A Array

platform\_shorttitle:

Affymetrix HT\_HG-U133A

platform\_summary:

hthgu133a

platform\_manufacturer:

Affymetrix

platform\_distribution:

commercial

platform\_accession:

GPL3921

warnings:

The following samples are likely from specimens also used in GSE26712: TCGA A.13.0725, TCGA.13.0885, TCGA.13.0887, TCGA.13.0890, TCGA.13.0886, TCGA.13.0714, TCGA.13.0727, TCGA.13.1817, TCGA.13.1499, TCGA.13.0883

Preprocessing: rma

featureData(eset):

An object of class 'AnnotatedDataFrame'

featureNames: A1CF A2M ... ZZZ3 (12981 total)

varLabels: probeset gene

varMetadata: labelDescription

## Details

assayData: 12981 features, 578 samples

Platform type: hthgu133a

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

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

21 observations deleted due to missingness

records	n.max	n.start	events	median	0.95LCL	0.95UCL
557.00	557.00	557.00	290.00	3.73	3.45	4.06

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

alt\_sample\_name:

Length	Class	Mode
578	character	character

```

unique_patient_ID:
  Length   Class      Mode
  578 character character

sample_type:
adjacentnormal      tumor
      8              570

histological_type:
ser NA's
568  10

primarysite:
other  ov  NA's
  4  564  10

summarygrade:
high  low  NA's
480  75  23

summarystage:
early  late  NA's
  43  520  15

tumorstage:
  1  2  3  4  NA's
16  27 436 84  15

substage:
  b  c  NA's
31 448 99

grade:
  1  2  3  4  NA's
  6  69 479 1  23

age_at_initial_pathologic_diagnosis:
  Min. 1st Qu.  Median    Mean 3rd Qu.  Max.  NA's
 26.00  51.00  59.00  59.70  68.25  89.00  10

pltx:
  n  y  NA's
19 492 67

tax:
  n  y  NA's
43 468 67

```

neo:

n	NA's
511	67

days\_to\_tumor\_recurrence:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
8.0	238.2	443.5	623.7	812.0	5480.0	56

recurrence\_status:

norecurrence	recurrence
279	299

days\_to\_death:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
8	349	881	1010	1446	5480	21

vital\_status:

deceased	living	NA's
290	270	18

site\_of\_tumor\_first\_recurrence:

locoregional	locoregional_plus_metastatic	NA's
153	3	
metastasis		NA's
143		279

primary\_therapy\_outcome\_success:

completeresponse	partialresponse	progressivedisease	stabledisease
318	65	41	30
NA's			
124			

debulking:

optimal	suboptimal	NA's
367	140	71

percent\_normal\_cells:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
0.000	0.000	0.000	2.385	0.000	55.000	19

percent\_stromal\_cells:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
0.00	5.00	10.00	12.85	20.00	70.00	25

percent\_tumor\_cells:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
0.00	75.00	85.00	80.64	90.00	100.00	22

batch:

9	11	12	13	14	15	17	18	19	21	22	24	27	40	NA's
45	37	45	47	46	22	47	47	47	46	47	46	6	49	1

uncurated\_author\_metadata:

Length	Class	Mode
578	character	character



# Index

## \*Topic **datasets**

- E.MTAB.386\_eset, 4
  - GSE12418\_eset, 7
  - GSE12470\_eset, 9
  - GSE13876\_eset, 11
  - GSE14764\_eset, 14
  - GSE17260\_eset, 17
  - GSE18520\_eset, 20
  - GSE19829.GPL570\_eset, 23
  - GSE19829.GPL8300\_eset, 25
  - GSE20565\_eset, 27
  - GSE2109\_eset, 30
  - GSE26712\_eset, 32
  - GSE30009\_eset, 36
  - GSE30161\_eset, 38
  - GSE32062.GPL6480\_eset, 42
  - GSE32063\_eset, 44
  - GSE6008\_eset, 47
  - GSE6822\_eset, 50
  - GSE9891\_eset, 52
  - PMID15897565\_eset (PMID15897565\_eset), 55
  - PMID17290060\_eset (PMID17290060\_eset), 58
  - PMID19318476\_eset (PMID19318476\_eset), 61
  - TCGA.mirna.8x15kv2\_eset (TCGA.mirna.8x15kv2\_eset), 64
  - TCGA\_eset (TCGA\_eset), 68
- curatedOvarianData (curatedOvarianData-package), 2
- curatedOvarianData-package, 2
- E.MTAB.386\_eset, 4
- GSE12418\_eset, 7
- GSE12470\_eset, 9
- GSE13876\_eset, 11
- GSE14764\_eset, 14
- GSE17260\_eset, 17
- GSE18520\_eset, 20
- GSE19829.GPL570\_eset, 23
- GSE19829.GPL8300\_eset, 25
- GSE20565\_eset, 27
- GSE2109\_eset, 30
- GSE26712\_eset, 32
- GSE30009\_eset, 36
- GSE30161\_eset, 38
- GSE32062.GPL6480\_eset, 42
- GSE32063\_eset, 44
- GSE6008\_eset, 47
- GSE30009\_eset (GSE30009\_eset), 36
- GSE30161\_eset (GSE30161\_eset), 38
- GSE32062.GPL6480\_eset (GSE32062.GPL6480\_eset), 42
- GSE32063\_eset (GSE32063\_eset), 44
- GSE6008\_eset (GSE6008\_eset), 47
- GSE6822\_eset (GSE6822\_eset), 50
- GSE9891\_eset (GSE9891\_eset), 52
- PMID15897565\_eset (PMID15897565\_eset), 55
- PMID17290060\_eset (PMID17290060\_eset), 58
- PMID19318476\_eset (PMID19318476\_eset), 61
- TCGA.mirna.8x15kv2\_eset (TCGA.mirna.8x15kv2\_eset), 64
- TCGA\_eset (TCGA\_eset), 68
- E.MTAB.386\_eset (E.MTAB.386\_eset), 4
- GSE12418\_eset (GSE12418\_eset), 7
- GSE12470\_eset (GSE12470\_eset), 9
- GSE13876\_eset (GSE13876\_eset), 11
- GSE14764\_eset (GSE14764\_eset), 14
- GSE17260\_eset (GSE17260\_eset), 17
- GSE18520\_eset (GSE18520\_eset), 20
- GSE19829.GPL570\_eset (GSE19829.GPL570\_eset), 23
- GSE19829.GPL8300\_eset (GSE19829.GPL8300\_eset), 25
- GSE20565\_eset (GSE20565\_eset), 27
- GSE2109\_eset (GSE2109\_eset), 30
- GSE26712\_eset (GSE26712\_eset), 32

GSE6822\_eset, [50](#)

GSE9891\_eset, [52](#)

PMID15897565\_eset, [55](#)

PMID17290060\_eset, [58](#)

PMID19318476\_eset, [61](#)

TCGA.mirna.8x15kv2\_eset, [64](#)

TCGA\_eset, [68](#)