Package 'breastCancerNKI'

November 10, 2026

Type Package

Title Genexpression dataset published by van't Veer et al. [2002] and van de Vijver et al. [2002] (NKI).

Version 1.50.0

Date 2011-02-10

Description Genexpression data from a breast cancer study published by van't Veer et al. in 2002 and van de Vijver et al. in 2002, provided as an eSet.

biocViews ExperimentData, CancerData, BreastCancerData, CGHData, MicroarrayData, OneChannelData, ChipOnChipData

Author Markus Schroeder, Benjamin Haibe-Kains, Aedin Culhane, Christos Sotiriou, Gianluca Bontempi, John Quackenbush

Maintainer Markus Schroeder <mschroed@jimmy.harvard.edu>, Benjamin Haibe-Kains <bhaibeka@jimmy.harvard.edu>

Depends R (>= 2.5.0)

Suggests survcomp, genefu, Biobase

LazyLoad yes

License Artistic-2.0

URL http://compbio.dfci.harvard.edu/

git_url https://git.bioconductor.org/packages/breastCancerNKI

git_branch RELEASE_3_23

git_last_commit 9b4b843

git_last_commit_date 2026-04-28

Repository Bioconductor 3.23

Date/Publication 2026-11-10

Contents

nki 2

Index **5**

nki *Gene expression, annotations and clinical data from van't Veer et al. 2002 and van de Vijver et al. 2002*

Description

This dataset contains the gene expression, annotations and clinical data as published in van't Veer et al. 2002 and van de Vijver et al. 2002.

Usage

```
data(nki)
```

Format

ExpressionSet with 24481 features and 337 samples, containing:

- `exprs(nki)`: Matrix containing gene expressions as measured by Agilent technology (dual-channel, oligonucleotides).
- `fData(nki)`: AnnotatedDataFrame containing annotations of Agilent microarray platform.
- `pData(nki)`: AnnotatedDataFrame containing Clinical information of the breast cancer patients whose tumors were hybridized.
- `experimentalData(nki)`: MIAME object containing information about the dataset.
- `annotation(nki)`: Name of the agilent chip.

Details

This dataset represents the study published by van't Veer et al. in 2002 and van de Vijver et al. in 2002.

- `van t Veer et al.`: Breast cancer patients with the same stage of disease can have markedly different treatment responses and overall outcome. The strongest predictors for metastases (for example, lymph node status and histological grade) fail to classify accurately breast tumours according to their clinical behaviour. Chemotherapy or hormonal therapy reduces the risk of distant metastases by approximately one-third; however, 70-80% of patients receiving this treatment would have survived without it. None of the signatures of breast cancer gene expression reported to date allow for patient-tailored therapy strategies. Here we used DNA microarray analysis on primary breast tumours of 117 young patients, and applied supervised classification to identify a gene expression signature strongly predictive of a short interval to distant metastases ('poor prognosis' signature) in patients without tumour cells in local lymph nodes at diagnosis (lymph node negative). In addition, we established a signature that identifies tumours of BRCA1 carriers. The poor prognosis signature consists of genes regulating cell cycle, invasion, metastasis and angiogenesis. This gene expression profile will outperform all currently used clinical parameters in predicting disease outcome. Our findings provide a strategy to select patients who would benefit from adjuvant therapy.

- van de Vijver et al.: Background: A more accurate means of prognostication in breast cancer will improve the selection of patients for adjuvant systemic therapy. Methods: Using microarray analysis to evaluate our previously established 70-gene prognosis profile, we classified a series of 295 consecutive patients with primary breast carcinomas as having a gene-expression signature associated with either a poor prognosis or a good prognosis. All patients had stage I or II breast cancer and were younger than 53 years old; 151 had lymph-node-negative disease, and 144 had lymph-node-positive disease. We evaluated the predictive power of the prognosis profile using univariable and multivariable statistical analyses. Results: Among the 295 patients, 180 had a poor-prognosis signature and 115 had a good-prognosis signature, and the mean (+-SE) overall 10-year survival rates were 54.6 +- 4.4 percent and 94.5 +- 2.6 percent, respectively. At 10 years, the probability of remaining free of distant metastases was 50.6 +- 4.5 percent in the group with a poor-prognosis signature and 85.2 +- 4.3 percent in the group with a good-prognosis signature. The estimated hazard ratio for distant metastases in the group with a poor-prognosis signature, as compared with the group with the good-prognosis signature, was 5.1 (95 percent confidence interval, 2.9 to 9.0; P<0.001). This ratio remained significant when the groups were analyzed according to lymph-node status. Multivariable Cox regression analysis showed that the prognosis profile was a strong independent factor in predicting disease outcome. Conclusions: The gene-expression profile we studied is a more powerful predictor of the outcome of disease in young patients with breast cancer than standard systems based on clinical and histologic criteria.

Source

http://www.rii.com/publications/2002/vantveer.html

References

Laura J. van't Veer and Hongyue Dai and Marc J. van de Vijver and Yudong D. He and Augustinus A.M. Hart and Mao Mao and Hans L. Peterse and Karin van der Kooy and Matthew J. Marton and Anke T. Witteveen and George J. Schreiber and Ron M. Kerkhoven and Chris Roberts and Peter S. Linsley an Rene Bernards and Stephen H. Friend (2002) "Gene expression profiling predicts clinical outcome of breast cancer", *Nature*, **415**:530-536

M. J. van de Vijver and Y. D. He and L. van't Veer and H. Dai and A. M. Hart and D. W. Voskuil and G. J. Schreiber and J. L. Peterse and C. Roberts and M. J. Marton and M. Parrish and D. Atsma and A. Witteveen and A. Glas and L. Delahaye and T. van der Velde and H. Bartelink and S. Rodenhuis and E. T. Rutgers and S. H. Friend and R. Bernards (2002) "A Gene Expression Signature as a Predictor of Survival in Breast Cancer", *New England Journal of Medicine*, **347**(25):1999-2009

Examples

```
## load Biobase package
library(Biobase)
## load the dataset
data(nki)
## show the first 5 rows and columns of the expression data
exprs(nki)[1:5,1:5]
## show the first 6 rows of the phenotype data
head(pData(nki))
## show first 20 feature names
featureNames(nki)[1:20]
```

```
## show the experiment data summary
experimentData(nki)
## show the used platform
annotation(nki)
## show the abstract for this dataset
abstract(nki)
```

Index

∗ **datasets**
nki, 2

nki, 2