Package 'breastCancerVDX'

November 10, 2026

Type Package

Title Gene expression datasets published by Wang et al. [2005] and Minn et al. [2007] (VDX).

Version 1.50.0

Date 2011-02-10

Description Gene expression data from a breast cancer study published by Wang et al. in 2005 and Minn et al. in 2007, provided as an eSet.

biocViews ExperimentData, Genome, Homo_sapiens_Data, CancerData, BreastCancerData, LungCancerData, MicroarrayData, OneChannelData, GEO

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/breastCancerVDX

git_branch RELEASE_3_23

git_last_commit a3b50f3

git_last_commit_date 2026-04-28

Repository Bioconductor 3.23

Date/Publication 2026-11-10

Contents

vdx 2

Index **5**

vdx *Gene expression, annotations and clinical data from Wang et al. 2005 and Minn et al 2007*

Description

This dataset contains the gene expression, annotations and clinical data as published in Wang et al. 2005 and Minn et al 2007.

Usage

```
data(vdx)
```

Format

ExpressionSet with 22283 features and 344 samples, containing:

- `exprs(vdx)`: Matrix containing gene expressions as measured by Affymetrix hgu133a technology (single-channel, oligonucleotides).
- `fData(vdx)`: AnnotatedDataFrame containing annotations of Affy microarray platform hgu133a.
- `pData(vdx)`: AnnotatedDataFrame containing Clinical information of the breast cancer patients whose tumors were hybridized.
- `experimentalData(vdx)`: MIAME object containing information about the dataset.
- `annotation(vdx)`: Name of the affy chip.

Details

This dataset represent the study published by Wang et al. 2005 and Minn et al. 2007.

- `Wang et al.`: Background: Genome-wide measures of gene expression can identify patterns of gene activity that subclassify tumours and might provide a better means than is currently available for individual risk assessment in patients with lymph-node-negative breast cancer. Methods: We analysed, with Affymetrix Human U133a GeneChips, the expression of 22 000 transcripts from total RNA of frozen tumour samples from 286 lymph-node-negative patients who had not received adjuvant systemic treatment. Findings: In a training set of 115 tumours, we identified a 76-gene signature consisting of 60 genes for patients positive for oestrogen receptors (ER) and 16 genes for ER-negative patients. This signature showed 93% sensitivity and 48% specificity in a subsequent independent testing set of 171 lymph-node-negative patients. The gene profile was highly informative in identifying patients who developed distant metastases within 5 years (hazard ratio 5.67 [95% CI 2.59-12.4]), even when corrected for traditional prognostic factors in multivariate analysis (5.55 [2.46-12.5]). The 76-gene profile also represented a strong prognostic factor for the development of metastasis in the subgroups of 84 premenopausal patients (9.60 [2.28-40.5]), 87 postmenopausal patients (4.04 [1.57-10.4]), and 79 patients with tumours of 10-20 mm (14.1 [3.34-59.2]), a group of patients for whom prediction of prognosis is especially difficult. Interpretation: The identified signature provides a powerful tool for identification of patients at high risk of distant recurrence. The ability to identify patients who have a favourable prognosis could, after independent confirmation, allow clinicians to avoid adjuvant systemic therapy or to choose less aggressive therapeutic options.

- Minn et al.: The association between large tumor size and metastatic risk in a majority of clinical cancers has led to questions as to whether these observations are causally related or whether one is simply a marker for the other. This is partly due to an uncertainty about how metastasis-promoting gene expression changes can arise in primary tumors. We investigated this question through the analysis of a previously defined "lung metastasis gene-expression signature" (LMS) that mediates experimental breast cancer metastasis selectively to the lung and is expressed by primary human breast cancer with a high risk for developing lung metastasis. Experimentally, we demonstrate that the LMS promotes primary tumor growth that enriches for LMS+ cells, and it allows for intravasation after reaching a critical tumor size. Clinically, this corresponds to LMS+ tumors being larger at diagnosis compared with LMS- tumors and to a marked rise in the incidence of metastasis after LMS+ tumors reach 2 cm. Patients with LMS-expressing primary tumors selectively fail in the lung compared with the bone or other visceral sites and have a worse overall survival. The mechanistic linkage between metastasis gene expression, accelerated tumor growth, and likelihood of metastatic recurrence provided by the LMS may help to explain observations of prognostic gene signatures in primary cancer and how tumor growth can both lead to metastasis and be a marker for cells destined to metastasize.

Source

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE2034

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE5327

References

Y. Wang and J. G. Klijn and Y. Zhang and A. M. Sieuwerts and M. P. Look and F. Yang and D. Talantov and M. Timmermans and M. E. Meijer-van Gelder and J. Yu and T. Jatkoe and E. M. Berns and D. Atkins and J. A. Foekens (2005) "Gene-Expression Profiles to Predict Distant Metastasis of Lymph-Node-Negative Primary Breast Cancer", *Lancet*, **365**:671-679

Minn, Andy J. and Gupta, Gaorav P. and Padua, David and Bos, Paula and Nguyen, Don X. and Nuyten, Dimitry and Kreike, Bas and Zhang, Yi and Wang, Yixin and Ishwaran, Hemant and Foekens, John A. and van de Vijver, Marc and Massague, Joan (2007) "Lung metastasis genes couple breast tumor size and metastatic spread", *Proceedings of the National Academy of Sciences*, **104**(16):6740-6745

Examples

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

```
annotation(vdx)
## show the abstract for this dataset
abstract(vdx)
```

Index

∗ **datasets**
vdx, 2

vdx, 2