CGHbase

October 5, 2010

CGHbase-package *CGHbase: Base functions and classes for arrayCGH data analysis.*

Description

CGHbase: Base functions and classes for arrayCGH data analysis.

Details

Main infrastructural classes: `cghRaw`, `cghSeg`, `cghCall`. Full help on methods and associated functions is available from withing class help pages.

Attached data sets: `WiltingData`, `WiltingRaw`, `WiltingNorm`, `WiltingSeg`, `WiltingCalled`.

Author(s)

Sjoerd Vosse <sjoerdvos@yahoo.com>

WiltingCalled *Cervical cancer arrayCGH data called with CGHcall*

Description

Cervical cancer arrayCGH data called with `CGHcall` with default settings, containing 3552 features for 5 samples.

Usage

```
WiltingCalled
```

Format

An object of class `cghCall`

Source

Wilting, S.M., Snijders, P.J., Meijer, G.A., Ylstra, B., van den IJssel, P.R., Snijders, A.M., Albertson, D.G., Coffa, J., Schouten, J.P., van de Wiel, M.A., Meijer, C.J., & Steenbergen, R.D. (2006). Increased gene copy numbers at chromosome 20q are frequent in both squamous cell carcinomas and adenocarcinomas of the cervix. *Journal of Pathology, 210*, 258-259.

Mark A. van de Wiel, Kyung In Kim, Sjoerd J. Vosse, Wessel N. van Wieringen, Saskia M. Wilting and Bauke Ylstra. CGHcall: calling aberrations for array CGH tumor profiles. *Bioinformatics, 23*, 892-894.

`WiltingData` *Cervical cancer arrayCGH data*

Description

A dataframe containing 4709 rows and 8 columns with arrayCGH data.

Usage

```
Wilting
```

Format

A dataframe containing the following 8 columns:

Name The unique identifiers of array elements.

Chromosome Chromosome number of each array element.

Position Chromosomal position in bp of each array element.

AdCA10 Raw log2 ratios for cervical cancer sample AdCA10.

SCC27 Raw log2 ratios for cervical cancer sample SCC27.

SCC32 Raw log2 ratios for cervical cancer sample SCC32.

SCC36 Raw log2 ratios for cervical cancer sample SCC36.

SCC39 Raw log2 ratios for cervical cancer sample SCC39.

Source

Wilting, S.M., Snijders, P.J., Meijer, G.A., Ylstra, B., van den IJssel, P.R., Snijders, A.M., Albertson, D.G., Coffa, J., Schouten, J.P., van de Wiel, M.A., Meijer, C.J., & Steenbergen, R.D. (2006). Increased gene copy numbers at chromosome 20q are frequent in both squamous cell carcinomas and adenocarcinomas of the cervix. *Journal of Pathology, 210*, 258-259.

`WiltingNorm` *Normalized log2 ratios from cervical cancer arrayCGH data.*

Description

Normalized log2 ratios frm cervical cancer arrayCGH data, containing 3552 features for 5 samples. These data have been normalized using the `normalize` function with default settings.

Usage

```
WiltingCalled
```

Format

An object of class `cghRaw`.

Source

Wilting, S.M., Snijders, P.J., Meijer, G.A., Ylstra, B., van den IJssel, P.R., Snijders, A.M., Albertson, D.G., Coffa, J., Schouten, J.P., van de Wiel, M.A., Meijer, C.J., & Steenbergen, R.D. (2006). Increased gene copy numbers at chromosome 20q are frequent in both squamous cell carcinomas and adenocarcinomas of the cervix. *Journal of Pathology, 210*, 258-259.

`WiltingRaw` *Raw log2 ratios from cervical cancer arrayCGH data.*

Description

Raw log2 ratios from cervical cancer arrayCGH data, containing 3552 features for 5 samples. These data have been preprocessed using `preprocess`.

Usage

```
WiltingCalled
```

Format

An object of class `cghRaw`.

Source

Wilting, S.M., Snijders, P.J., Meijer, G.A., Ylstra, B., van den IJssel, P.R., Snijders, A.M., Albertson, D.G., Coffa, J., Schouten, J.P., van de Wiel, M.A., Meijer, C.J., & Steenbergen, R.D. (2006). Increased gene copy numbers at chromosome 20q are frequent in both squamous cell carcinomas and adenocarcinomas of the cervix. *Journal of Pathology, 210*, 258-259.

WiltingRegions — *Regions of cervical cancer arrayCGH data as defined by CGHregions*

Description

Regions of cervical cancer arrayCGH data as defined by `CGHregions` with default settings, containing 90 regions over 5 samples.

Usage

```
WiltingRegions
```

Format

An object of class `cghRegions`

Source

Wilting, S.M., Snijders, P.J., Meijer, G.A., Ylstra, B., van den IJssel, P.R., Snijders, A.M., Albertson, D.G., Coffa, J., Schouten, J.P., van de Wiel, M.A., Meijer, C.J., & Steenbergen, R.D. (2006). Increased gene copy numbers at chromosome 20q are frequent in both squamous cell carcinomas and adenocarcinomas of the cervix. *Journal of Pathology, 210*, 258-259.

Mark A. van de Wiel and Wessel N. van Wieringen (2007). CGHregions: Dimension Reduction for Array CGH Data with Minimal Information Loss. *Cancer Informatics, 2*, 55-63.

WiltingSeg — *Segmented log2 ratios from cervical cancer arrayCGH data.*

Description

Segmented log2 ratios from cervical cancer arrayCGH data, containing 3552 features for 5 samples. These data have been segmented using `segmentData` with default settings.

Usage

```
WiltingCalled
```

Format

An object of class `cghSeg`.

Source

Wilting, S.M., Snijders, P.J., Meijer, G.A., Ylstra, B., van den IJssel, P.R., Snijders, A.M., Albertson, D.G., Coffa, J., Schouten, J.P., van de Wiel, M.A., Meijer, C.J., & Steenbergen, R.D. (2006). Increased gene copy numbers at chromosome 20q are frequent in both squamous cell carcinomas and adenocarcinomas of the cervix. *Journal of Pathology, 210*, 258-259.

avedist *Retrieve regions information from cghRegions object.*

Description

This function accesses the regions information stored in the featureData of an object derived from the `cghRegions-class`.

Usage

```
avedist(object)
nclone(object)
```

Arguments

`object`	Object derived from class `cghRegions`

Value

`avedist` returns a vector containing the Average L1-distance of clone signatures to the medoid signature; `nclone` returns a vector containing the number of clones that is included in each region;

Author(s)

Sjoerd Vosse

See Also

`cghRegions-class`

chromosomes *Retrieve feature position data from cgh objects.*

Description

These generic functions access the position data stored in the featureData of an object derived from the `cghRaw-class`, `cghSeg-class` or `cghCall-class`.

Usage

```
chromosomes(object)
bpstart(object)
bpend(object)
```

Arguments

`object`	Object derived from class `cghRaw`, `cghSeg`, or `cghCall`

Value

chromosomes returns a vector of chromosome numbers; bpstart returns a vector of basepair start positions; bpend returns a vector of basepair end positions;

Author(s)

Sjoerd Vosse

See Also

cghRaw-class, cghSeg-class, cghCall-class

---

cghCall *Class to contain and describe called array comparative genomic hybridization data.*

Description

Container for aCGH data and experimental metadata. cghCall class is derived from eSet, and requires the following matrices of equal dimension as assayData members:

- copynumber Log2 copynumber ratios.
- segmented Segmented log2 ratios.
- calls Called copynumer values.
- probloss Loss probabilities as returned by CGHcall.
- probnorm Normal probabilities as returned by CGHcall.
- probgain Gain probabilities as returned by CGHcall.
- probamp Optional amplification probabilities as returned by CGHcall when run with nclass=4.

Furthermore, columns named Chromosome, Start, and End are required as featureData members, containing feature position information.

Extends

Directly extends class eSet.

Creating Objects

```
new('cghCall', phenoData = [AnnotatedDataFrame], experimentData = [MIAME],
annotation = [character], copynumber = [matrix], segmented = [matrix],
calls = [matrix], probloss = [matrix], probnorm = [matrix], probgain
= [matrix], featureData = [AnnotatedDataFrame], ...)
```

An object of class cghCall is generally obtained as output from CGHcall.

Slots

Inherited from eSet:

- assayData: Contains matrices with equal dimensions, and with column number equal to nrow(phenoData). assayData must contain the following matrices
  - copynumber Log2 copynumber ratios.
  - segmented Segmented log2 ratios.
  - calls Called copynumer values.
  - probloss Loss probabilities as returned by CGHcall.
  - probnorm Normal probabilities as returned by CGHcall.
  - probgain Gain probabilities as returned by CGHcall.
  - probamp Optional amplification probabilities as returned by CGHcall when run with nclass=4.

  with rows represening array probes and columns representing samples. Additional matrices of identical size (e.g., representing measurement errors) may also be included in assayData. Class:AssayData-class
- phenoData: See eSet
- featureData: An AnnotatedDataFrame with columns Chromosome, Start, and End containing array element position data.
- experimentData: See eSet
- annotation: See eSet

Methods

Class-specific methods.

- copynumber(cghCall), copynumber(cghCall,matrix)<- Access and set elements named copynumber in the AssayData-class slot.
- segmented(cghCall), segmented(cghCall,matrix)<- Access and set elements named segmented in the AssayData-class slot.
- calls(cghCall), calls(cghCall,matrix)<- Access and set elements named calls in the AssayData-class slot.
- probloss(cghCall), probloss(cghCall,matrix)<- Access and set elements named probloss in the AssayData-class slot.
- probnorm(cghCall), probnorm(cghCall,matrix)<- Access and set elements named probnorm in the AssayData-class slot.
- probgain(cghCall), probgain(cghCall,matrix)<- Access and set elements named probgain in the AssayData-class slot.
- chromosomes, bpstart, bpend Access the chromosomal positions stored in featureData
- **plot.cghCall** Create a plot containing log2ratios, segments and call probabilities ordered by chromosomal position
- **summaryPlot** Create a plot summarizing the call probabilities of all samples

See eSet for derived methods.

Author(s)

Sjoerd Vosse

See Also

eSet-class, cghRaw-class, cghSeg-class

Examples

```

# create an instance of cghCall
new("cghCall")

# load an instance of cghCall
data(WiltingCalled)

  # plot the first sample
  plot.cghCall(WiltingCalled[,1])

  # plot the first chromosome of the first sample
  plot.cghCall(WiltingCalled[chromosomes(WiltingCalled)==1,1])

  # get the copynumber values of the third and fourth sample
  log2ratios <- copynumber(WiltingCalled[,3:4])

  # get the names of the samples
  sampleNames(WiltingCalled)

  # get the names of the array elements
  featureNames(WiltingCalled)
```

cghRaw — *Class to contain and describe raw or normalized array comparative genomic hybridization data.*

Description

Container for aCGH data and experimental metadata. `cghRaw` class is derived from `eSet`, and requires a matrix named `copynumber` as assayData member. Furthermore, columns named `Chromosome`, `Start`, and `End` are required as featureData members, containing feature position information.

Extends

Directly extends class `eSet`.

Creating Objects

```
new('cghRaw', phenoData = [AnnotatedDataFrame], experimentData = [MIAME],
annotation = [character], copynumber = [matrix], featureData = [AnnotatedDataFra
...)
```

`make_cghRaw` is a function to convert a dataframe or textfile to an object of class `cghRaw`. The input should be either a dataframe or a tabseparated textfile (textfiles must contain a header). The first three columns should contain the name, chromosome and position in bp for each array target respectively. The chromosome and position column must contain numbers only. Following these is a column with log2 ratios for each of your samples. If the input type is a textfile, missing values should be represented as 'NA' or an empty field.

Slots

Inherited from `eSet`:

`assayData`**:** Contains matrices with equal dimensions, and with column number equal to `nrow(phenoData)`. `assayData` must contain a matrix `copynumber` with rows represening array probes and columns representing samples. Additional matrices of identical size (e.g., representing measurement errors) may also be included in `assayData`. Class:`AssayData-class`

`phenoData`**:** See `eSet`

`featureData`**:** An `AnnotatedDataFrame` with columns `Chromosome`, `Start`, and `End` containing array element position data.

`experimentData`**:** See `eSet`

`annotation`**:** See `eSet`

Methods

Class-specific methods.

`copynumber(cghRaw)`, `copynumber(cghRaw,matrix)<-` Access and set elements named `copynumber` in the `AssayData-class` slot.

`chromosomes`, `bpstart`, `bpend` Access the chromosomal positions stored in `featureData`

plot.cghRaw Create a plot containing log2ratios ordered by chromosomal position

See `eSet` for derived methods. Annotation functionality is not yet supported.

Author(s)

Sjoerd Vosse

See Also

`eSet-class`, `cghSeg-class`, `cghCall-class`

Examples

```
# create an instance of cghRaw
new("cghRaw")

# create an instance of cghRaw from a dataframe
data(WiltingData)
cghobj <- cghRaw(WiltingData)

# plot the first sample
plot.cghRaw(cghobj[,1])
# first three chromosomes
plot.cghRaw(cghobj[chromosomes(cghobj)==1,1])

# get the copynumber values of the third and fourth sample
log2ratios <- copynumber(cghobj[,3:4])

# get the names of the samples
sampleNames(cghobj)

# get the names of the array elements
featureNames(cghobj)
```

cghRegions *Class to contain and describe array comparative genomic hybridization regions data.*

Description

Container for aCGH regions data and experimental metadata. cghRegions class is derived from eSet, and requires a matrix named regions as assayData member. Furthermore, columns named Chromosome, Start, End, Nclone, and Avedist are required as featureData members, containing region and position information.

Extends

Directly extends class eSet.

Creating Objects

```
new('cghRegions', phenoData = [AnnotatedDataFrame], experimentData
= [MIAME], annotation = [character], regions = [matrix], featureData
= [AnnotatedDataFrame], ...)
```

An object of this class is generally obtained by running the function CGHregions.

Slots

Inherited from eSet:

assayData**:** Contains matrices with equal dimensions, and with column number equal to nrow(phenoData). assayData must contain a matrix regions with rows represening regions and columns representing samples. Additional matrices of identical size (e.g., representing measurement errors) may also be included in assayData. Class:AssayData

phenoData**:** See eSet

featureData**:** An AnnotatedDataFrame with columns Chromosome, Start, End, Nclone, and Avedist containing region and position information.

experimentData**:** See eSet

annotation**:** See eSet

Methods

Class-specific methods.

regions(cghRegions), regions(cghRegions,matrix)<- Access and set elements named regions in the AssayData-class slot.

chromosomes, bpstart, bpend, nclone, avedist Access the region and position information stored in featureData

plot.cghRegions Create a plot displaying chromosomes on the Y-axis and base pair position on the X-axis. A new region is displayed by a slight jump with respect to the previous region. Each region is displayed as a bi-colored segment, the lower and upper part of which correspond to the proportions pl and pg of samples with a loss (red) or gain (green), respectively. The color coding is displayed as well: 1: pl (pg) < 10%; 2: 10% = pl (pg) < 30%; 3:30% = pl (pg) < 50%; 4: pl (pg) = 50%.

frequencyPlot Create a frequency plot

See `eSet` for derived methods. Annotation functionality is not yet supported.

Author(s)

Sjoerd Vosse

See Also

`eSet`, `cghRaw`, `cghSeg`, `cghCall`

Examples

```
# create an instance of cghRegions
new("cghRegions")

# load an instance of cghRegions
data(WiltingRegions)

# plot all region data
plot.cghRegions(WiltingRegions)
# make a frequency plot
frequencyPlot(WiltingRegions)

# extract the region values
values <- regions(WiltingRegions)

# get the names of the samples
sampleNames(WiltingRegions)
```

`cghSeg` *Class to contain and describe segmented array comparative genomic hybridization data.*

Description

Container for aCGH data and experimental metadata. `cghSeg` class is derived from `eSet`, and requires a matrix named `copynumber` as well as a matrix named `segmented` as assayData members of equal dimensions. Furthermore, columns named `Chromosome`, `Start`, and `End` are required as featureData members, containing feature position information.

Extends

Directly extends class `eSet`.

Creating Objects

```
new('cghSeg', phenoData = [AnnotatedDataFrame], experimentData = [MIAME],
annotation = [character], copynumber = [matrix], segmented = [matrix],
featureData = [AnnotatedDataFrame], ...)
```

An object of class `cghSeg` is generally obtained as output from `segmentData`.

Slots

Inherited from `eSet`:

`assayData`: Contains matrices with equal dimensions, and with column number equal to `nrow(phenoData)`. `assayData` must contain matrices `copynumber` and `segmented` with rows represening array probes and columns representing samples. Additional matrices of identical size (e.g., representing measurement errors) may also be included in `assayData`. Class:`AssayData-class`

`phenoData`: See `eSet`

`featureData`: An `AnnotatedDataFrame` with columns `Chromosome`, `Start`, and `End` containing array element position data.

`experimentData`: See `eSet`

`annotation`: See `eSet`

Methods

Class-specific methods.

`copynumber(cghSeg)`, `copynumber(cghSeg,matrix)<-` Access and set elements named `copynumber` in the `AssayData-class` slot.

`segmented(cghSeg)`, `segmented(cghSeg,matrix)<-` Access and set elements named `segmented` in the `AssayData-class` slot.

`chromosomes`, `bpstart`, `bpend` Access the chromosomal positions stored in `featureData`

plot.cghSeg Create a plot containing log2ratios and segments ordered by chromosomal position

See `eSet` for derived methods.

Author(s)

Sjoerd Vosse

See Also

`eSet-class`, `ExpressionSet-class`, `cghRaw-class`, `cghCall-class`

Examples

```
# create an instance of cghSeg
new("cghSeg")

# load an instance of cghSeg
data(WiltingSeg)

# plot the first sample
# plot.cghSeg(WiltingSeg[,1])
# first three chromosomes
# plot.cghSeg(WiltingSeg[chromosomes(WiltingSeg)==1,1])

# get the copynumber values of the third and fourth sample
log2ratios <- copynumber(WiltingSeg[,3:4])

# get the names of the samples
```

```
sampleNames(WiltingSeg)

# get the names of the array elements
featureNames(WiltingSeg)
```

copynumber *Retrieve copynumber data from cgh objects.*

Description

These generic functions access the copynumber values of assay data stored in an object derived from the cghRaw-class, cghSeg-class or cghCall-class.

Usage

```
copynumber(object)
copynumber(object) <- value
segmented(object)
segmented(object) <- value
calls(object)
calls(object) <- value
```

Arguments

object	Object derived from class cghRaw, cghSeg, or cghCall
value	Matrix with rows representing features and columns samples.

Value

copynumber returns a matrix of copynumber values;

Author(s)

Sjoerd Vosse

See Also

cghRaw-class, cghSeg-class, cghCall-class

Examples

```
data(WiltingCalled)
log2ratios <- copynumber(WiltingCalled)
segments <- segmented(WiltingCalled)
calls <- calls(WiltingCalled)
```

frequencyPlot *Visualization of aCGH regions.*

Description

This function creates a frequency plot for aCGH regions.

Usage

```
frequencyPlot(x, y, ...)
```

Arguments

`x`	An object of class `cghRegions`.
`y`	This argument is not used and should be missing.
`...`	Arguments `plot`.

Details

We find plotted on the x-axis the array probes sorted by chromosomal position. The vertical bars represent the frequency of gains and losses across your samples. The black bars represent gains, the gray bars represent losses.

Value

This function creates a plot.

Author(s)

Mark van de Wiel and Sjoerd Vosse

References

Mark A. van de Wiel and Wessel N. van Wieringen (2007). CGHregions: Dimension Reduction for Array CGH Data with Minimal Information Loss. *Cancer Informatics, 2*, 55-63.

Examples

```
## Not run:
data(WiltingRegions)
frequencyPlot(WiltingRegions)

## End(Not run)
```

`plot.cghRaw` *Plot aCGH data.*

Description

Please see the class descriptions for more details on the plot functions.

Usage

```
plot.cghRaw(x, y, ...)
plot.cghSeg(x, y, ...)
plot.cghCall(x, y, ...)
plot.cghRegions(x, y, ...)
```

Arguments

`x`	An object of class `cghRaw`, `cghSeg`, `cghCall`, or `cghSeg`.
`y`	This argument is not used and should be missing.
`...`	Arguments `plot`.

Author(s)

Sjoerd Vosse

See Also

`cghRaw-class`, `cghSeg-class`, `cghCall-class`, `cghRegions-class`

`probloss` *Retrieve call probabilities from a cghCall object.*

Description

These generic functions access the call probabilities from assay data stored in a object derived from the `cghCall-class`.

Usage

```
probloss(object)
probloss(object) <- value
probnorm(object)
probnorm(object) <- value
probgain(object)
probgain(object) <- value
probamp(object)
probamp(object) <- value
```

Arguments

`object`	Object derived from class `cghCall`
`value`	Matrix with rows representing features and columns samples.

Value

These functions return matrices of call probabilities.

Author(s)

Sjoerd Vosse

See Also

`cghCall-class`

`regions` *Retrieve regions data from cghRegions object.*

Description

This function accesses the regions values of assay data stored in an object derived from the `cghRegions-class`.

Usage

```
regions(object)
regions(object) <- value
```

Arguments

`object`	Object derived from class `cghRegions`
`value`	Matrix with rows representing features and columns samples.

Value

`regions` returns a matrix of regions values;

Author(s)

Sjoerd Vosse

See Also

`cghRegions-class`

summaryPlot *Visualization of aCGH profiles.*

Description

This function creates a summary plot for aCGH profiles.

Usage

```
summaryPlot(x, y, ...)
```

Arguments

x	An object of class cghCall.
y	This argument is not used and should be missing.
...	Arguments plot.

Details

We find plotted on the x-axis the array probes sorted by chromosomal position. The vertical bars represent the average probability that the positions they cover are gained (green bars) or lost (red bars). The green bars represent gains, the red bars represent losses. When 4 levels have been used for calling, amplifications are indicated with a blue tickmark at the top of the plot.

Value

This function creates a plot.

Author(s)

Sjoerd Vosse & Mark van de Wiel

References

Mark A. van de Wiel, Kyung In Kim, Sjoerd J. Vosse, Wessel N. van Wieringen, Saskia M. Wilting and Bauke Ylstra. CGHcall: calling aberrations for array CGH tumor profiles. *Bioinformatics, 23*, 892-894.

Examples

```
 ## Not run:
   data(Wilting)
   rawcgh <- make_cghSeg(Wilting)
   normalized <- normalize(rawcgh)
   segmented <- segmentData(normalized)
   called <- CGHcall(segmented)
   summaryPlot(called)

## End(Not run)
```

Index