# ChromHeatMap

April 19, 2010

---

`ALLs.chr22` *Chromosome 22 subset of ALL data for ALL1/AF4 and E2A/PBX1*

---

## Description

This is a greatly reduced subset of the Chiaretti et al. ALL data set (available in its entirety as the Bioconductor ALL package). The data in this subset consist of microarrays from 15 different individuals with acute lymphoblastic leukemia (ALL). The data are further restricted to chromosome 22 only. This data set is intended for demonstration purposes only.

## Usage

```
ALLs.chr22
```

## Format

An ExpressionSet with the following covariates:

- ageThe age of the patient in years.
- mol.biolThe assigned molecular biology of the cancer (mainly for those with B-cell ALL). In this data set this is restricted to ALL1/AF4 and E2A/PBX1.

## Source

The ALL Bioconductor data package

## References

Sabina Chiaretti, Xiaochun Li, Robert Gentleman, Antonella Vitale, Marco Vignetti, Franco Mandelli, Jerome Ritz, and Robin Foa Gene expression profile of adult T-cell acute lymphocytic leukemia identifies distinct subsets of patients with different response to therapy and survival. Blood, 1 April 2004, Vol. 103, No. 7.

## chrdata — *The ALLs.chr22 ExpressionSet, reformatted as a ChrStrandData object*

### Description

This is a greatly reduced subset of the Chiaretti et al. ALL data set (available in its entirety as the Bioconductor ALL package). The data in this subset consist of microarrays from 15 different individuals with acute lymphoblastic leukemia (ALL). The data are further restricted to chromosome 22 only. This data set is intended for demonstration purposes only. See the documentation for `makeChrStrandData` for a description of the ChrStrandData object format. This format directly associates the ExpressionSet data with chromosome location, speeding up retrieval of data during heat map plotting.

### Usage

```
chrdata
```

### Format

A ChrStrandData object

### Source

The ALL Bioconductor data package

### References

Sabina Chiaretti, Xiaochun Li, Robert Gentleman, Antonella Vitale, Marco Vignetti, Franco Mandelli, Jerome Ritz, and Robin Foa Gene expression profile of adult T-cell acute lymphocytic leukemia identifies distinct subsets of patients with different response to therapy and survival. Blood, 1 April 2004, Vol. 103, No. 7.

## chrHeatMap — *Plot ChrStrandMatrix objects as heat maps along a chromosome*

### Description

Plots a either one or two ChrStrandMatrix objects (typically constructed using the `createChrMatrix` function) as heat maps along a specified chromosome, optionally clustering samples and including an idiogram.

### Usage

```
chrHeatMap (strand.data, cytopaint.func=NULL, col = "heat.colors",
            start, end, breaks, RowSideColors, title=TRUE,
            margins = c(6, 6), cexCyto = 0.8, srtCyto=90, lmat = NULL, lhei = NU
            lwid = NULL, ...)
```

## Arguments

| | |
|---|---|
| strand.data | A ChrStrandMatrix object, or a list of such objects, one per strand to be plotted (or a single matrix for 'both' strands), created using the createChrMatrix function. |
| cytopaint.func | A function closure taking a single argument, 'boxwidth', and plotting its enclosed idiogram data at that width. See plotChrMap for the code used to generate this closure. |
| col | A vector of colors to use for the heat map, or the name of a function generating such a vector. |
| start | The starting genome coordinate for the plot. |
| end | The ending genome coordinate for the plot. |
| breaks | A vector of numeric break points indicating the boundaries between the col colors. |
| RowSideColors | A vector of colors to use for a color band indicating e.g. sample categories. |
| title | If TRUE, this causes the function to include default heat map subtitles indicating which chromosome and strand has been plotted. If FALSE or NULL, subtitles will left blank. If this argument is set to a character vector of the same length and order as strand.data its contents will be used as heat map subtitles. |
| margins | A numeric vector indicating the c(bottom, left) margins of the plot containing X and Y axes labels. |
| cexCyto | A positive number used to control the font size for the idiogram plot. For plots spanning just a few cytobands it may be worth setting this to a larger number, and srtCyto, below, to zero. |
| srtCyto | A number indicating the degree to which the idiogram text labels should be rotated. This defaults to 90 degrees, but for more detailed plots a setting of zero here often looks better. |
| lmat | An optional matrix to be passed to layout. |
| lhei | An optional vector of layout row heights. |
| lwid | An optional vector of layout row widths. |
| ... | Additional arguments are passed to the drawMapDendro function. |

## Details

Typically this function should not be called directly, but rather via the wrapper plotChrMap function. This function uses cytoband data from the UCSC genome annotation database and code adapted from the quantsmooth package to draw an idiogram of the chromosome, or a subset thereof.

## Value

This function is executed for its side effects.

## Author(s)

Tim F Rayner

## References

lodplot and quantsmooth packages

## See Also

plotChrMap, createChrMatrix, drawMapDendro

## Examples

```
data('demo')
stranddata <- createChrMatrix( chrdata, chr=22, strand='forward', start=21925000, end=243
chrHeatMap(stranddata)
```

---

# ChrMapPlot *Class containing a mapping between plot location and probe identifier.*

---

## Description

ChrMapPlot objects are generated as an output from the main `plotChrMap` function, which users can then pass to the `grabChrMapProbes` function.

## Creating Objects

Objects of this class are created using the `plotChrMap` function:

```
plotChrMap(chrdata, '22')
```

## Slots

**labels** An array of probe identifiers, with names corresponding to chromosome coordinates.

**start** The leftmost interval number (most usually 1).

**end** The rightmost interval number.

## Methods

Standard generic methods:

`show(ChrMapPlot)` Generates a short description of the ChrMapPlot object.

## Author(s)

Tim F Rayner

## See Also

plotChrMap, grabChrMapProbes.

## Examples

```
data('demo')
plotmap <- plotChrMap(chrdata, '22', cytoband='q11.23')
probes <- grabChrMapProbes(plotmap)
library('hgu95av2.db')
genes <- mget(probes, hgu95av2SYMBOL, ifnotfound=NA)
```

chrNames *Retrieve chromosome names from an object.*

## Description

This generic function simply returns the names of all the chromosomes represented by a given ChrStrandData or ChrStrandMatrix object. Note that not every sample associated with a ChrStrandData object need have data from every chromosome.

## Usage

```
chrNames(object)
```

## Arguments

| | |
|---|---|
| object | Object derived from class ChrStrandData or ChrStrandMatrix |

## Value

`chrNames(object)` returns a character vector listing the chromosomes.

## Author(s)

Tim F Rayner

## See Also

ChrStrandData-class

ChrStrandData *Class to contain data associated with chromosome coordinates across a whole genome.*

## Description

Container for data from high-throughput assays mapped to chromosome locations.

## Creating Objects

The most convenient way to create a ChrStrandData object is to use the `makeChrStrandData` function, which can be used to convert data stored in either an `ExpressionSet` or data frame into a `ChrStrandData` object:

```
makeChrStrandData(ALL, lib = "hgu95av2.db")
```

### Slots

**data** a 'list', whose components correspond to samples in the same order as appearing in the columns of 'expr'. Each component is also a 'list', named by chromosomes "1"-"22", "X" and "Y". Each named component is again a 'list' with two elements named "posS" and "negS", corresponding to the forward and reverse strands of a chromosome, each of which is a list containing start coordinates ("x"), end coordinates("xe") and the corresponding data values ("y").

**lib** A string giving the name of the annotation data package to use.

**chrs** The list of chromosomes represented in the object.

### Methods

Class-specific methods.

`annotation(ChrStrandData)` Returns the name of the AnnotationDbi library used to annotate the object.

`chrNames(ChrStrandData)` Returns a list of the chromosomes represented in the object.

`sampleNames(ChrStrandData)` Returns the names of the samples associated with the object.

Standard generic methods:

`show(ChrStrandData)` Generates a short description of the ChrStrandData object.

`summary(ChrStrandData)` Generates a summary of the data available for each chromosome in the ChrStrandData object.

### Author(s)

Tim F Rayner

### See Also

`makeChrStrandData`, `ChrStrandMatrix-class`.

### Examples

```
data('demo')
chrdata <- makeChrStrandData(exprs(ALLs.chr22), lib = "hgu95av2.db")
```

---

## ChrStrandMatrix *Class to contain data associated with genome locations for a specific chromosome.*

---

### Description

Container for chromosome-specific subsets of data selected from an genome-wide ChrStrandData object, suitable for use with `chrHeatMap`.

### Creating Objects

Typically, objects of this class are created and used internally by the `createChrMatrix` and `chrHeatMap` functions. Objects can be created in a similar fashion by end-users:

```
createChrMatrix(chrdata, chr=22, strand='forward', start=21925000,
end=24300000, interval=5000)
```

Note that this function may combine data from multiple probes (taking the mean) into a single chromosomal locus based on the size of the specified interval. If this happens the combined probe identifiers are concatenated in the output object, separated by a semicolon.

### Slots

**data** The data matrix, arranged with samples in columns and genomic locations in rows.

**probeID** An array of probe identifiers associated with the data. The names attached to this array correspond with chromosome coordinate (specifically, the starting coordinates, i.e. the left-hand edges). These identifiers will ultimately be returned by e.g. the `grabChrMapProbes` function.

**chr** The chromosome name or number.

**strand** The chromosome strand ('forward', 'reverse' or 'both').

**start** The starting chromosome coordinates for each genomic location.

**end** The ending chromosome coordinates for each genomic location.

### Methods

Class-specific methods.

`chrNames(ChrStrandMatrix)` Returns the name of the chromosome for the object.

`strandName(ChrStrandMatrix)` Returns the chromosome strand for the object.

`sampleNames(ChrStrandMatrix)` Returns the names of the samples associated with the object.

`featureNames(ChrStrandMatrix)` Returns the probe identifiers associated with the object.

`exprs(ChrStrandMatrix)` Returns the chromosome-specific data matrix for the object.

Standard generic methods:

`show(ChrStrandMatrix)` Generates a short description of the ChrStrandMatrix object.

`summary(ChrStrandMatrix)` Generates a summary of the data available for each sample in the ChrStrandMatrix object.

### Author(s)

Tim F Rayner

### See Also

`createChrMatrix`, `ChrStrandData-class`.

### Examples

```
data('demo')
stranddata <- createChrMatrix( chrdata, chr=22, strand='forward', start=21925000, end=243
```

`createChrMatrix` *Generate chromosome-based subset matrices from the mapped data structures generated by makeChrStrandData*

## Description

Given a data object from makeChrStrandData, generate a matrix containing a subset of the data from a given region of a given chromosome strand, with data binned at appropriate intervals along the chromosome. The minimum width of the binning interval is controlled using the "interval" argument, which can therefore be used to control the output resolution of the data.

## Usage

```
createChrMatrix(data, chr, strand = c('forward','reverse','both'), subset = NULL
                start=1, end, interval=ceiling((end - start)/500))
```

## Arguments

| | |
|---|---|
| `data` | A ChrStrandData object (e.g. generated by `makeChrStrandData`). |
| `chr` | The name of the chromosome to plot. |
| `strand` | The chromosome strand to plot ('both' indicates that both strands should be overlaid in a single heatmap). |
| `subset` | An optional numeric vector indicating which samples should be plotted. |
| `start` | The starting chromosome coordinate from which to plot. |
| `end` | The ending chromosome coordinate. |
| `interval` | The (optional) size of the data bins to use along the chromosome, in bases. |

## Details

Typically this function will not be called directly, but rather via the wrapper `plotChrMap` function.

Note that this function may combine data from multiple probes (taking the mean) into a single chromosomal locus based on the size of the specified interval. If this happens the combined probe identifiers are concatenated in the output object, separated by a semicolon.

## Value

A ChrStrandMatrix object suitable for use with `chrHeatMap` and `drawMapDendro`.

## Author(s)

Tim F Rayner

## See Also

`plotChrMap`, `chrHeatMap`, `drawMapDendro`, `ChrStrandMatrix-class`, `ChrStrandData-class`

## Examples

```
data('demo')
stranddata <- createChrMatrix( chrdata, chr=22, strand='forward', start=21925000, end=243
```

## cytobands *Cytoband location information*

### Description

This data set contains cytoband information for a range of species, taken directly from the UCSC genome annotation database. This data set is designed to be easily extendable to cover new species.

### Usage

```
cytobands
```

### Format

A list of data frames, one per species, each with one row per cytoband and the following columns:

- chrThe chromosome number for the cytoband, prefixed with 'chr'.
- startThe start coordinate for the cytoband.
- endThe end coordinate for the cytoband.
- bandThe cytoband number (i.e., the '23.3' in '1q23.3').
- stainThe cytoband stain (see the `stains` data set).
- armThe chromosome arm for the cytoband (i.e., the 'q' in '1q23.3').

The list names (i.e. `names(cytobands)`) should correspond to species names in the AnnotationDbi packages used.

### Source

The UCSC genome annotation database: http://hgdownload.cse.ucsc.edu/downloads.html

## drawMapDendro *Draw a heatmap and dendrogram for a strand-specific data matrix generated by createChrMatrix*

### Description

Given a data matrix, cluster by sample (if desired), and plot the dendrogram and heatmap along chromosome coordinates. This function reuses code from the gplots `heatmap.2` function. Note that this function makes assumptions about the current layout of the display device, and so should generally be called only via `plotChrMap`.

### Usage

```
drawMapDendro(x, start, end, col = "heat.colors", dendrogram = TRUE, Rowv = TRUE
              margins = c(6, 6), na.rm=TRUE, hclustfun = hclust, distfun = dist,
              breaks, RowSideColors, cexRow, cexCol,
              xlab, ylab, labRow, labCol, na.color = 'gray', ...)
```

## Arguments

| | |
|---|---|
| x | The strand-specific data matrix to cluster and plot, usually generated using createChrMatrix. |
| start | The starting genome coordinate for the plot. |
| end | The ending genome coordinate for the plot. |
| col | A character vector of colors to use in the heat map, or the name of a function generating such a vector. |
| dendrogram | A boolean flag indicating whether or not to draw the dendrogram. |
| Rowv | Determines if and how the sample dendrogram should be reordered. If a dendrogram, then it is used "as-is", i.e., without any reordering. If a vector of integers, then the dendrogram is computed and reordered based on the order of the vector. Set this argument to FALSE or NULL to draw the heatmap without any sample reordering. |
| margins | A numeric vector indicating the c(bottom, left) margins of the plot containing X and Y axes labels. |
| na.rm | Whether or not to remove NA from calculations. |
| hclustfun | Function used to compute the hierarchical clustering when Rowv is not a dendrogram object. Defaults to hclust. |
| distfun | Function used to compute the distance (dissimilarity) between both rows and columns. Defaults to dist. |
| breaks | (Optional) Either a numeric vector indicating the splitting points for binning x into colors, or a integer number of break points to be used, in which case the break points will be spaced equally between min(x) and max(x). |
| RowSideColors | (Optional) Character vector of length nrow(x) containing the color names for a vertical side bar that may be used to annotate the rows of x. |
| cexRow, cexCol | (Optional) Positive numbers, used as cex.axis in for the row or column axis labeling. If these arguments are omitted the function will try and calculate a sane axis font size based on the number of rows or columns respectively. |
| xlab, ylab | X- and Y- axis titles; defaults to none. |
| labRow, labCol | Character vectors with row and column labels to use; these default to rownames(x) or colnames(x), respectively. |
| na.color | Color to use for missing value (NA). Defaults to gray. |
| ... | Additional arguments are passed to the image function. |

## Details

This function makes assumptions about the plot layout, usually set by the enclosing chrHeatMap function. Typically neither of these functions should be called directly, but rather via the wrapper plotChrMap function.

## Value

This function is executed for its side effects.

## Author(s)

Tim F Rayner

### See Also

plotChrMap, createChrMatrix, chrHeatMap

### Examples

```
data('demo')
stranddata <- createChrMatrix( chrdata, chr=22, strand='forward',
  start=21925000, end=24300000 )
layout(matrix(1:2, ncol=2), widths=c(0.1,1))
drawMapDendro( stranddata, margins=c(0,0) )
```

---

## grabChrMapProbes *Identify the probes plotted using plotChrMap*

---

### Description

Allows the user to interactively select regions of the plotChrMap heatmap, identifying all the probes plotted in those regions.

### Usage

```
grabChrMapProbes( plotmap )
```

### Arguments

| | |
|---|---|
| plotmap | The output of the plotChrMap function. |

### Details

This function takes the output of the plotChrMap function and uses it to identify the probes responsible for the signals plotted on the plotChrMap heatmap. It asks the user to select two points on either side of the heatmap bands of interest (specifically, boundary for inclusion of a given band is its left-hand edge), and returns a vector of probe identifiers. This can be passed directly to AnnotationDbi::mget to yield gene symbols and other annotation.

Note that the plotting area layout() and par() values are not reset on exit, so that this function can be reused as many times as is desired.

### Value

A character vector of probe identifiers. If multiple probe identifiers have been averaged into a single band these identifiers will be string concatenated, separated by semicolons. In such cases the start, end and interval arguments to plotChrMap can be used to split the probes into separate bands.

### Author(s)

Tim F Rayner

### See Also

plotChrMap

### Examples

```
data('demo')
plotmap <- plotChrMap(chrdata, '22', cytoband='q11.23')
probes <- grabChrMapProbes(plotmap)
library('hgu95av2.db')
genes <- mget(probes, hgu95av2SYMBOL, ifnotfound=NA)
```

---

## makeChrStrandData-methods

*Map a data matrix onto chromosome coordinates*

---

### Description

Given a data matrix with row names corresponding to the probe IDs in an accompanying annotation package, returns a data structure that can be used with the `plotChrMap` function. Based on the Makesense method from the geneplotter package.

### Methods

**expr = "ExpressionSet"** Given an ExpressionSet object, returns a ChrStrandData object.

**expr = "matrix"** Given a matrix object (where `rownames(expr)` yields the probe identifiers), returns a ChrStrandData object.

---

## makeChrStrandData *Map a data matrix onto chromosome coordinates*

---

### Description

Given an ExpressionSet, or a data matrix with row names corresponding to the probe IDs in an accompanying annotation package, this function returns a data structure that can be used with the `plotChrMap` function. This code is based on the Makesense method from the geneplotter package, extended to use both the CHRLOC and CHRLOCEND annotation environments from recent AnnotationDbi packages.

### Usage

```
makeChrStrandData(expr, lib)
```

### Arguments

| | |
|---|---|
| `expr` | The ExpressionSet or data matrix to remap. |
| `lib` | The name of the annotation package to use. |

### Value

A ChrStrandData object suitable for use with plotChrMap.

### Author(s)

Tim F Rayner

### References

geneplotter, annotate and AnnotationDbi packages

### See Also

plotChrMap, ChrStrandData-class

### Examples

```
data('demo')
chrdata <- makeChrStrandData(exprs(ALLs.chr22), lib = "hgu95av2.db")
```

---

## makeRangedDataList *Plot expression data as tracks in the UCSC genome browser*

---

### Description

Creates a RangedDataList object suitable for uploading to the UCSC genome browser using the rtracklayer package.

### Usage

```
makeRangedDataList( data, chr, start = 1, end, genome, subset = NULL,
                    cytoband, plot=FALSE, session )
```

### Arguments

| | |
|---|---|
| data | A ChrStrandData object, output from the makeChrStrandData function. |
| chr | Chromosomal id, chromosome to plot 1:22,X,Y. |
| start | Optional start chromosome position from which to commence plotting. |
| end | Optional end chromosome position. |
| genome | The name of the genome from which the data coordinates are taken (e.g. "hg18"). Passed to GenomicData in the rtracklayer package. |
| subset | Optional numeric vector listing the samples from data to plot. |
| cytoband | Optional cytological band to plot (e.g. 'q23'). |
| plot | An optional flag indicating whether to automatically plot the resulting RangedDataList on the UCSC browser or not. |
| session | An optional rtracklayer UCSCSession object. Ignored unless plot=TRUE. |

### Details

This function is used to create RangedDataList objects from ChrStrandData objects (see the makeChrStrandData function). If the plot argument is set to TRUE, the data is also uploaded to a UCSC browser session using default settings. See the rtracklayer package for more information on RangedData and UCSCSession objects.

### Value

A RangedDataList object containing the data for the specified genome region. See the rtracklayer package for more information on this object class.

### Author(s)

Tim F Rayner

### References

rtracklayer package

### See Also

makeChrStrandData, RangedDataList plotChrMap,

### Examples

```
data('demo')
r <- makeRangedDataList( data=chrdata, chr=22, cytoband='q11.23', genome='hg18' )
```

---

## plotChrMap *Plot data as an annotated heat map along a chromosome*

---

### Description

Given a ChrStrandData object (produced by the makeChrStrandData function), this function plots a heat map of its data values along a specified chromosome, optionally clustering samples and including an idiogram.

### Usage

```
plotChrMap( data, chr, start = 1, end, subset = NULL,
            cytoband, interval = ceiling((end-start)/500),
            strands = c('forward', 'reverse'), ... )
```

### Arguments

| | |
|---|---|
| data | A ChrStrandData object, output from the makeChrStrandData function. |
| chr | Chromosomal id, chromosome to plot 1:22,X,Y. |
| start | Optional start chromosome position from which to commence plotting. |
| end | Optional end chromosome position. |
| subset | Optional numeric vector listing the samples from data to plot. |
| cytoband | Optional cytological band to plot (e.g. 'q23'). |
| interval | An optional interval size controlling the plot detail level. |
| strands | The chromosome strands to plot (a one- or two-element character vector, values 'forward', 'reverse', or 'both'). |
| ... | Additional arguments are passed to the chrHeatMap function. |

## Details

This function is used to plot ChrStrandData objects (the output of the `makeChrStrandData` function) as heatmaps arranged along genome coordinates. The default heat map will plot the entire forward strand for the chosen chromosome at the top of the figure, with an idiogram and the reverse strand below it. To plot both strands overlaid, use the `strands='both'` argument. Probe signals are averaged over a window size controlled by `interval`, such that the default length of each heat map segment is 1/500 the total heat map width. This can be varied as required to control the resolution of the plot. This function uses both the start and end chromosomal locations for each gene to plot heatmap positions, and as such will not work with older AnnotationDbi packages.

See the related functions from this package for further plotting arguments which may be passed to this function. In particular, see the `drawMapDendro` documentation for arguments used to control sample clustering and plot axis font sizes, and `chrHeatMap` for arguments relating to the idiogram plot. Note that the plotting area layout() and par() values are not reset on exit, so that `grabChrMapProbes` can be subsequently used on the output.

Idiogram plotting is currently only supported for data mapping to human, mouse and rat genomes. In principle this is extendable to any organism for which the UCSC genome browser includes cytoband information. Please contact the maintainer of this package for help in such cases.

## Value

A ChrMapPlot object containing a list of probe identifiers mapped to their corresponding display locations, for use with `grabChrMapProbes`.

## Author(s)

Tim F Rayner

## References

annotate package

## See Also

`drawMapDendro`, `chrHeatMap`, `makeChrStrandData`, `grabChrMapProbes`

## Examples

```
data('demo')
plotChrMap(chrdata, '22', cytoband='q11', labRow=ALLs.chr22$mol.biol,
cexCol=0.8, cexCyto=1.2, srtCyto=0)
```

---

# stains *Cytoband display information*

---

## Description

This is a data set describing the display parameters used to plot cytoband data.

## Usage

```
stains
```

### Format

A data frame with one row per cytoband type, and the following columns:

- typeThe cytoband type. This must correspond to the "stain" column in the cytoband data frame (see the `cytobands` documentation).
- bandcolThe shade of gray used to colour the cytobands. A number between 0 (black) and 1 (white). Passed as the "col" argument to `rect`.
- textcolThe shade of gray used for the cytoband text labels. A number between 0 (black) and 1 (white). Passed as the "col" argument to `text`.
- banddensThe shading density to use for the band colour. Passed as the "density" argument to `rect`.
- bandbordThe shade of gray used for the plotted cytoband borders. A number between 0 (black) and 1 (white). Passed as the "border" argument to `rect`.

### Source

Developed based on the design of the idiogram Bioconductor package

---

## `strandName` *Retrieve strand information from a ChrStrandMatrix object.*

---

### Description

This generic function simply returns the chromosome strand which name of all the chromosomes represented by a given ChrStrandData object. Note that not every sample associated with the object need have data from every chromosome.

### Usage

```
strandName(object)
```

### Arguments

| | |
|---|---|
| `object` | Object derived from class ChrStrandMatrix |

### Value

`strandName(object)` returns the name of the strand from which the object data is taken.

### Author(s)

Tim F Rayner

### See Also

`ChrStrandMatrix-class`

# Index