Package 'ASCAT'

November 19, 2015

Type Package

Title Allele-Specific Copy Number Analysis of Tumours

Version 2.4

Date 2015-09-22

Author Peter Van Loo [aut, cre], Gro Nielsen [ctb], Jiqiu Cheng [ctb], Stefan Dentro [ctb], Kerstin Haase [ctb]

Maintainer Kerstin Haase <Kerstin.Haase@crick.ac.uk>

Description R package of ASCAT as published in http://www.ncbi.nlm.nih.gov/pubmed/20837533

Depends R (>= 2.13.0),
RColorBrewer

License GPL-3

LazyLoad yes

R topics documented:

ascat.aspcf *ascat.aspcf*

Description

run ASPCF segmentation

Usage

```
ascat.aspcf(ASCATobj, selectsamples = 1:length(ASCATobj$samples),
  ascat.gg = NULL, penalty = 25)
```

Arguments

ASCATobj	an ASCAT object
selectsamples	a vector containing the sample number(s) to PCF. Default = all
ascat.gg	germline genotypes (NULL if germline data is available)
penalty	penalty of introducing an additional ASPCF breakpoint (expert parameter, don't adapt unless you know what you're doing)

Details

This function can be easily parallelised by controlling the selectsamples parameter
it saves the results in LogR_PCFed[sample]_[segment].txt and BAF_PCFed[sample]_[segment].txt; if these files exist, the results are read from them.
Hence, after parallelisation, copy all the files into the current directory, and run this function to read in the results

Value

output: ascat data structure containing:
1. Tumor_LogR data matrix
2. Tumor_BAF data matrix
3. Tumor_LogR_segmented: matrix of LogR segmented values
4. Tumor_BAF_segmented: list of BAF segmented values; each element in the list is a matrix containing the segmented values for one sample (only for probes that are germline homozygous)
5. Germline_LogR data matrix
6. Germline_BAF data matrix
7. SNPpos: position of all SNPs
8. ch: a list containing vectors with the indices for each chromosome (e.g. Tumor_LogR[ch[[13]],] will output the Tumor_LogR data of chromosome 13
9. chr: a list containing vectors with the indices for each distinct part that can be segmented separately (e.g. chromosome arm, stretch of DNA between gaps in the array design)

`ascat.GCcorrect` *ascat.GCcorrect*

Description

Corrects logR of the tumour sample(s) with genomic GC content

Usage

```
ascat.GCcorrect(ASCATobj, GCcontentfile = NULL)
```

Arguments

`ASCATobj`	an ASCAT object
`GCcontentfile`	File containing the GC content around every SNP for increasing window sizes

Details

Note that probes not present in the GCcontentfile will be lost from the results

Value

ASCAT object with corrected tumour logR

`ascat.loadData` *ascat.loadData*

Description

Function to read in SNP array data

Usage

```
ascat.loadData(Tumor_LogR_file, Tumor_BAF_file, Germline_LogR_file = NULL,
  Germline_BAF_file = NULL, chrs = c(1:22, "X", "Y"), gender = NULL,
  sexchromosomes = c("X", "Y"))
```

Arguments

`Tumor_LogR_file`	file containing logR of tumour sample(s)
`Tumor_BAF_file`	file containing BAF of tumour sample(s)
`Germline_LogR_file`	file containing logR of germline sample(s), NULL
`Germline_BAF_file`	file containing BAF of germline sample(s), NULL
`chrs`	a vector containing the names for the chromosomes (e.g. c(1:22,"X"))
`gender`	a vector of gender for each cases ("XX" or "XY"). Default = all female ("XX")
`sexchromosomes`	a vector containing the names for the sex chromosomes

Details

germline data files can be NULL - in that case these are not read in

Value

ascat data structure containing:
1. Tumor_LogR data matrix
2. Tumor_BAF data matrix
3. Tumor_LogR_segmented: placeholder, NULL
4. Tumor_BAF_segmented: placeholder, NULL
5. Germline_LogR data matrix
6. Germline_BAF data matrix
7. SNPpos: position of all SNPs
8. ch: a list containing vectors with the indices for each chromosome (e.g. Tumor_LogR[ch[[13]],] will output the Tumor_LogR data of chromosome 13
9. chr: a list containing vectors with the indices for each distinct part that can be segmented separately (e.g. chromosome arm, stretch of DNA between gaps in the array design)
10. gender: a vector of gender for each cases ("XX" or "XY"). Default = NULL: all female ("XX")

`ascat.plotAscatProfile`

ascat.plotAscatProfile

Description

ascat.plotAscatProfile

Usage

```
ascat.plotAscatProfile(n1all, n2all, heteroprobes, ploidy, rho, goodnessOfFit,
  nonaberrant, y_limit = 5, nAfull, ch, lrr, bafsegmented, chrs)
```

Arguments

`n1all`	copy number major allele
`n2all`	copy number minor allele
`heteroprobes`	probes with heterozygous germline
`ploidy`	ploidy of the sample
`rho`	purity of the sample
`goodnessOfFit`	estimated goodness of fit
`nonaberrant`	boolean flag denoting non-aberrated samples
`y_limit`	Optional parameter determining the size of the y axis in the nonrounded plot and ASCAT profile. Default=5
`nAfull`	copy number major allele
`ch`	a list containing c vectors, where c is the number of chromosomes and every vector contains all probe numbers per chromosome
`lrr`	(unsegmented) log R, in genomic sequence (all probes), with probe IDs

bafsegmented	B Allele Frequency, segmented, in genomic sequence (only probes heterozygous in germline), with probe IDs
chrs	a vector containing the names for the chromosomes (e.g. c(1:22,"X"))

Value

plot showing the ASCAT profile of the sample

`ascat.plotGenotypes` *ascat.plotGenotypes*

Description

ascat.plotGenotypes

Usage

```
ascat.plotGenotypes(ASCATobj, title, Tumor_BAF_noNA, Hom, ch_noNA)
```

Arguments

ASCATobj	an ASCAT object
title	main title of the plot
Tumor_BAF_noNA	B-allele frequencies of the tumour sample with removed NA values
Hom	Boolean vector denoting homozygous SNPs
ch_noNA	vector of probes per chromosome (NA values excluded)

Value

plot showing classified BAF per sample, with unused SNPs in green, germline homozygous SNPs in blue and all others in red

---

`ascat.plotNonRounded` *ascat.plotNonRounded*

Description

ascat.plotNonRounded

Usage

```
ascat.plotNonRounded(ploidy, rho, goodnessOfFit, nonaberrant, nAfull, nBfull,
  y_limit = 5, bafsegmented, ch, lrr)
```

Arguments

ploidy	ploidy of the sample
rho	purity of the sample
goodnessOfFit	estimated goodness of fit
nonaberrant	boolean flag denoting non-aberrated samples
nAfull	copy number major allele
nBfull	copy number minor allele
y_limit	Optional parameter determining the size of the y axis in the nonrounded plot and ASCAT profile. Default=5
bafsegmented	B Allele Frequency, segmented, in genomic sequence (only probes heterozygous in germline), with probe IDs
ch	a list containing c vectors, where c is the number of chromosomes and every vector contains all probe numbers per chromosome
lrr	(unsegmented) log R, in genomic sequence (all probes), with probe IDs

Value

plot showing the nonrounded copy number profile, using base plotting function

---

ascat.plotRawData *ascat.plotRawData*

Description

Plots SNP array data

Usage

```
ascat.plotRawData(ASCATobj)
```

Arguments

ASCATobj	an ASCAT object (e.g. data structure from ascat.loadData)

Value

Produces png files showing the logR and BAF values for tumour and germline samples

ascat.plotSegmentedData *ascat.plotSegmentedData*

Description

plots the SNP array data before and after segmentation

Usage

```
ascat.plotSegmentedData(ASCATobj)
```

Arguments

ASCATobj	an ASCAT object (e.g. from ascat.aspcf)

Value

png files showing raw and segmented tumour logR and BAF

ascat.plotSunrise *ascat.plotSunrise*

Description

ascat.plotSunrise

Usage

```
ascat.plotSunrise(d, psi_opt1, rho_opt1)
```

Arguments

d	distance matrix for a range of ploidy and tumour percentage values
psi_opt1	optimal ploidy
rho_opt1	optimal aberrant cell fraction

Value

plot visualising range of ploidy and tumour percentage values

ascat.predictGermlineGenotypes
ascat.predictGermlineGenotypes

Description

predicts the germline genotypes of samples for which no matched germline sample is available

Usage

```
ascat.predictGermlineGenotypes(ASCATobj, platform = "AffySNP6")
```

Arguments

ASCATobj	an ASCAT object
platform	used array platform

Details

Currently possible values for platform:
AffySNP6 (default)
Custom10k
Illumina109k
IlluminaCytoSNP
Illumina610k
Illumina660k
Illumina700k
Illumina1M
Illumina2.5M
IlluminaOmni5
Affy10k
Affy100k
Affy250k_sty
Affy250k_nsp
AffyOncoScan
AffyCytoScanHD

Value

predicted germline genotypes

ascat.runAscat *ascat.runAscat*

Description

ASCAT main function, calculating the allele-specific copy numbers

Usage

```
ascat.runAscat(ASCATobj, gamma = 0.55, pdfPlot = F, y_limit = 5,
  circos = NA, rho_manual = NA, psi_manual = NA)
```

Arguments

ASCATobj	an ASCAT object from ascat.aspcf
gamma	technology parameter, compaction of Log R profiles (expected decrease in case of deletion in diploid sample, 100% aberrant cells; 1 in ideal case, 0.55 of Illumina 109K arrays)
pdfPlot	Optional flag if nonrounded plots and ASCAT profile in pdf format are desired. Default=F
y_limit	Optional parameter determining the size of the y axis in the nonrounded plot and ASCAT profile. Default=5
circos	Optional file to output the non-rounded values in Circos track format. Default=NA
rho_manual	optional argument to override ASCAT optimization and supply rho parameter (not recommended)
psi_manual	optional argument to override ASCAT optimization and supply psi parameter (not recommended)

Details

Note: for copy number only probes, nA contains the copy number value and nB = 0.

Value

an ASCAT output object, containing:
1. nA: copy number of the A allele
2. nB: copy number of the B allele
3. aberrantcellfraction: the aberrant cell fraction of all arrays
4. ploidy: the ploidy of all arrays
5. failedarrays: arrays on which ASCAT analysis failed
6. nonaberrantarrays: arrays on which ASCAT analysis indicates that they show virtually no aberrations
7. segments: an array containing the copy number segments of each sample (not including failed arrays)
8. segments_raw: an array containing the copy number segments of each sample without any rounding applied

base.gw.plot *base.gw.plot*

Description

base.gw.plot

Usage

```
base.gw.plot(bafsegmented, nAfullPlot, nBfullPlot, colourTotal, colourMinor,
  maintitle, chr.segs, lrr, y_limit, twoColours = FALSE)
```

Arguments

bafsegmented	B Allele Frequency, segmented, in genomic sequence (only probes heterozygous in germline), with probe IDs
nAfullPlot	Total segment copy number
nBfullPlot	Segment copy number minor allele
colourTotal	Colour to plot total copy number
colourMinor	Colour to plot minor allele
maintitle	Title comprising ploidy, rho, goodness of fit
chr.segs	Vector comprising chromosome segments
lrr	(unsegmented) log R, in genomic sequence (all probes), with probe IDs
y_limit	Optional parameter determining the size of the y axis in the nonrounded plot and ASCAT profile. Default=5
twoColours	Optional flag to specify colours, if TRUE colour is paler for CN values > y_limit

Value

basic plot containing chromosome positions and names, plots copy number for either ASCAT non rounded or BB average

runASCAT *runASCAT*

Description

the ASCAT main function

Usage

```
runASCAT(lrr, baf, lrrsegmented, bafsegmented, gender, SNPpos, chromosomes,
  chrnames, sexchromosomes, failedqualitycheck = F, distancepng = NA,
  copynumberprofilespng = NA, nonroundedprofilepng = NA,
  aberrationreliabilitypng = NA, gamma = 0.55, rho_manual = NA,
  psi_manual = NA, pdfPlot = F, y_limit = 5, circos = NA)
```

Arguments

lrr	(unsegmented) log R, in genomic sequence (all probes), with probe IDs
baf	(unsegmented) B Allele Frequency, in genomic sequence (all probes), with probe IDs
lrrsegmented	log R, segmented, in genomic sequence (all probes), with probe IDs
bafsegmented	B Allele Frequency, segmented, in genomic sequence (only probes heterozygous in germline), with probe IDs
gender	a vector of gender for each cases ("XX" or "XY"). Default = NULL: all female ("XX")
SNPpos	position of all SNPs
chromosomes	a list containing c vectors, where c is the number of chromosomes and every vector contains all probe numbers per chromosome
chrnames	a vector containing the names for the chromosomes (e.g. c(1:22,"X"))
sexchromosomes	a vector containing the names for the sex chromosomes
failedqualitycheck	did the sample fail any previous quality check or not?
distancepng	if NA: distance is plotted, if filename is given, the plot is written to a .png file
copynumberprofilespng	if NA: possible copy number profiles are plotted, if filename is given, the plot is written to a .png file
nonroundedprofilepng	if NA: copy number profile before rounding is plotted (total copy number as well as the copy number of the minor allele), if filename is given, the plot is written to a .png file
aberrationreliabilitypng	if NA: aberration reliability score is plotted, if filename is given, the plot is written to a .png file
gamma	technology parameter, compaction of Log R profiles (expected decrease in case of deletion in diploid sample, 100% aberrant cells; 1 in ideal case, 0.55 of Illumina 109K arrays)
rho_manual	optional argument to override ASCAT optimization and supply rho parameter (not recommended)
psi_manual	optional argument to override ASCAT optimization and supply psi parameter (not recommended)
pdfPlot	Optional flag if nonrounded plots and ASCAT profile in pdf format are desired. Default=F
y_limit	Optional parameter determining the size of the y axis in the nonrounded plot and ASCAT profile. Default=5
circos	Optional file to output the non-rounded values in Circos track format. Default=NA

Value

list containing optimal purity and ploidy

Index