

# Package ‘Ppurple’

January 23, 2018

**Title** P(robabilistic)PU(rity)PL(oidy)E(stimation)  
Operations

**Version** 0.2.0

**Description** Purity and ploidy inference from WGS, panel, WES from total read density and hets.

**Depends** R (>= 3.1.0),  
GenomicRanges (>= 1.18),  
gUtils,  
bamUtils,  
data.table (>= 1.9)

**Imports** IRanges (>= 2.0),  
S4Vectors (>= 0.4),  
GenomeInfoDb (>= 1.2),  
parallel,  
BiocGenerics(>= 0.12),  
methods,  
Matrix,  
stringr

**Suggests** BSgenome.Hsapiens.UCSC.hg19,  
testthat,  
rtracklayer

**License** GPL-2

**BugReports** <http://github.com/mskilab/Ppurple>

**LazyData** true

**RoxygenNote** 6.0.1.9000

## R topics documented:

ppurple . . . . .	2
Index	3

ppurple

*Probabilistic purity ploidy estimation***Description**

Computes posterior probability of purity ploidy for data with coverage cov (granges with value \$y specifying coverage) hets with \$alt, \$tref, \$nref, \$nalt specifying tumor and normal alt and ref allelic counts, and (optional) segments via EM algorithm.

If segments not specified, then it is inferred from the segmentation of cov via CBS (using DNA copy)

EM begins with initial (user specified) grid of values and then refines at 10 fold (or value of refine arg, >1) around the modes, meaning that it reruns the calculation with a finer grid completely surrounding the modes to output the final solution which is a data.table mapping \$purity and \$ploidy values to a posterior probability \$p

**Usage**

```
ppurple(cov, hets = NULL, segs = NULL, purities = seq(0, 1, 0.1),
        ploidy = seq(1, 5, 0.2), refine = 10, K = 20, verbose = TRUE,
        min.p = 1e-04, mc.cores = 1, numchunks = mc.cores)
```

**Arguments**

cov	GRanges or data.table of genome wide coverage tiles with field \$y specifying normalized coverage
hets	GRanges or data.table of hets with fields \$alt, \$tref specifying alt and ref counts of hets in tumor
segs	GRanges of pre-computed segments (optional, if NULL these will be computed via DNA copy of cov)
purities	numeric vector of ploidies to sweep in grid (default from 1 to 5, 0.2 increment)
ploidy	numeric vector of purities to sweep in grid (default from 0 to 1, 0.1 increment)
refine	integer scalar of how many fold refinement of purities x ploidies grid to perform after initial run (default = 10, careful to not make too big)
K	integer scalar number of copy states to model (default 20)
verbose	logical flag
mc.cores	integer scalar to parallelize (default = 1)
numchunks	how many chunks to parallelize over (default = mc.cores)

**Author(s)**

Marcin Imielinski

# Index

ppurple, [2](#)