

Package ‘MultiSV’

August 27, 2014

Encoding UTF-8

Type Package

Title MultiSV: an R package for identification of structural variations in multiple populations based on whole genome resequencing

Version 0.0-67

Date 2014-08-27

Author Khurram Maqbool

Maintainer Khurram Maqbool <khurram.maqbool@outlook.com>

Description MultiSV is an R package for identification of structural variations in multiple populations based on whole genome resequencing. It fits linear mixed model and identifies structural variations in multiple populations using whole genome sequencing data. It could also be manipulated to use on RNA-seq data for differential gene expression (implementation in future releases). Main steps for analysis include generating read depth in bins using ComputeBinCounts, conversion of bins to MultiSV format using Bin2MultiSV. Finally, identification of structural variations using CallMultiSV.

Depends R (>= 2.10)

Imports nlme,reshape

License LGPL (>= 3)

LazyLoad yes

LazyData yes

NeedsCompilation no

Repository CRAN

Date/Publication 2014-08-27 18:52:26

R topics documented:

Bin2MultiSV	2
CallMultiSV	3
CmptMltPvl	4
ComputeBinCounts	5
GetTestCfgFile	6
gtExEs	6
IdfMltSV	7
MultiSVData	8
MultiSVExample	8
PrbMlt	9
PrepMultiDt	10
PrepMultiDtLgMn	10
ProcMutlDt	11
WriteMultiSV	12
Index	13

Bin2MultiSV	<i>Read data using configuration file</i>
-------------	---

Description

Read data using configuration file

Usage

```
Bin2MultiSV(CfgFile)
```

Arguments

CfgFile : A tab separated file with the following format (with no header)
 Contrast SampleID PathtoRDbinFile
 Treat Sample1 ~/Sample1.WindowRD-1000.count

Value

Bin2MultiSV: read configuration file

Author(s)

Khurram Maqbool

See Also

Bin2MultiSV

Examples

```
## Not run:
ConfigFile <- "~/MultiSV.conf
Bin2MultiSV(ConfigFile)

## End(Not run)
```

CallMultiSV	<i>A Method to identify MultiSV from RDbIn files for different samples defined in Configuration File</i>
-------------	--

Description

This is the main function to identify SVs. It works after the data has been prepared using Bin2MultiSV from bam files. Bin2MultiSV will prepare the data for MultiSV package and CallMultiSV will identify SVs.

Usage

```
CallMultiSV(CfgFile, LgLim, SVSize, MtlSVOut)
```

Arguments

CfgFile	: A tab separated file with the following format (with no header) Contrast SampleID PathtoRDbInFile Treat Sample1 ~/Sample1.WindowRD-1000.count Control Ref1 ~/Ref1.WindowRD-1000.count Contrast SampleID PathtoRDbInFile
LgLim	: log2 limit, set 0.6 or higher
SVSize	: Minimum size of SV to identify, set 5000 or higher
MtlSVOut	: Output file

Details

Data Input from read depth (RD) bins

Value

CallMultiSV:

Author(s)

Khurram Maqbool

See Also

Bin2MultiSV

Examples

```
## Not run:  
CallMultiSV(CfgFile, LgLim, SVSize, MtlSVOut)  
  
## End(Not run)
```

CmptMltPvl

Compute Pvals of MultiSVs

Description

Compute Pvals of MultiSVs

Usage

```
CmptMltPvl(MultiData,CfgFile)
```

Arguments

CfgFile : A tab separated file with the following format (with no header)
Contrast SampleID PathtoRDbInFile
Treat Sample1 ~/Sample1.WindowRD-1000.count

MultiData : MultiSV Data object

Value

CmptMltPvl: 0.05

Author(s)

Khurram Maqbool

See Also

CmptMltPvl

Examples

```
## Not run:  
CmptMltPvl("MultiSV.conf")  
  
## End(Not run)
```

ComputeBinCounts *Compute read counts in bins from read hits.*

Description

Compute read counts in bins from read hits.

Usage

```
ComputeBinCounts(RBinSampleFile,  
RBinChrSizeFile,RBinSampleName,  
RBinWindowSize,OutFolder)
```

Arguments

```
RBinSampleFile  
                  : File containing read hits  
RBinChrSizeFile  
                  : File containing list of chromosomes with their sizes  
RBinSampleName  
                  : Name of the Individual or population  
RBinWindowSize  
                  : Size of bin  
OutFolder  
                  : output directory to save the counts file
```

Value

ComputeBinCounts: Compute read counts in bins

Author(s)

Khurram Maqbool

See Also

Bin2MultiSV

Examples

```
## Not run:  
ComputeBinCounts(RBinSampleFile, RBinChrSizeFile, RBinSampleName ,  
RBinWindowSize, OutFolder)  
  
## End(Not run)
```

GetTestCfgFile

Function with MultiSV input example

Description

Function with MultiSV input example

Usage

GetTestCfgFile()

Author(s)

Khurram Maqbool

Examples

```
##' @export
`MultiSVExample.default` <- function(MultiSVData){
  LgLim = 0.6
  SVSize = 5000
  MultiData <- PrepMultiDt(MultiSVData)
  MultiData <- PrepMultiDtLgMn(MultiData)
  MultiData <- ProcMutlDt(MultiData)
  IdfMltSV(MultiData,LgLim,SVSize)
}
```

gtExEs*proc information*

Description

proc information

Usage

gtExEs(proc)

Arguments

proc : Name of proc

Value

gtExEs: proc

Author(s)

Khurram Maqbool

See Also

Bin2MultiSV

Examples

```
## Not run:  
gtExEs(proc)  
  
## End(Not run)
```

IdfMltSV

Call MultiSVs

Description

Call MultiSVs

Usage

```
IdfMltSV(MultiData,LgLim,SVSize)
```

Arguments

MultiData : MultiSV Data object
LgLim : log2 threshold
SVSize : Minimum size of SV to identify, set 5000 or higher

Value

IdfMltSV: read configuration file

Author(s)

Khurram Maqbool

Examples

```
## Not run:  
IdfMltSV(MultiData,Cfg,LgLim,SVSize)  
  
## End(Not run)
```

MultiSVData

RD Data

Description

The data set is read depth (RD) from six domestic populations including ugc355, ugc356, ugc357, ugc358, ugc376, ugc377 and one wild (WB) population of pigs (SS10.2). The reads for the populations are available from SRA accession number SRA057461. RD from each population is in 1kb bins obtained after aligning the reads from each population to SS10.2 reference genome assembly.

Author(s)

Khurram Maqbool

References

Rubin, C.-J., H.-J.Megens, A.Martinez Barrio, K.Maqbool, S. Sayyab, D. Schwochow, C.Wang, Ö. Carlborg, P. Jern, C.B. Jørgensen, A. L.Archibald, M. Fredholm, M. A. M. Groenen, and L.Andersson. 2012. Strong signatures of selection in the domestic pig genome. Proc. Natl. Acad. Sci. USA 109:19529–19536, DOI: 10.1073/pnas.1217149109

MultiSVExample

MultiSV Example

Description

MultiSV Example

Usage

MultiSVExample(MultiSVData)

Arguments

MultiSVData : MultiSV Data object

Value

MultiSVExample: Run a small example to identify MultiSV

Author(s)

Khurram Maqbool

See Also

CallMultiSV

Examples

```
## Not run:  
MultiSVExample(MultiSVData)  
  
## End(Not run)
```

PrbMlt	<i>Fit linear mixed model</i>
--------	-------------------------------

Description

Fit linear mixed model

Usage

```
PrbMlt(MultiData)
```

Arguments

MultiData : MultiSV Data object

Value

PrbMlt: read configuration file

Author(s)

Khurram Maqbool

See Also

PrbMlt

Examples

```
## Not run:  
PrbMlt(MultiData)  
  
## End(Not run)
```

PrepMultiDt *Prepare MultiSV data*

Description

Prepare MultiSV data

Usage

PrepMultiDt(MultiData)

Arguments

MultiData : MultiSV Data object

Value

PrepMultiDt: Prepare MultiSV data

Author(s)

Khurram Maqbool

See Also

PrepMultiDt

Examples

```
## Not run:  
PrepMultiDt(MultiData)  
  
## End(Not run)
```

PrepMultiDtLgMn *Prepare MultiSV data*

Description

Prepare MultiSV data

Usage

PrepMultiDtLgMn(MultiData)

Arguments

MultiData : MultiSV Data object

Value

PrepMultiDtLgMn: Prepare MultiSV data

Author(s)

Khurram Maqbool

See Also

PrepMultiDt

Examples

```
## Not run:  
PrepMultiDtLgMn(MultiData)  
  
## End(Not run)
```

ProcMutilDt

Process Multi Data

Description

Process Multi Data

Usage

ProcMutilDt(MultiData)

Arguments

MultiData : MultiSV Data object

Value

ProcMutilDt: Process MultiSV data

Author(s)

Khurram Maqbool

See Also

ProcMutilDt

Examples

```
## Not run:  
ProcMutilDt(MultiData)  
  
## End(Not run)
```

WriteMultiSV

Output MultiSVs

Description

Output MultiSVs

Usage

```
WriteMultiSV(MultiData,File)
```

Arguments

```
MultiData      : MultiSV Data object  
File           : path to output File e.g. MultiSV.gff
```

Value

WriteMultiSV: read configuration File

Author(s)

Khurram Maqbool

Examples

```
## Not run:  
WriteMultiSV(MultiData,File)  
  
## End(Not run)
```

Index

- *Topic **CNV**
 - MultiSVData, 8
- *Topic **Deletion**
 - MultiSVData, 8
- *Topic **Duplication**
 - MultiSVData, 8
- *Topic **MultiSV**
 - MultiSVData, 8
- *Topic **Population**
 - MultiSVData, 8

- Bin2MultiSV, 2

- CallMultiSV, 3
- CmpMltPvl, 4
- ComputeBinCounts, 5

- GetTestCfgFile, 6
- gtExEs, 6

- IdfMltSV, 7

- MultiSVData, 8
- MultiSVExample, 8

- PrbMlt, 9
- PrepMultiDt, 10
- PrepMultiDtLgMn, 10
- ProcMutildt, 11

- WriteMultiSV, 12