

# Package ‘MethylMix’

April 10, 2015

**Version** 1.0.0

**Date** 2014-08-25

**Type** Package

**Title** MethylMix: Identifying methylation driven cancer genes.

**Author** Olivier Gevaert

**Maintainer** Olivier Gevaert <olivier.gevaert@gmail.com>

**Description** MethylMix is an algorithm implemented to identify hyper and hypomethylated genes for a disease. MethylMix is based on a beta mixture model to identify methylation states and compares them with the normal DNA methylation state. MethylMix uses a novel statistic, the Differential Methylation value or DM-value defined as the difference of a methylation state with the normal methylation state. Finally, matched gene expression data is used to identify, besides differential, functional methylation states by focusing on methylation changes that effect gene expression.

**License** GPL-2

**LazyLoad** yes

**Depends** R (>= 3.1.1)

**Imports** foreach,parallel,doParallel,RColorBrewer,optimx,RPMM

**biocViews**

DNAMethylation,StatisticalMethod,DifferentialMethylation,GeneRegulation,GeneExpression,MethylationArray,  
DifferentialExpression, Pathways, Network

**Suggests** BiocStyle

## R topics documented:

MAcancer . . . . .	2
MAcancer_CDH1 . . . . .	2
METcancer . . . . .	3
METcancer_CDH1 . . . . .	4
MethylMix . . . . .	4

MethylMix_PlotModel . . . . .	5
METnormal . . . . .	7
METnormal_CDHI . . . . .	7

<b>Index</b>	<b>9</b>
--------------	----------

---

MAcancer	<i>Cancer Gene expression data of glioblastoma patients from the TCGA project.</i>
----------	--

---

### Description

Cancer Gene expression data of glioblastoma patients from the TCGA project. A set of 14 genes that have been shown in the literature to be involved in differential methylation in glioblastoma were selected as an example to try out MethylMix.

### Usage

MAcancer

### Format

This describes the format of MAcancer.

### Source

This gives the source of MAcancer

### References

Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature*. 2008 Oct 23;455(7216):1061-8. doi: 10.1038/nature07385. Epub 2008 Sep 4. Erratum in: *Nature*. 2013 Feb 28;494(7438):506. PubMed PMID: 18772890; PubMed Central PMCID: PMC2671642.

Also see TCGA: The Cancer Genome Atlas: <http://cancergenome.nih.gov/>

---

MAcancer_CDHI	<i>Cancer Gene expression data of CDHI for breast cancer patients from the TCGA project.</i>
---------------	--

---

### Description

Cancer Gene expression data for the gene CDHI of breast cancer patients from the TCGA project.

### Usage

MAcancer\_CDHI

**Format**

This describes the format of MAcancer.

**Source**

This gives the source of MAcancer

**References**

Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012 Oct 4;490(7418):61-70. doi: 10.1038/nature11412. Epub 2012 Sep 23. PubMed PMID: 23000897; PubMed Central PMCID: PMC3465532. Also see TCGA: The Cancer Genome Atlas: <http://cancergenome.nih.gov/>

---

METcancer

*DNA methylation data from cancer tissue from glioblastoma patients from the TCGA project.*

---

**Description**

Cancer Gene expression data of glioblastoma patients from the TCGA project. A set of 14 genes that have been shown in the literature to be involved in differential methylation in glioblastoma were selected as an example to try out MethylMix.

**Usage**

METcancer

**Format**

This describes the format of METcancer.

**Source**

This gives the source of METcancer

**References**

Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature*. 2008 Oct 23;455(7216):1061-8. doi: 10.1038/nature07385. Epub 2008 Sep 4. Erratum in: *Nature*. 2013 Feb 28;494(7438):506. PubMed PMID: 18772890; PubMed Central PMCID: PMC2671642. Also see TCGA: The Cancer Genome Atlas: <http://cancergenome.nih.gov/>

---

METcancer_CDH1	<i>DNA methylation data for CDH1 from the breast cancer TCGA project.</i>
----------------	---

---

**Description**

Raw DNA methylation data for CDH1 from the 450k Infinium platform from breast cancer patients from the TCGA project.

**Usage**

METcancer\_CDH1

**Format**

This describes the format of METcancer.

**Source**

This gives the source of METcancer

**References**

Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012 Oct 4;490(7418):61-70. doi: 10.1038/nature11412. Epub 2012 Sep 23. PubMed PMID: 23000897; PubMed Central PMCID: PMC3465532. Also see TCGA: The Cancer Genome Atlas: <http://cancergenome.nih.gov/>

---

MethylMix	<i>Mixture model for DNA methylation data in cancer.</i>
-----------	--

---

**Description**

MethylMix identifies DNA methylation driven genes by modeling DNA methylation data in cancer vs. normal and looking for homogeneous subpopulations. In addition matched gene expression data (e.g. from microarray technology or RNA sequencing) is used to identify functional DNA methylation events by requiring a negative correlation between methylation and gene expression of a particular gene.

**Usage**

MethylMix(METcancer, METnormal, MACancer, OutputRoot, Parallel)

**Arguments**

METcancer	This a matrix with the methylation data of cancer tissue with genes in rows and samples in columns
METnormal	This is a matrix with the normal methylation data of the same genes as in MET-cancer. Again genes in rows and samples in columns. The samples do not have to match with the cancer data.
MAcancer	This is the matched gene expression data for the same samples as in METcancer.
OutputRoot	Path to store the MethylMix results object.
Parallel	If true MethylMix will be run using parallel processing.

**Value**

MethylMixResults is a list with the following components:

MethylationStates	Matrix with for all genes the Methylation states using DM-value (i.e. Differential methylation values) that are defined as the methylation value with respect to the average normal methylation for a gene.
NrComponents	The number of methylation states for each gene.
Models	Beta mixture model parameters for each gene.
MethylationDrivers	Genes identified as functional and differential by MethylMix.
MixtureStates	A list with the DM-values for each gene that is functional and differential.

**Examples**

```
# load the three data sets needed for MethylMix
data(METcancer)
data(METnormal)
data(MAcancer)

# run methylmix on a small set of example data
MethylMixResults=MethylMix(METcancer,METnormal,MAcancer)

# try the parallel toolbox to speed up MethylMix modeling
MethylMixResults=MethylMix(METcancer,METnormal,MAcancer,Parallel=TRUE)
```

---

MethylMix\_PlotModel *Plotting a mixture model for a gene.*

---

**Description**

MethylMix identifies DNA methylation driven genes by modeling DNA methylation data in cancer vs. normal and looking for homogeneous subpopulations. In addition matched gene expression data (e.g. from microarray technology or RNA sequencing) is used to identify functional DNA methylation events by requiring a negative correlation between methylation and gene expression of a particular gene.

**Usage**

```
MethylMix_PlotModel(GeneName, METdata, MixtureModelResults, MAdata=0,
METnormal=0, FileName="")
```

**Arguments**

GeneName	Name of the gene for which to create a MethylMix plot.
METdata	This a matrix with the methylation data of cancer tissue with genes in rows and samples in columns
MixtureModelResults	The results object from a MethylMix run.
METnormal	This is a matrix with the normal methylation data of the same genes as in METcancer. Again genes in rows and samples in columns. The samples do not have to match with the cancer data.
MAdata	This is the matched gene expression data for the same samples as in METcancer.
FileName	Filename to export the figure. If empty figure is shown in console.

**Examples**

```
# load the three data sets needed for MethylMix
data(METcancer)
data(METnormal)
data(MAcancer)

# run methylmix on a small set of example data
MethylMixResults=MethylMix(METcancer, METnormal, MAcancer)

# Plot the most famous methylated gene for glioblastoma
MethylMix_PlotModel(MGMT, METcancer, MethylMixResults)

# plot MGMT also with its normal methylation variation
MethylMix_PlotModel(MGMT, METcancer, MethylMixResults, MAdata=0, METnormal)

# plot a MethylMix model for another gene
MethylMix_PlotModel(ZNF217, METcancer, MethylMixResults, MAdata=0, METnormal)

# also plot the inverse correlation with gene expression
# this creates two separate plots
MethylMix_PlotModel(MGMT, METcancer, MethylMixResults, MAdata=MAcancer)
```

```
# plot all functional and differential genes
for (i in 1:length(MethylMixResults$MethylationDrivers)) {
  MethylMix_PlotModel(MethylMixResults$MethylationDrivers[i],METcancer,
    MethylMixResults,MAdata=0,METnormal)
}
```

---

METnormal	<i>DNA methylation data from normal tissue from glioblastoma patients.</i>
-----------	--

---

### Description

Normal tissue DNA methylation data of glioblastoma patients. These normal tissue samples were run on the same platform and are described in the publication referenced below.

### Usage

METnormal

### Format

This describes the format of METnormal.

### Source

This gives the source of METnormal

### References

Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K, Berman BP, Pan F, Pelloski CE, Sulman EP, Bhat KP, Verhaak RG, Hoadley KA, Hayes DN, Perou CM, Schmidt HK, Ding L, Wilson RK, Van Den Berg D, Shen H, Bengtsson H, Neuvial P, Cope LM, Buckley J, Herman JG, Baylin SB, Laird PW, Aldape K; Cancer Genome Atlas Research Network. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell*. 2010 May 18;17(5):510-22. doi: 10.1016/j.ccr.2010.03.017. Epub 2010 Apr 15. PubMed PMID: 20399149; PubMed Central PMCID: PMC2872684

---

METnormal_CDH1	<i>DNA methylation data for CDH1 from the breast cancer TCGA project.</i>
----------------	---

---

### Description

Raw DNA methylation data for CDH1 from the 450k Infinium platform from breast cancer patients from the TCGA project.

**Usage**

METnormal\_CDHI

**Format**

This describes the format of METcancer.

**Source**

This gives the source of METcancer

**References**

Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012 Oct 4;490(7418):61-70. doi: 10.1038/nature11412. Epub 2012 Sep 23. PubMed PMID: 23000897; PubMed Central PMCID: PMC3465532.

Also see TCGA: The Cancer Genome Atlas: <http://cancergenome.nih.gov/>



# Index

MAcancer, [2](#)  
MAcancer\_CDH1, [2](#)  
METcancer, [3](#)  
METcancer\_CDH1, [4](#)  
MethylMix, [4](#)  
MethylMix\_PlotModel, [5](#)  
METnormal, [7](#)  
METnormal\_CDH1, [7](#)