

# Package ‘MPFE’

April 25, 2025

**Type** Package

**Title** Estimation of the amplicon methylation pattern distribution from bisulphite sequencing data

**Version** 1.45.0

**Date** 2014-11-28

**Author** Peijie Lin, Sylvain Foret, Conrad Burden

**biocViews** HighThroughputSequencingData, DNAMethylation, MethylSeq

**Maintainer** Conrad Burden <conrad.burden@anu.edu.au>

**Description** Estimate distribution of methylation patterns from a table of counts from a bisulphite sequencing experiment given a non-conversion rate and read error rate.

**License** GPL (>= 3)

**git\_url** <https://git.bioconductor.org/packages/MPFE>

**git\_branch** devel

**git\_last\_commit** 146b3d9

**git\_last\_commit\_date** 2025-04-15

**Repository** Bioconductor 3.22

**Date/Publication** 2025-04-24

## Contents

|                            |          |
|----------------------------|----------|
| MPFE-package . . . . .     | 2        |
| estimatePatterns . . . . . | 2        |
| patternMap . . . . .       | 4        |
| patternsExample . . . . .  | 5        |
| plotPatterns . . . . .     | 5        |
| <b>Index</b>               | <b>7</b> |

---

|              |             |
|--------------|-------------|
| MPFE-package | <i>MPFE</i> |
|--------------|-------------|

---

### Description

Estimate distribution of methylation patterns from a table of counts from a bisulphite sequencing experiment given a non-conversion rate and sequencing error rate.

### Details

Package: MPFE  
 Type: Package  
 License: GPL(>=3)

The main component of this package is the function [estimatePatterns](#), which reads a table of read counts of bisulphite sequencing data for a given amplicon and generates a table and plot of the estimated distribution over methylation patterns.

### Author(s)

Peijie Lin, Sylvain Foret, Conrad Burden  
 Maintainer: conrad.burden@anu.edu.au

### Examples

```
data(patternsExample)
estimates <- estimatePatterns(patternsExample, epsilon=0.02, eta=0.01)
estimates
plotPatterns(estimates[[2]])
```

---

|                  |                                                      |
|------------------|------------------------------------------------------|
| estimatePatterns | <i>Estimate distribution of methylation patterns</i> |
|------------------|------------------------------------------------------|

---

### Description

Estimate distribution of methylation patterns from a table of counts from a bisulphite sequencing experiment given a non-conversion rate and a sequencing error rate.

### Usage

```
estimatePatterns(patternCounts,
                 epsilon=0,
                 eta=0,
                 column=NULL,
                 fast=TRUE,
                 steps=20000,
                 reltol=1e-12)
```

**Arguments**

|               |                                                                                                                                                                                                                       |
|---------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| patternCounts | data frame with methylation patterns in first column and pattern counts in subsequent columns.                                                                                                                        |
| epsilon       | non-conversion rate, a value between 0 and 1.                                                                                                                                                                         |
| eta           | error rate, either a vector of numbers between 0 and 1 of length equal to the number of CpG sites or a single value between 0 and 1 for a single error rate across all sites.                                         |
| column        | a vector that specifies the indices of the columns of 'patternCounts' to process. Its entries are integer values from 1 to the number of pattern counts columns in 'patternCounts'. If NULL, defaults to all columns. |
| fast          | logical, if TRUE, fast version implemented (default).                                                                                                                                                                 |
| steps         | number of steps for the optimiser, passed to <a href="#">constrOptim</a> . If NULL, defaults to 20000 steps.                                                                                                          |
| reltol        | relative tolerance for the optimiser, passed to <a href="#">constrOptim</a> . If NULL, defaults to 1e-12.                                                                                                             |

**Value**

The function returns a list of data frames.

The data frames contain the following columns:

|                       |                                                               |
|-----------------------|---------------------------------------------------------------|
| pattern               | the list of input patterns (factor)                           |
| coverage              | the number of reads for each pattern (integer)                |
| observedDistribution  | the observed frequencies of each pattern (numeric)            |
| estimatedDistribution | the estimated frequencies (numeric)                           |
| spurious              | indicates whether the patterns are real or spurious (logical) |

**Author(s)**

Peijie Lin, Sylvain Foret, Conrad Burden

**Examples**

```
data(patternsExample)
estimatePatterns(patternsExample,
  epsilon=0.02,
  eta=0.01)
estimatePatterns(patternsExample,
  epsilon=0.01,
  eta=c(0.015, 0.01, 0.01, 0.01, 0.015),
  column=2)
```

---

patternMap

---

*Plot a representation of the patterns and their frequencies*


---

## Description

Plot the observed distribution and the estimated distribution of the methylation patterns

## Usage

```
patternMap(patterns,
            minFreq=0,
            maxFreq=1,
            noSpurious=TRUE,
            estimatedDistribution=TRUE,
            topDown=TRUE,
            allTicks=FALSE,
            methCol='black',
            unMethCol='white',
            ...)
```

## Arguments

|                       |                                                                                                                                                                       |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| patterns              | A data frame obtained from the output of the function <a href="#">estimatePatterns</a> .                                                                              |
| minFreq               | Only plot patterns with at least minFreq frequency.                                                                                                                   |
| maxFreq               | Only plot patterns with more maxFreq frequency or more.                                                                                                               |
| noSpurious            | Don't plot spurious patterns (only relevant if estimatedDistribution is FALSE).                                                                                       |
| estimatedDistribution | Use the frequencies from the estimated distribution. If FALSE, use the observed distribution.                                                                         |
| topDown               | Put the most abundant patterns at the top. If FALSE the most abundant patterns are at the bottom.                                                                     |
| allTicks              | Draw a tick under every position.                                                                                                                                     |
| methCol               | The colour for the methylated positions. Can be a single colour, a vector of colours (recycled), or a function (for instance from <a href="#">colorRampPalette</a> ). |
| unMethCol             | As methCol but for un-methylated positions.                                                                                                                           |
| ...                   | Other arguments passed to <a href="#">plot</a>                                                                                                                        |

## Details

This function draws a map of the different pattern and their frequencies based on the values returned by [estimatePatterns](#).

## Author(s)

Peijie Lin, Sylvain Foret, Conrad Burden

**Examples**

```
data(patternsExample)
estimates <- estimatePatterns(patternsExample,
                              epsilon=0.02,
                              eta=0.01)

patternMap(estimates[[1]])
```

---

|                 |                        |
|-----------------|------------------------|
| patternsExample | <i>patternsExample</i> |
|-----------------|------------------------|

---

**Description**

A data frame which contains a column of methylation patterns and two columns of counts. This data was obtained as described in Lyko, F., Forest, S., Kucharski, R., Wolf, S., Falckenhayn, C., and Maleszka, R. (2010). The honey bee epigenomes: differential methylation of brain DNA in queens and workers. PLoS Biol, 8(11), e1000506.

**Usage**

```
data(patternsExample)
```

**Format**

This data frame contains the following columns:

**mPattern** methylation patterns  
**k1** first column of counts  
**k2** second column of counts

---

|              |                                              |
|--------------|----------------------------------------------|
| plotPatterns | <i>Plot methylation pattern distribution</i> |
|--------------|----------------------------------------------|

---

**Description**

Plot the observed distribution and the estimated distribution of the methylation patterns

**Usage**

```
plotPatterns(compareData, yLimit1=NULL, yLimit2=NULL)
```

**Arguments**

|             |                                                                                                       |
|-------------|-------------------------------------------------------------------------------------------------------|
| compareData | data frame, obtained from the output of the function <a href="#">estimatePatterns</a>                 |
| yLimit1     | upper limit of y-axis on left hand scale of the first graph. If NULL, defaults to show all patterns   |
| yLimit2     | upper limit of y-axis on left hand scale of the second graph. If NULL, defaults to show most patterns |

## Details

The two graphs in the output plot are the same but have different ranges. The parameters `yLimit1` and `yLimit2` control the range of the `y_axis` on the plots produced.

**Value**

A plot that compares the observed read distribution with the estimated distribution.

**Author(s)**

Peijie Lin, Sylvain Foret, Conrad Burden

## Examples

```
data(patternsExample)
estimates <- estimatePatterns(patternsExample,
                              epsilon=0.02,
                              eta=0.01)
plotPatterns(estimates[[1]])
```

# Index

- \* **amplicon**,
  - MPFE-package, [2](#)
- \* **bisufite sequencing**,
  - MPFE-package, [2](#)
- \* **bisulphite sequencing**,
  - MPFE-package, [2](#)
- \* **methylation**
  - MPFE-package, [2](#)
- \*
  - MPFE-package, [2](#)
- colorRampPalette, [4](#)
- constrOptim, [3](#)
- estimatePatterns, [2](#), [2](#), [4](#), [5](#)
- MPFE (MPFE-package), [2](#)
- MPFE-package, [2](#)
- patternMap, [4](#)
- patternsExample, [5](#)
- plot, [4](#)
- plotPatterns, [5](#)