# Lab 3A: An Introduction to Bioconductor's `marray` packages

June 4, 2003

In this lab, we demonstrate the main functions in the `marray` suite of packages for diagnostic plots and normalization of two-color spotted microarray data. To load the packages

```
> library(marrayNorm)
```

```
Loading required package: marrayClasses
Creating a new generic function for "print" in package
marrayClasses
Creating a new generic function for "rbind" in package
marrayClasses
Creating a new generic function for "cbind" in package
marrayClasses
Loading required package: stepfun
Loading required package: marrayInput
Loading required package: marrayPlots
```

For a more detailed introduction, consult the package vignettes which can be listed by the command `openVignette()`. A demo for `marrayPlots` can also be accessed by `demo(marrayPlots)`. We will work with the sample dataset `swirl`; for a description of `swirl`, type `? swirl`. To load this dataset

```
> data(swirl)
```

## 1 Basic classes and methods: `marrayClasses` package

One of the main classes in `marrayClasses` is the `marrayLayout` class; it is used to keep track of important layout parameters, such as the total number of spotted probe sequences on the array, the dimensions of the spot and grid matrices, the plate origin of the probes, information on spotted control sequences. For details on this class consult the help file, ? marrayLayout. Two other important classes are marrayRaw and marrayNorm, which represent, respectively, pre-normalization and post-normalization intensity data for a batch of spotted microarrays. Methods for manipulating instances of these classes are also described in the help files.

The object swirl is an instance of the class marrayRaw. Try the following commands to obtain information on this object

```
> class(swirl)

[1] "marrayRaw"

> slotNames(swirl)

[1] "maRf"      "maGf"      "maRb"      "maGb"      "maW"       "maLayout"
[7] "maGnames"  "maTargets" "maNotes"

> swirl

Pre-normalization intensity data:        Object of class marrayRaw.

Number of arrays:        4 arrays.

A) Layout of spots on the array:
Array layout:        Object of class marrayLayout.

Total number of spots:                       8448
Dimensions of grid matrix:                4 rows by 4 cols
Dimensions of spot matrices:                22 rows by 24 cols

Currently working with a subset of 8448 spots.

Control spots:
There are   2 types of controls :
Control       N
    768    7680



Notes on layout:
No Input File

B) Samples hybridized to the array:
Object of class marrayInfo.
```

```
  maLabels # of slide        Names experiment Cy3 experiment Cy5      date
1       81          81 swirl.1.spot          swirl      wild type 2001/9/20
2       82          82 swirl.2.spot      wild type          swirl 2001/9/20
3       93          93 swirl.3.spot          swirl      wild type 2001/11/8
4       94          94 swirl.4.spot      wild type          swirl 2001/11/8
  comments
1       NA
2       NA
3       NA
4       NA

Number of labels:  4
Dimensions of maInfo matrix:  4  rows by  6  columns

Notes:
C:/GNU/R/rw1041/library/marrayInput/data/SwirlSample.txt

C) Summary statistics for log-ratio distribution:
              Min. 1st Qu. Median  Mean 3rd Qu. Max.
swirl.1.spot -2.73   -0.79  -0.58 -0.48   -0.29 4.42
swirl.2.spot -2.72   -0.15   0.03  0.03    0.21 2.35
swirl.3.spot -2.29   -0.75  -0.46 -0.42   -0.12 2.65
swirl.4.spot -3.21   -0.46  -0.26 -0.27   -0.06 2.90

D) Notes on intensity data:

   To access individual slots

> maLayout(swirl)

Array layout:         Object of class marrayLayout.

Total number of spots:                       8448
Dimensions of grid matrix:               4 rows by 4 cols
Dimensions of spot matrices:               22 rows by 24 cols

Currently working with a subset of 8448 spots.

Control spots:
There are   2 types of controls :
Control       N
    768    7680
```

```
Notes on layout:
No Input File

> maGnames(swirl)

Object of class marrayInfo.

   maLabels    "ID" "Name"
1     geno1 control  geno1
2     geno2 control  geno2
3     geno3 control  geno3
4     3XSSC control  3XSSC
5     3XSSC control  3XSSC
6      EST1 control   EST1
7     geno1 control  geno1
8     geno2 control  geno2
9     geno3 control  geno3
10    3XSSC control  3XSSC
...

Number of labels:  8448
Dimensions of maInfo matrix:  8448  rows by  2  columns

Notes:
C:/GNU/R/rw1041/library/marrayInput/data/fish.gal
```

As with other microarray objects in Bioconductor packages, you can use subsetting commands for marrayRaw objects

```
> sw <- swirl[1:100, 2]
> class(sw)

[1] "marrayRaw"

> sw

Pre-normalization intensity data:         Object of class marrayRaw.

Number of arrays:         1 arrays.

A) Layout of spots on the array:
Array layout:         Object of class marrayLayout.
```

```
Total number of spots:                        8448
Dimensions of grid matrix:                4 rows by 4 cols
Dimensions of spot matrices:                22 rows by 24 cols

Currently working with a subset of 100 spots.

Control spots:
There are   2 types of controls :
Control       N
     48      52


Notes on layout:
No Input File

B) Samples hybridized to the array:
Object of class marrayInfo.

  maLabels # of slide        Names experiment Cy3 experiment Cy5      date
2       82         82 swirl.2.spot      wild type          swirl 2001/9/20
  comments
2       NA

Number of labels:  1
Dimensions of maInfo matrix:  1  rows by  6  columns

Notes:
C:/GNU/R/rw1041/library/marrayInput/data/SwirlSample.txt

C) Summary statistics for log-ratio distribution:
              Min. 1st Qu. Median  Mean 3rd Qu. Max.
swirl.2.spot -1.08   -0.26  -0.12 -0.12       0 1.72

D) Notes on intensity data:
```

You can access red and green foreground and background intensities, and log ratios as follows

```
> Gb <- maGb(swirl)
> dim(Gb)

[1] 8448    4
```

```
> Gb[1:5, ]

     [,1] [,2] [,3] [,4]
[1,]  182  175   86   97
[2,]  171  183   86   85
[3,]  153  183   86   85
[4,]  153  142   71   87
[5,]  153  142   71   87

> Rf <- maRf(swirl)
> dim(Rf)

[1] 8448    4

> Rf[1:5, ]

     swirl.1.spot swirl.2.spot swirl.3.spot swirl.4.spot
[1,]    19538.470    16138.720    2895.1600   14054.5400
[2,]    23619.820    17247.670    2976.6230   20112.2600
[3,]    21579.950    17317.150    2735.6190   12945.8500
[4,]     8905.143     6794.381     318.9524     524.0476
[5,]     8676.095     6043.542     780.6667     304.6190

> M <- maM(swirl)
> dim(M)

[1] 8448    4
```

## 2 Reading in data: `marrayInput` package

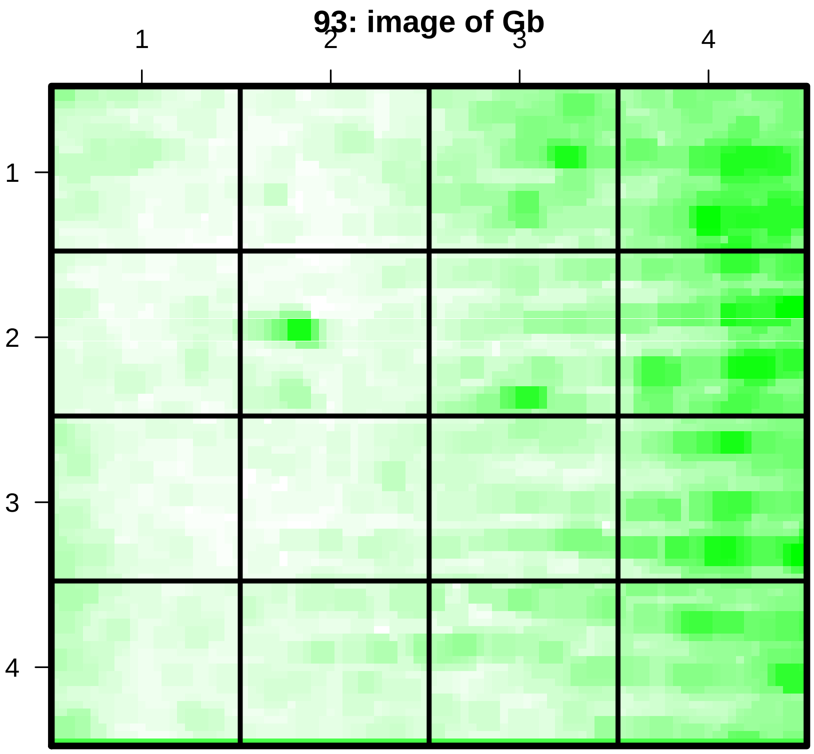
Functions and widgets for reading in spotted array data are provided in the marrayInput package, see for example ? read.marrayRaw or ? widget.marrayRaw.

## 3 Diagnostic plots: `marrayPlots` package

The marrayPlots package provides functions for diagnostic plots of microarray spot statistics.

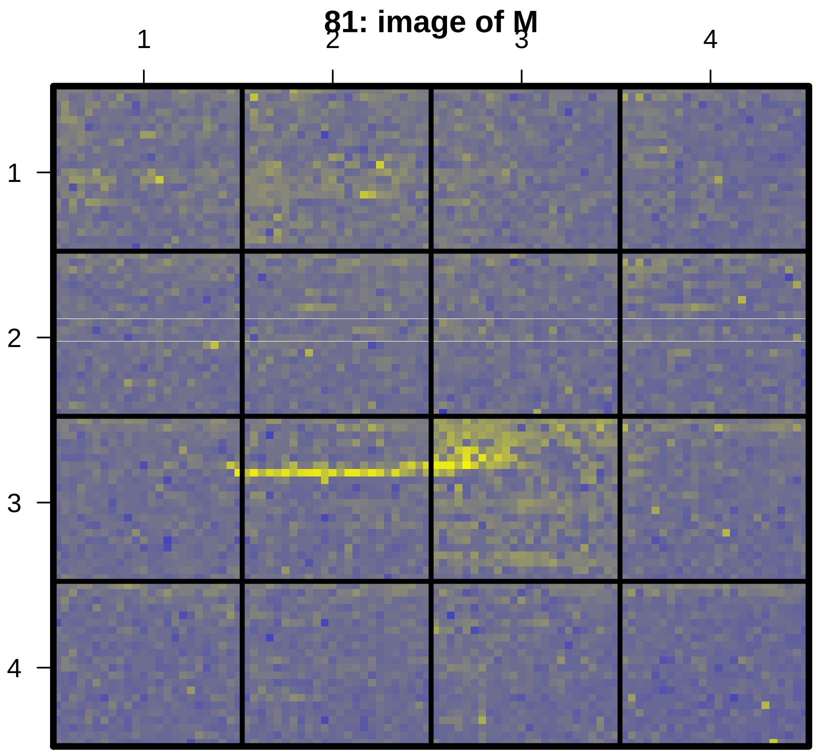
To produce a spatial image of background intensities for the Cy3 channel in the third array

```
> tmp <- maImage(swirl[, 3], x = "maGb", bar = FALSE)
```



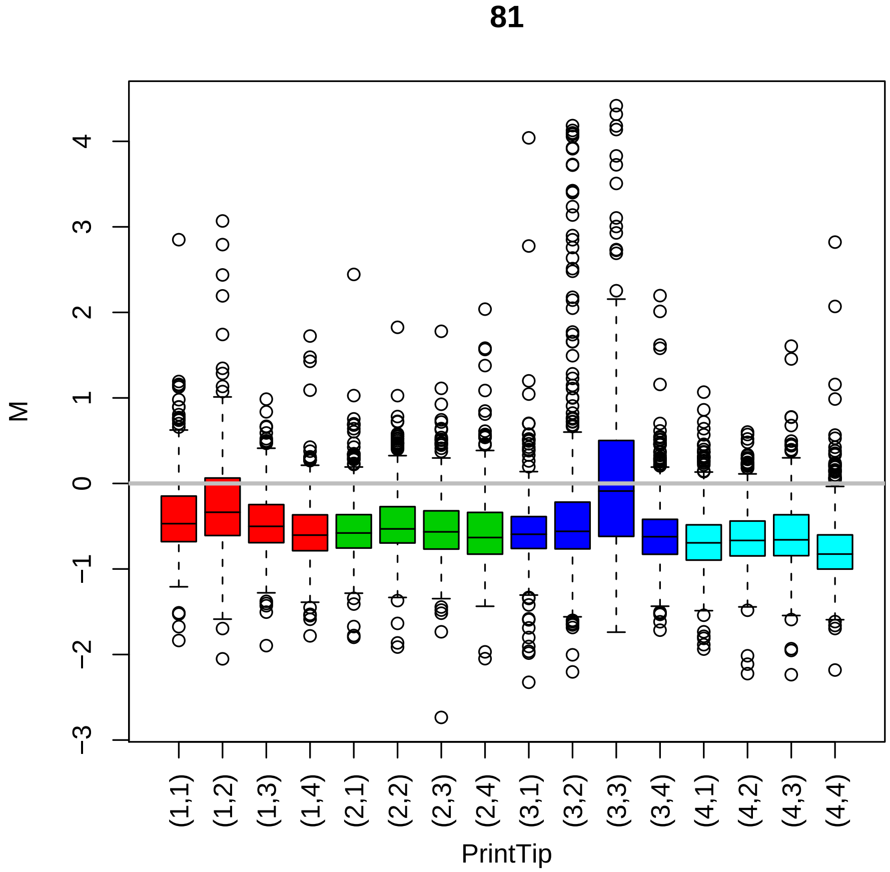
To produce a spatial image of log ratios for the first array in the batch

```
> tmp <- maImage(swirl[, 1], col = maPalette(low = "blue", high = "yellow"),
+     bar = FALSE)
```



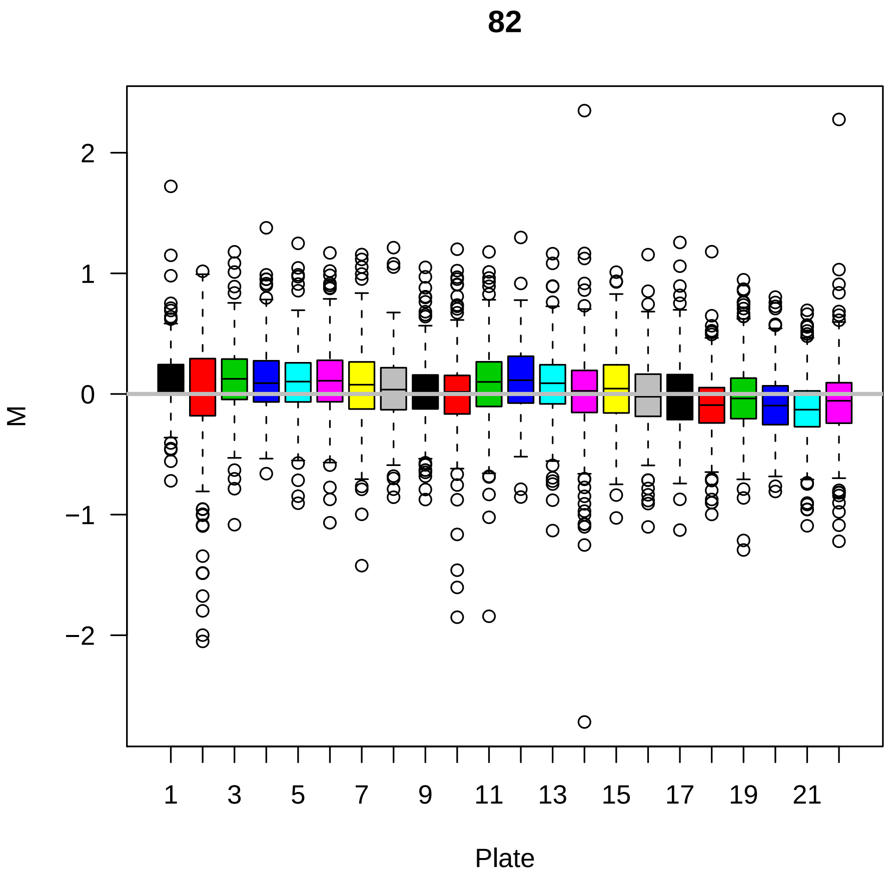
To produce boxplots of log ratios by sector for the first array in the batch

```
> maBoxplot(swirl[, 1])
```



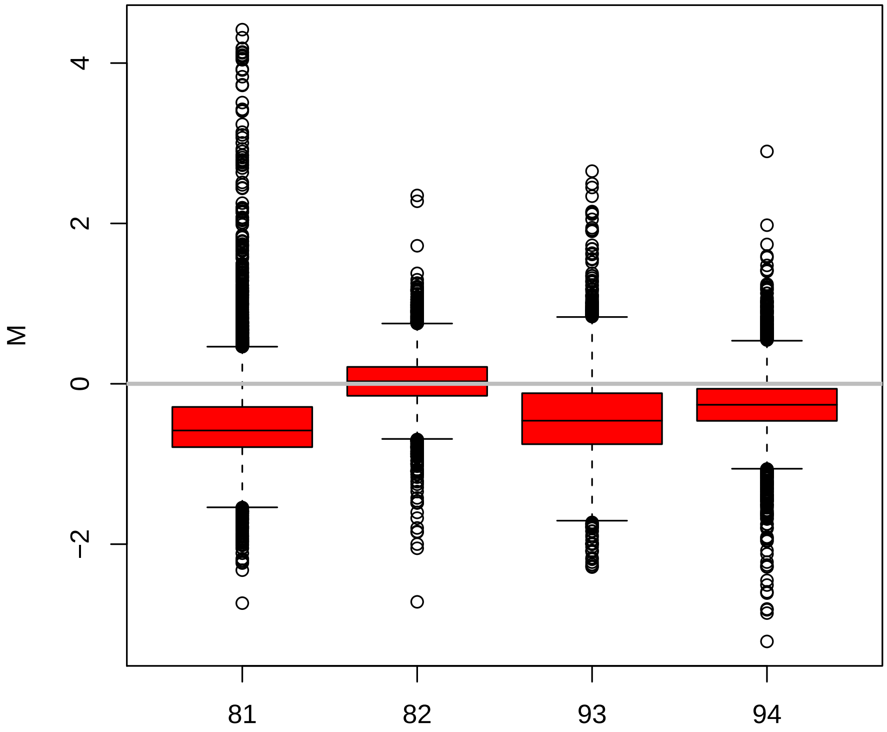
To produce boxplots of log ratios by plate for the second array in the batch

```
> maPlate(swirl) <- maCompPlate(swirl, n = 384)
> maBoxplot(swirl[, 2], x = "maPlate", names = NULL)
```



For boxplots of log ratios for all four arrays

```
> maBoxplot(swirl)
```



## 4 Normalization: `marrayNorm` package

The marrayNorm package implements robust adaptive location and scale normalization procedures, which correct for different types of dye biases (e.g., intensity, spatial, plate biases).

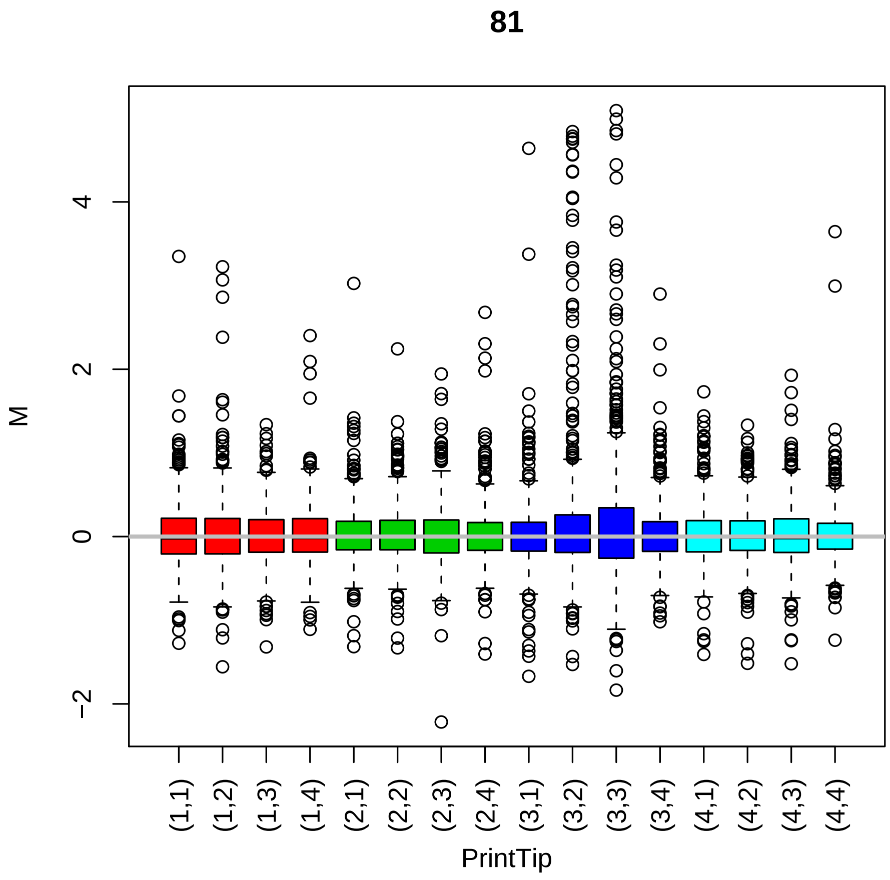
The main location and scale normalization function is maNormMain. Simpler wrapper functions are provided in maNorm and maNormScale. The function operate on objects of class marrayRaw (or possibly marrayNorm, if normalization is performed in several steps) and return objects of class marrayNorm.

For within-print-tip-group loess location normalization of batch swirl

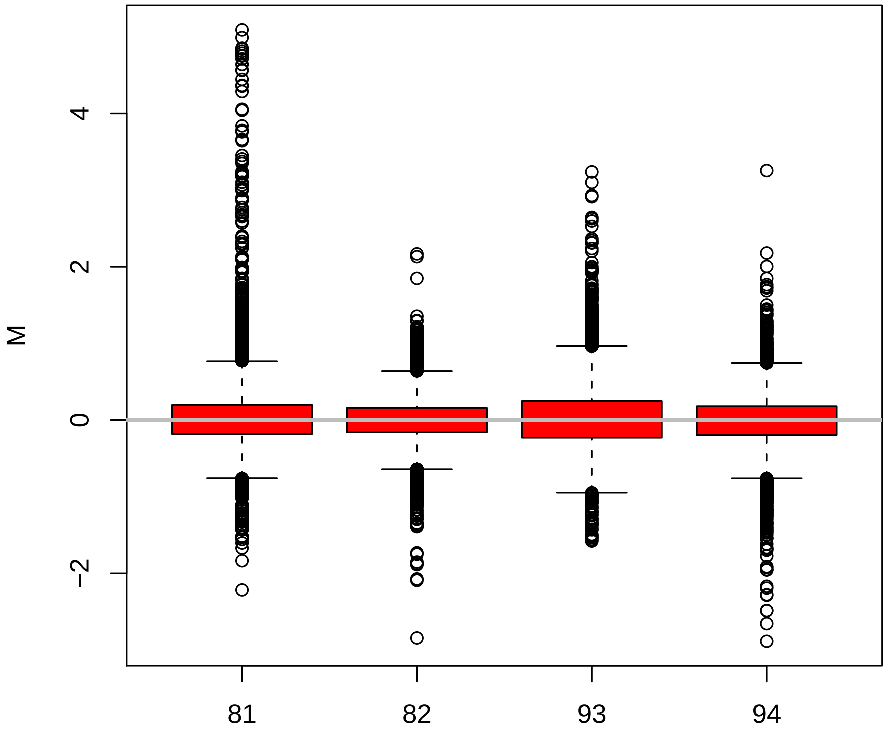
```
> swirl.norm <- maNormMain(swirl)
```

For boxplots of post-normalization log-ratios

```
> maBoxplot(swirl.norm[, 1])
```



```
> maBoxplot(swirl.norm)
```


M
4
2
0
−2
81
82
93
94