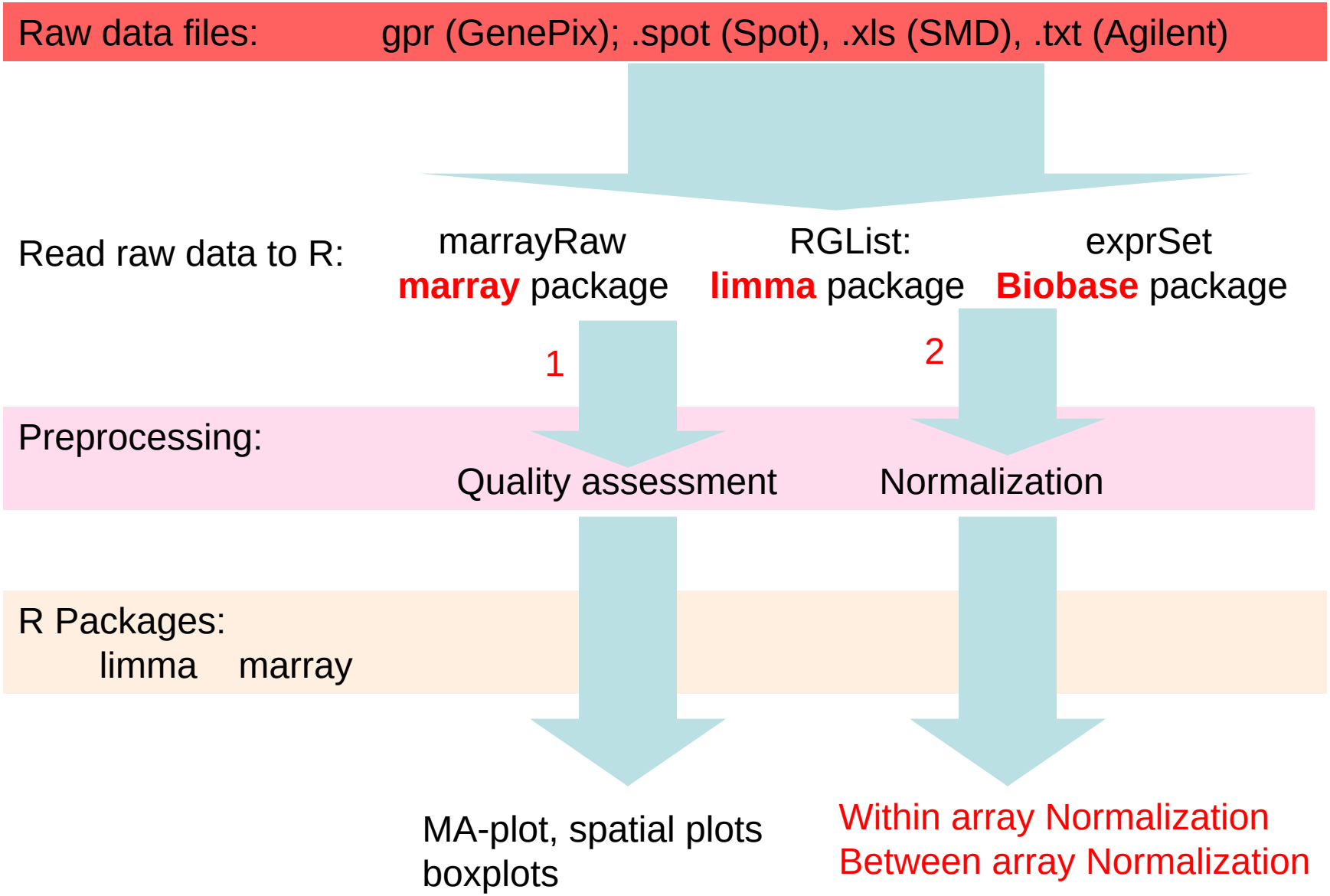
Introduction



Format of array file

```
## read array data from NCBI
library("annotate")
samp.6Hs.166 <- readGEOAnn("GSM16689")
```

Data Information
23184 spots
1760 control spots
Print-tip: 12*4
Spots matrix: 21*23

‘annotate’ package另外還要安裝 ’AnnotationDbi’,’RSQLite’,’ xtable’

samp.6Hs.166

資料編輯器

	ID_REF	X	Y	Dia.	F635 Median	F635 Mean	F635 SD	B635	B635 Median	B635 Mean	B635 SD
1	1	1700	14710	120	139	154	70	137	137	148	66
2	2	1880	14710	70	181	207	109	142	142	149	57
3	3	2070	14710	60	238	244	81	140	140	156	83
4	4	2260	14710	90	228	242	85	146	146	154	70
5	5	2450	14710	90	218	238	94	138	138	147	61
6	6	2640	14710	110	7185	7258	811	140	140	156	203
7	7	2830	14710	110	445	471	163	142	142	155	202
8	8	3030	14710	110	870	872	203	135	135	141	56
9	9	3220	14710	100	255	254	85	138	138	146	64
10	10	3420	14700	80	218	215	79	144	144	148	69

read target information

```
library("marray")
datadir <- system.file("swirldata", package="marray")
swirl.samples <- read.marrayInfo(file.path(datadir, "SwirlSample.txt"))
```

```
An object of class "marrayInfo"
@maLabels
[1] "swirl.1.spot" "swirl.2.spot" "swirl.3.spot" "swirl.4.spot"

@maInfo
        Names slide number experiment Cy3 experiment Cy5      date comments
1 swirl.1.spot        81          swirl       wild type 2001/9/20       NA
2 swirl.2.spot        82      wild type           swirl 2001/9/20       NA
3 swirl.3.spot        93          swirl       wild type 2001/11/8       NA
4 swirl.4.spot        94      wild type           swirl 2001/11/8       NA

@maNotes
[1] "C:/PROGRA~1/R/R-26~1.2/library/marray/swirldata/SwirlSample.txt"
```

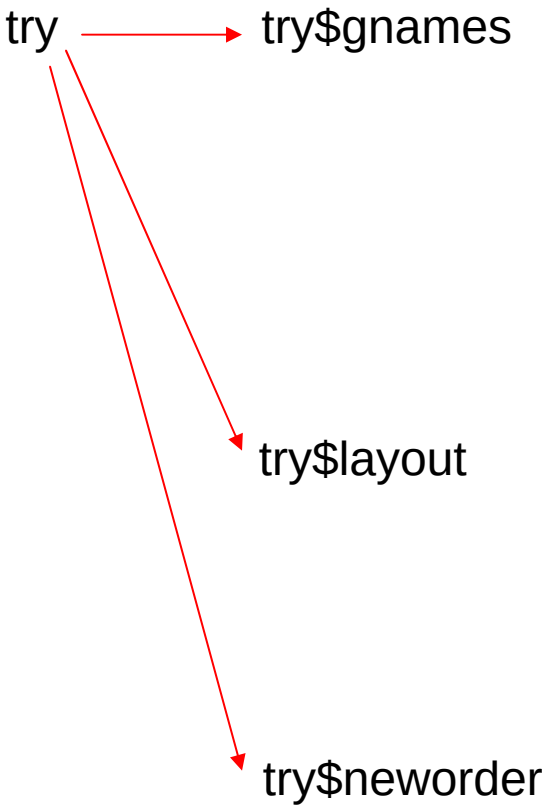
swap

swap

Reading probe-related information

Probe information file type: .gal (for GenePix or Spot)

```
library(marray)
datadir <- system.file("swirldata", package="marray")
try <- read.Galfile(galfile="fish.gal", path=datadir)
```




```
@maLabels
[1] "control" "control" "control" "control" "control"
8443 more elements ...
@maInfo
                ID  Name
control    control geno1
control.1 control geno2
control.2 control geno3
@maNotes
[1] ""
```

```
@maNgr [1] 4
@maNgc [1] 4
@maNsr [1] 22
@maNsc [1] 24
@maNspots [1] 8448
.
.
```

Reading probe and background intensity

read.marrayRaw → read.GenePix, read.SMD, read.Spot, read.Agilent

```
datadir <- system.file("swirldata", package="marray")
swirl.targets <- read.marrayInfo(file.path(datadir, "SwirlSample.txt"))
data <- read.Spot(path=datadir, targets=swirl.targets)
```

data →

maRf = ..., # Object of class matrix
maGf = ..., # Object of class matrix
maRb = ..., # Object of class matrix
maGb = ..., # Object of class matrix
maW = ..., # Object of class matrix
maLayout = ..., # Object of class marrayLayout
maGnames = ..., # Object of class marrayInfo
maTargets = ..., # Object of class marrayInfo
maNotes = ..., # Object of class character

maLayout

maNgr = ..., # Object of class numeric
maNgc = ..., # Object of class numeric
maNsr = ..., # Object of class numeric
maNsc = ..., # Object of class numeric
maNspots = ..., # Object of class numeric
maSub = ..., # Object of class logical
maPlate = ..., # Object of class factor
maControls = ..., # Object of class factor
maNotes = ..., # Object of class character

maGnames

maLabels = ..., # Object of class character
maInfo = ..., # Object of class data.frame
maNotes = ..., # Object of class character

maTargets

maLabels = ..., # Object of class character
maInfo = ..., # Object of class data.frame
maNotes = ..., # Object of class character

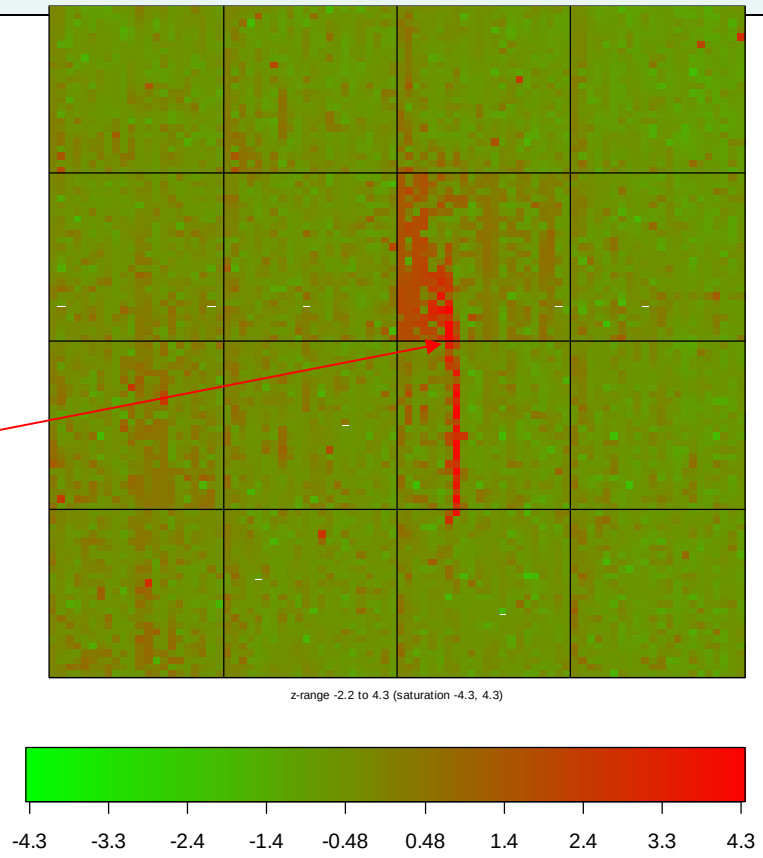
Quality Assessment-1

Array Image

```
#open a widows
nf <- layout(matrix(c(1,2),2,1), 5, c(5,2), TRUE)
layout.show(nf)
# plot the data set
M <- log2(data[,1]@maRf/data[,1]@maGf)
Limcol <- maPalette(low="green", high="red", k=100)
Lim <- round(max(abs(range(M))), 1)
imageplot(M,layout=list(ngrid.r=4,ngrid.c=4,nspot.r=22,nspot.c=24), low="green", high="red", ncolors=100)
maColorBar(Limcol, scale=c(Lim, -Lim), mar=c(3,1,1,1))
```

“M” can be the foreground or background to test the error of array

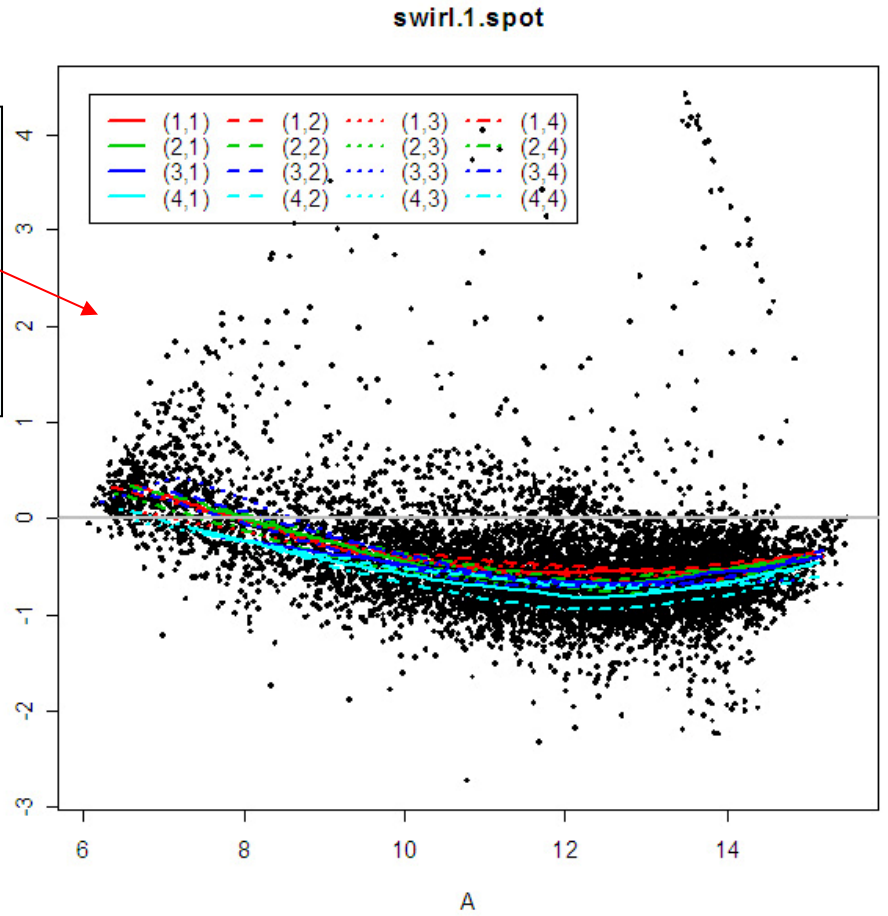
The array may have some problem in this area

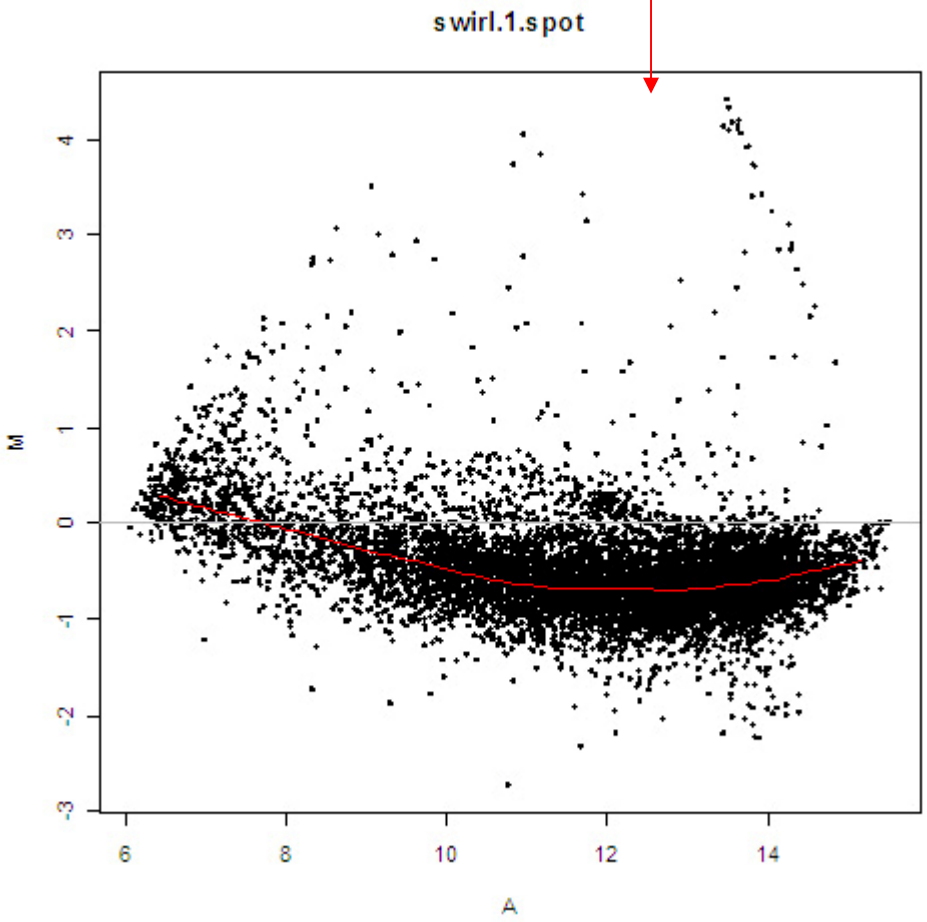


Quality Assessment-2

MAplot

```
# Lowess fit using print-tip spots
maPlot(data[,1])
# Lowess fit using all spots
maPlot(data [,1], z=NULL,
legend.func=NULL)
```





Quality Assessment-3

boxplot

```
boxplot(data, col="red",
        xlab = "Array Channels",
        ylab = "LogRatio",
        main = "Raw data distribution")
```

Q1+1.5*IQR

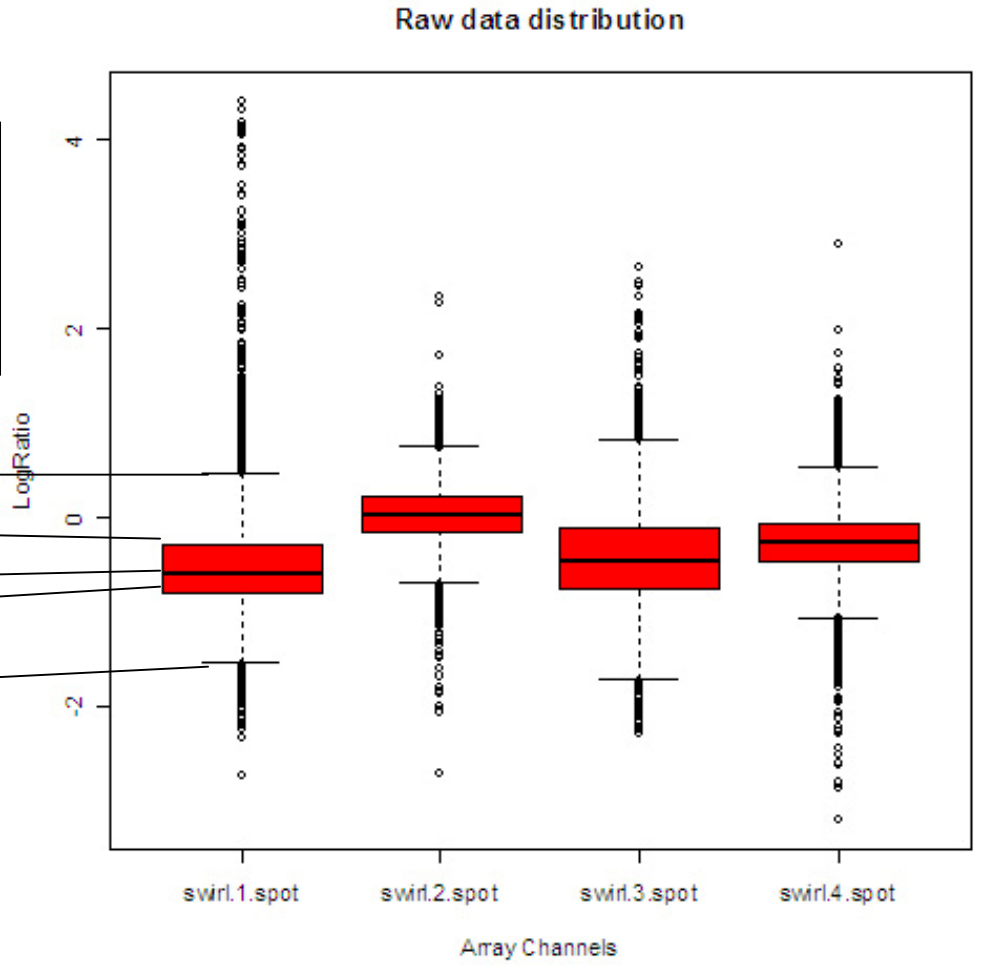
Q1

Medium

Q3

Q3+1.5*IQR

IQR = *Q3* − *Q1*



Normalization

Remove the artifactual biases due to technical factors

1. *Dyes effects*
2. *Scanning parameter*
3. *Print-tip different*
4. *Spatial effects*

Normalization is closely related to “Quality Assessment”

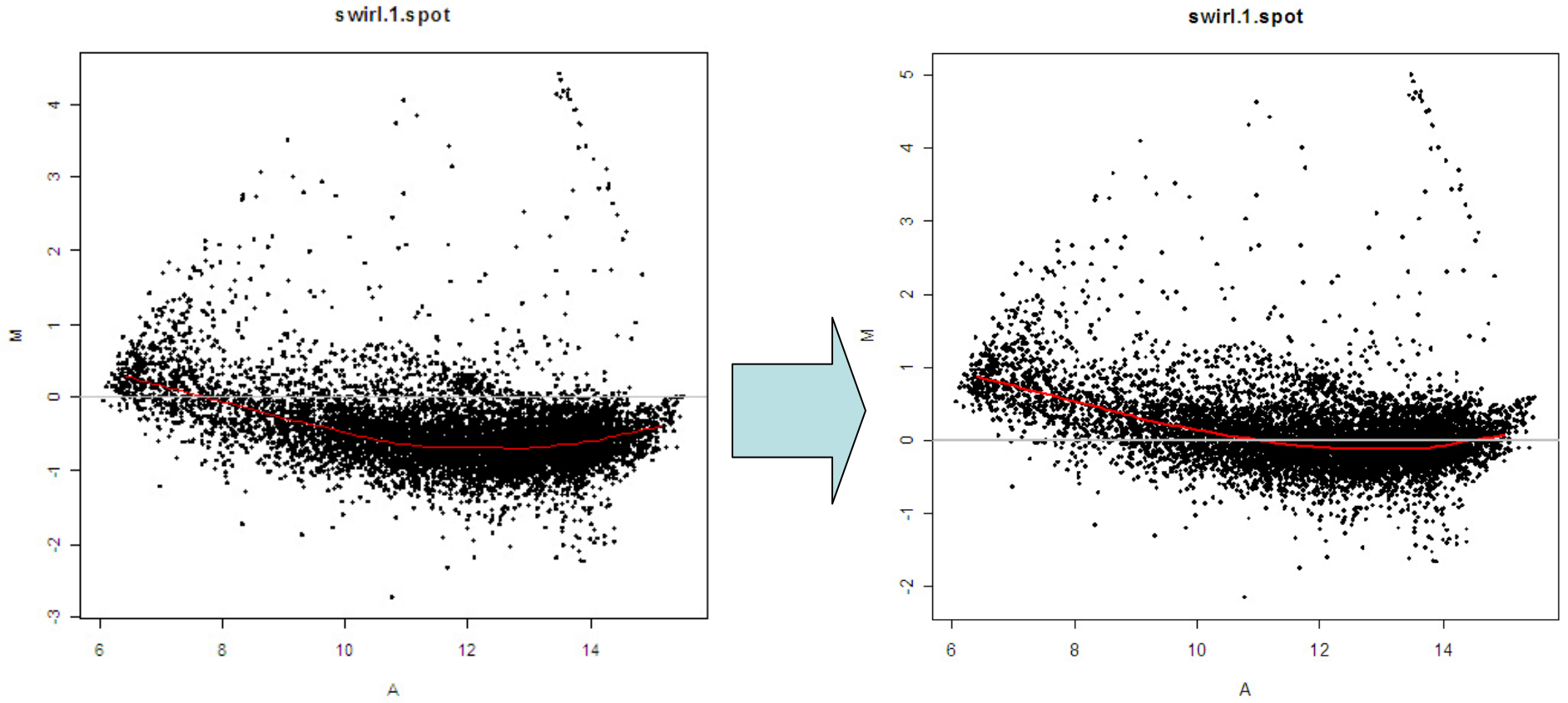
1. *Two-channel Normalization (within array) {marray}*
2. *Separate channel Normalization (Between arrays) {limma}*

Two-channel Normalization

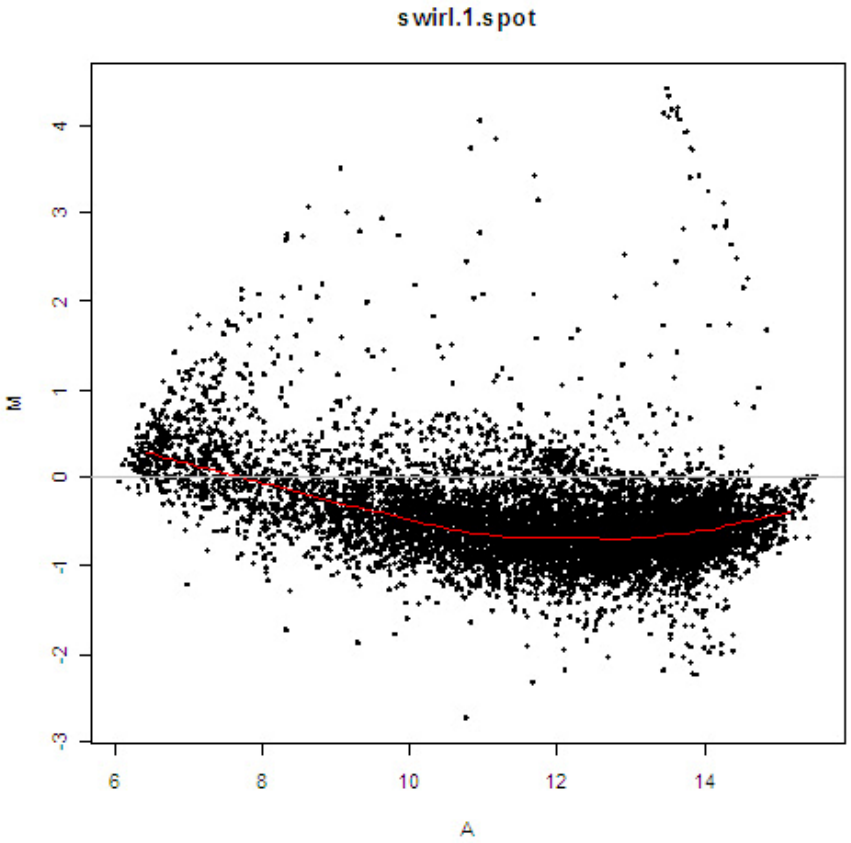
Medium Normalization
Loess Normalization
Print-tip Loess Normalization

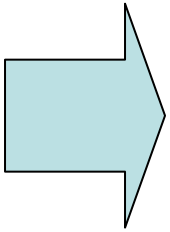
```
data1_M_Norm <- maNorm(data[,1], norm="m")
data1_L_Norm <- maNorm(data[,1], norm="l")
data1_P_Norm <- maNorm(data[,1], norm="p")
```

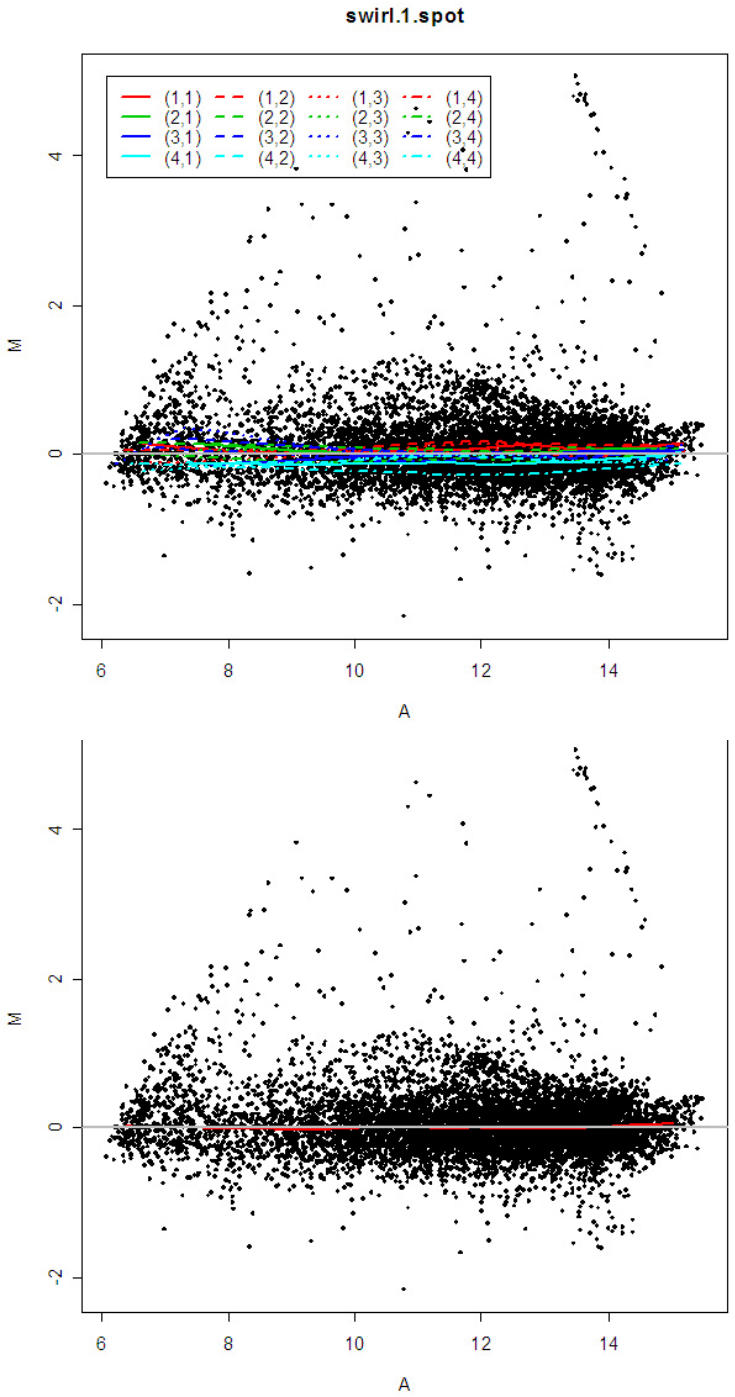
Medium Normalization



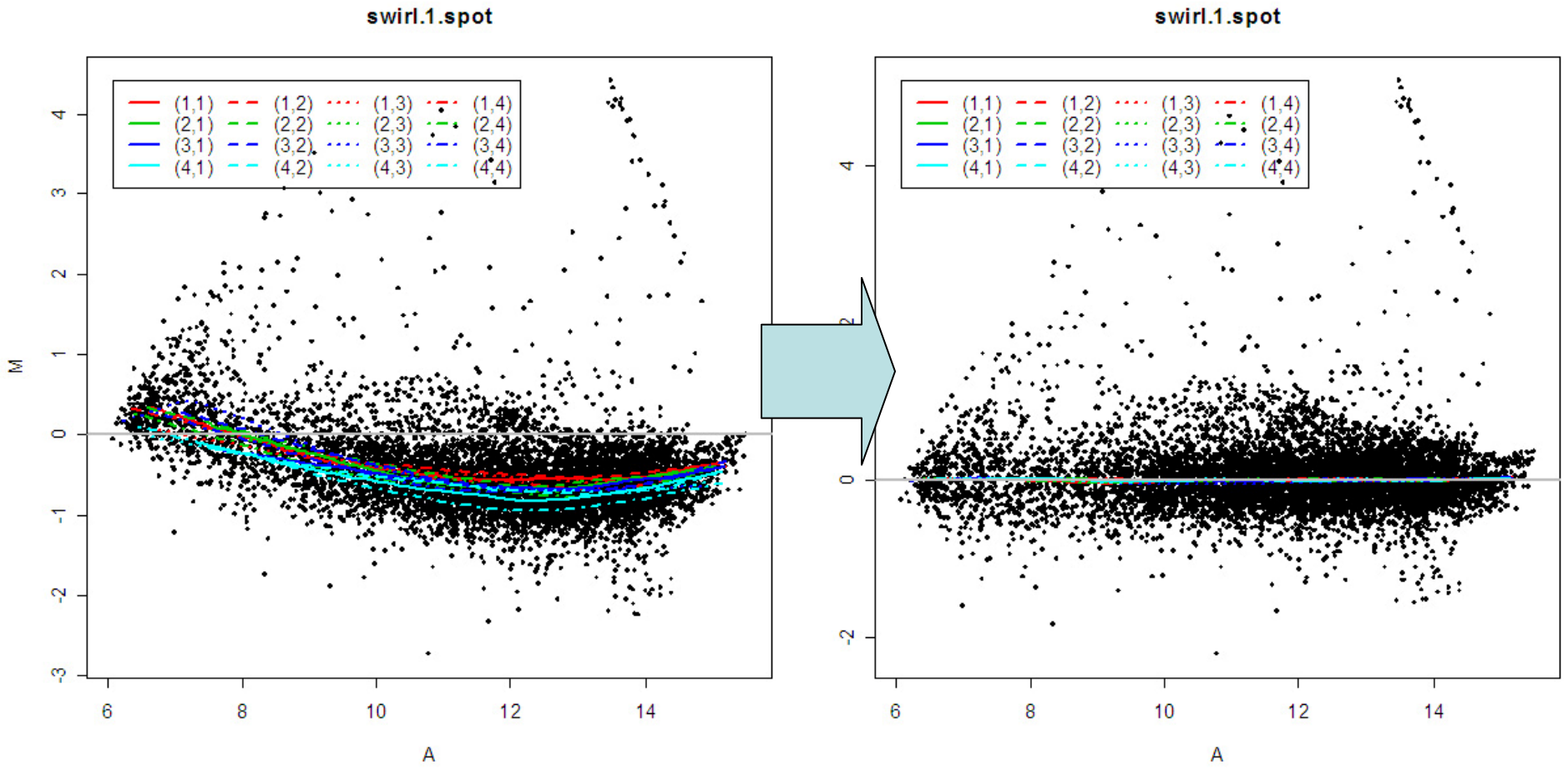
Loess Normalization







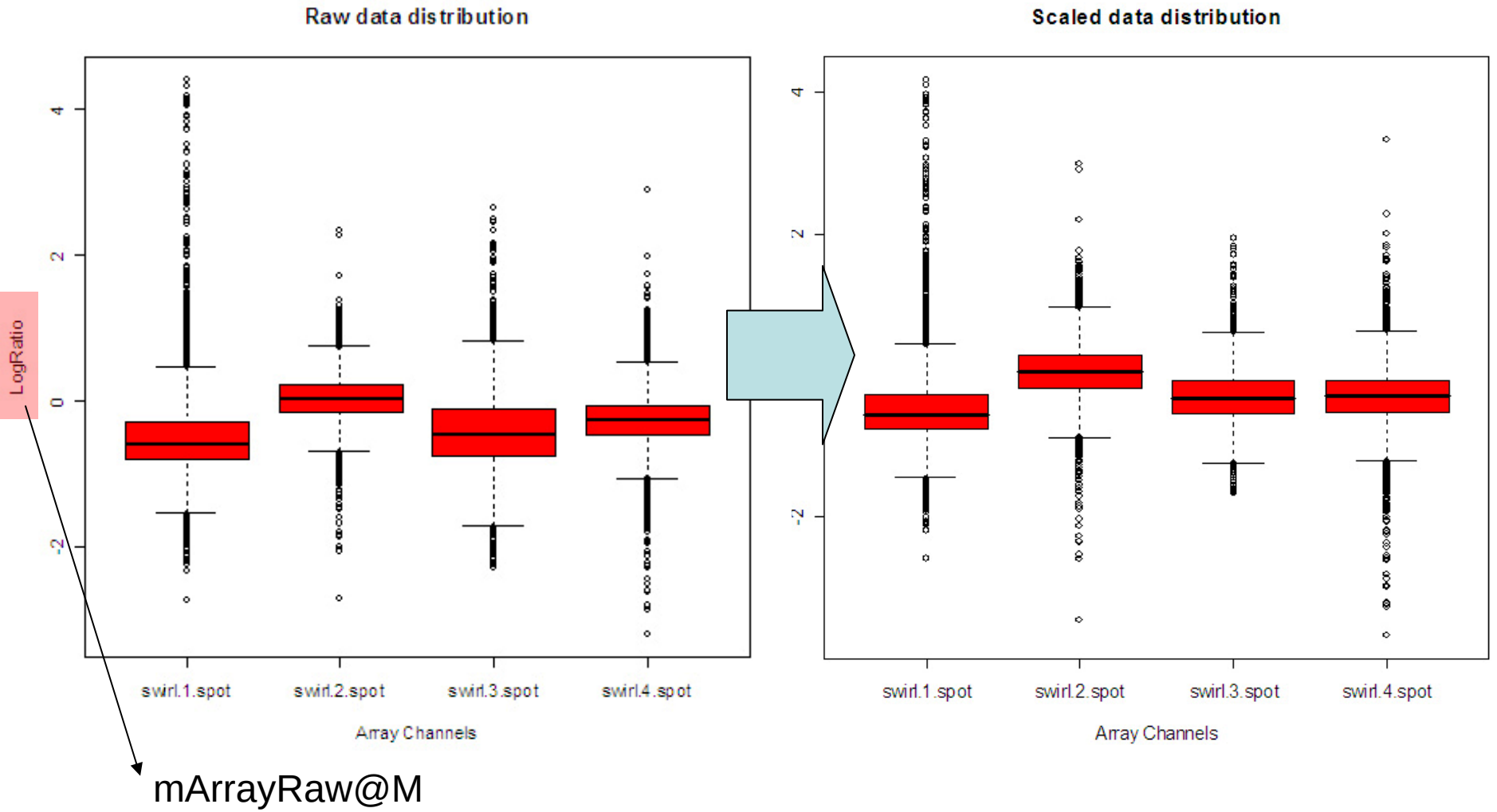
Print-tip Loess Normalization



Two-channel Normalization-between array

Scale Normalization

```
data_S_Norm <- maNormScale(data)
```

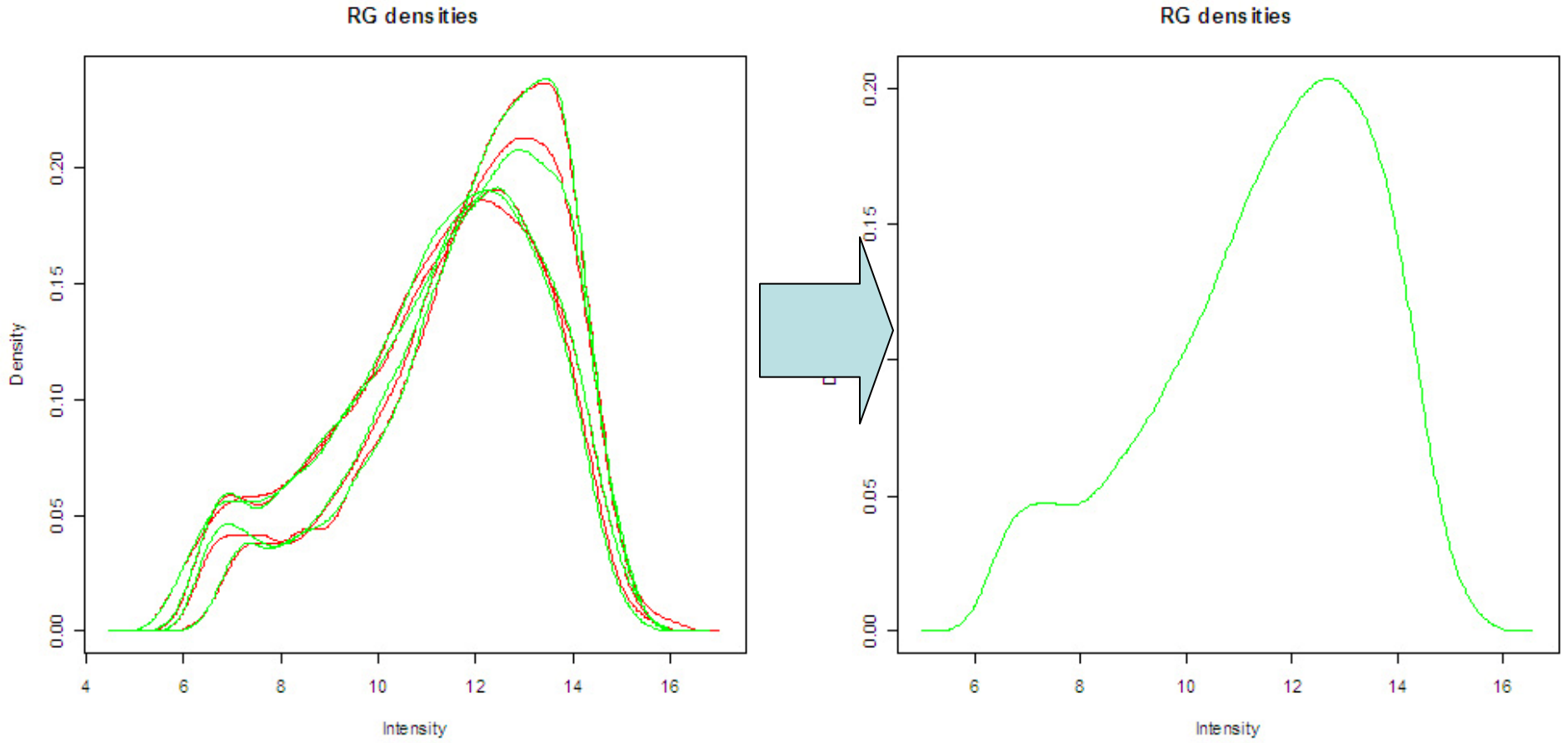


mArrayRaw@M

Separate-channel Normalization

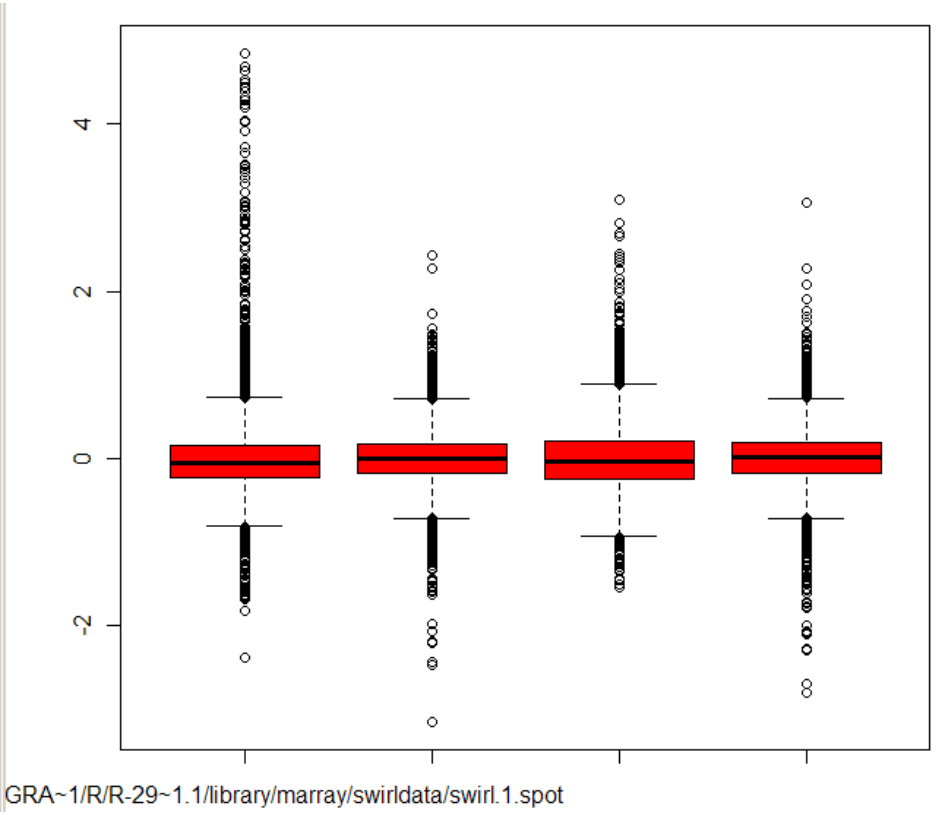
Quantile Normalization

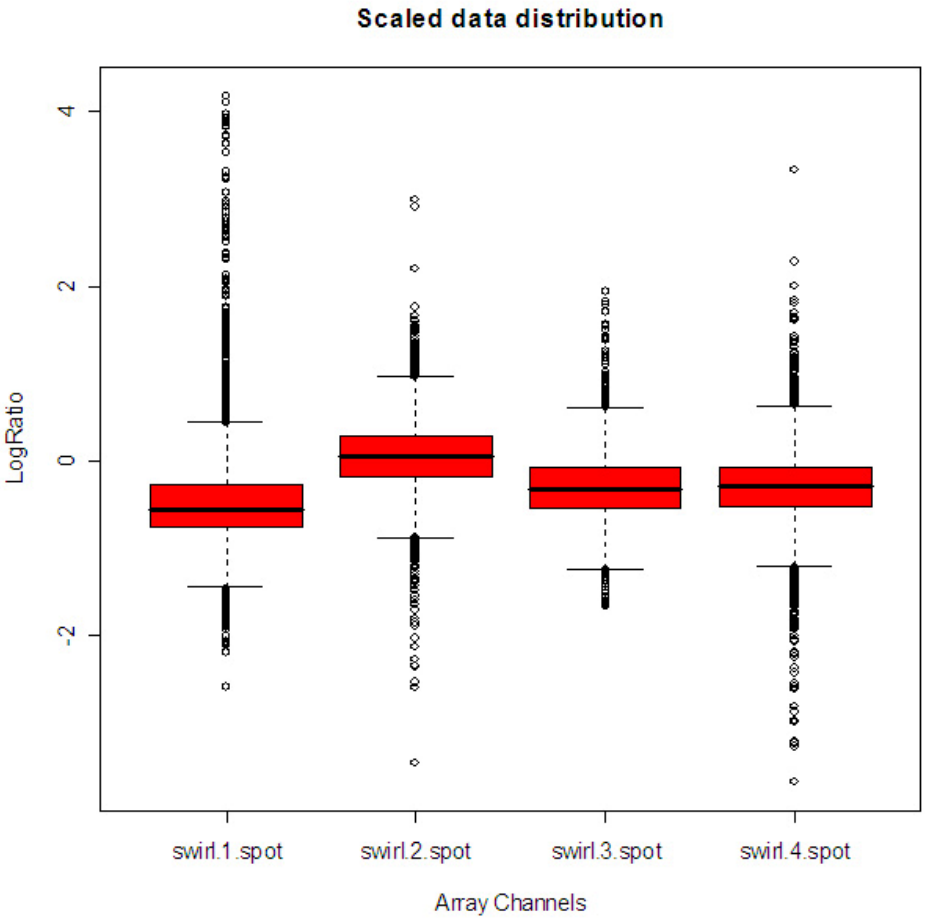
```
data_norm <- maNorm(data, norm="p")
data_MA <- as.MAList(data_norm)
data_MA_B <- normalizeBetweenArrays(data_MA, method="quantile")
plotDensities(data_MA_B)
plotDensities(data_MA)
```



Quantile Normalization ***V.S.*** Scale Normalization

```
boxplot(data_MA_B$M, col="red")
```





Get Normalized data from MAList object

由MAplot公式推知:

Log2(R)=(A*2+M)/2

Log2(G)=(A*2-M)/2

```
data_r=(data_MA_B$A*2+data_MA_B$M)/2
data_g=(data_MA_B$A*2-data_MA_B$M)/2
```

（以data_MA_B當作normalized data）

```
swirl.samples
```

	Names	slide number	experiment Cy3	experiment Cy5	
1	swirl.1.spot	81	swirl	wild type	2001/9/20
2	swirl.2.spot	82	wild type	swirl	2001/9/20
3	swirl.3.spot	93	swirl	wild type	2001/11/8
4	swirl.4.spot	94	wild type	swirl	2001/11/8

swap

swap

Get data from R and G channels

```
wild_type=cbind((data_r[,1]+data_g[,2])/2,(data_r[,3]+data_g[,4])/2)
swirl=cbind((data_g[,1]+data_r[,2])/2,(data_g[,3]+data_r[,4])/2)
```

Count the t test p value to select DEGs

Count p value

```
t_test=c() # pre-set the p value output variable
for(i in 1:nrow(wild_type)){
        t_test[i]=t.test(wild_type[i,],swirl[i,])[['p.value']]
}
```

Set a threshold

```
f1=t_test < 0.05
```

Filter#1

```
> t_test < 0.05
 [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[12] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[23] FALSE FALSE FALSE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE  TRUE
[34] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
```

Filter by labels

```
f2=data_MA_B$genes$Labels != 'control'
```

Filter#2

```
> data_MA_B$genes$Labels != 'control'
 [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[12] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[23] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[34] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[45] FALSE FALSE FALSE FALSE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
[56]  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
```

Count the t test p value to select DEGs

How many are the DEGs??

sum(Filter#1 and Filter#2)

sum(f1 & f2)

Get the names of DEGs

data_MA_B$genes$ID[f1 & f2]

```
> unique(data_MA_B$genes$ID[f1 & f2])
  [1] "fb26a05" "fb32a09" "fb35a03" "fb61a03" "fb84a07" "fb85a01" "fb26b03"
  [8] "fb26b07" "fb50b05" "fb50b07" "fb94b07" "fb97b11" "fc05b03" "fc08b07"
 [15] "fc11b07" "fc13b07" "fb18a06" "fb22a12" "fb32a04" "fb33a10" "fb42a08"
 [22] "fb42a10" "fb48a10" "fb55a08" "fb57a06" "fb85a06" "fb97a12" "fc05a12"
 [29] "fc21a10" "fc24a08" "fb17b10" "fb24b10" "fb25b06" "fb26b10" "fb36b02"
 [36] "fb41b10" "fb48b12" "fb57b04" "fb65b04" "fb98b10" "fc06b10" "fb25c05"
```