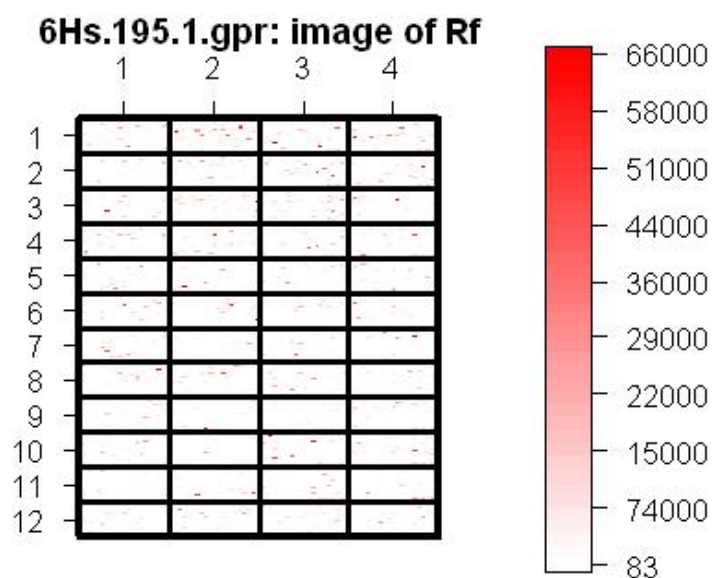


5.5 Representación gráfica de datos de experimentos de microarreglos.

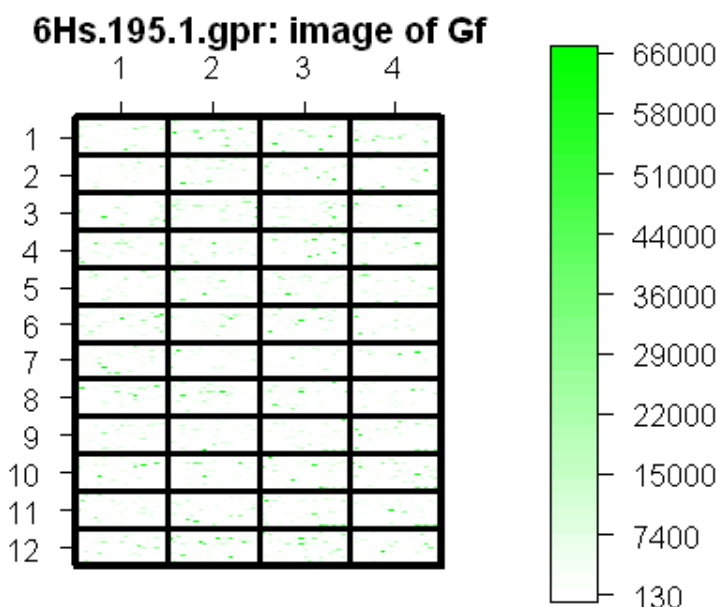
5.5.1 Gráficos espaciales usando la función `image` del paquete `marray`

Los gráficos espaciales son gráficos de dos dimensiones en el cual cada punto representa un estadístico del spot del microarreglo. Se obtienen mediante la función `image` y se utilizan para explorar la calidad de hibridación de cada arreglo.

```
> image (beta7[,1],xvar="maRf")
```



```
> image (beta7[,1],xvar="maGf")
```



Las figuras permiten visualizar la distribución de las intensidades del primer microarreglo para cada color luego del procesamiento de imágenes. En general se espera que estas imágenes no presenten zonas más intensas que otras.

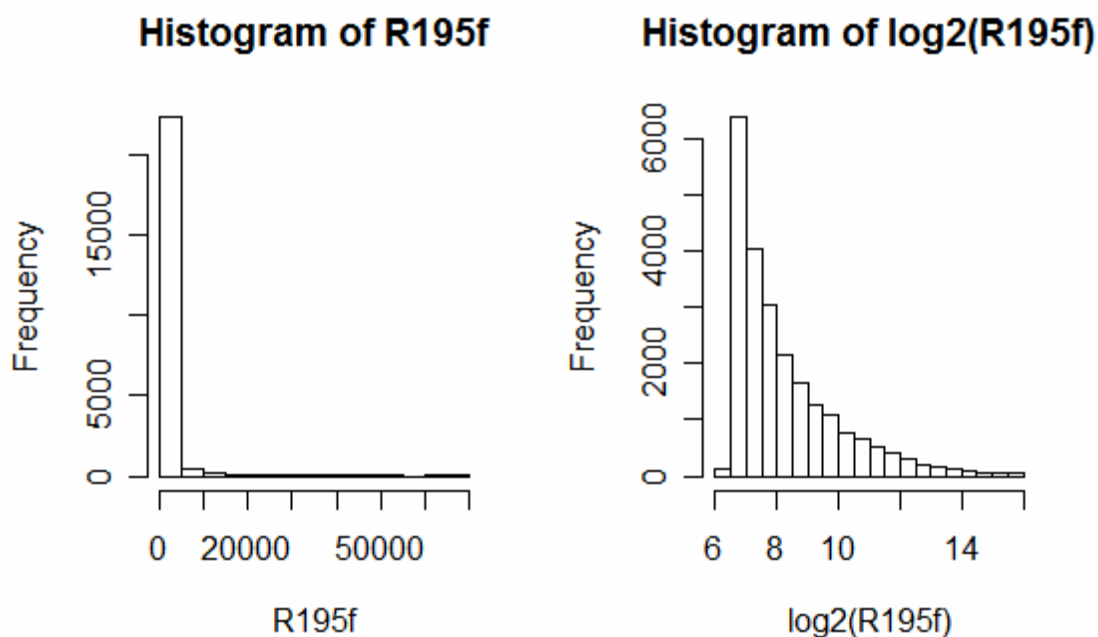
5.5.2 Histogramas

Asignamos a los vectores **R195f** y **G195f** los valores del foreground de los canales Rojo y Verde del primer arreglo respectivamente para obtener el histograma de las correspondientes intensidades utilizando la función **hist** de R.

```
R195f <- slot(beta7,"maRf")[,1] #primera columna de la matriz maRf
G195f <- slot(beta7,"maGf")[,1] #primera columna de la matriz maGf
```

Ya hemos visto que los datos de 6Hs.195.1.gpr están en la primera columna de beta7

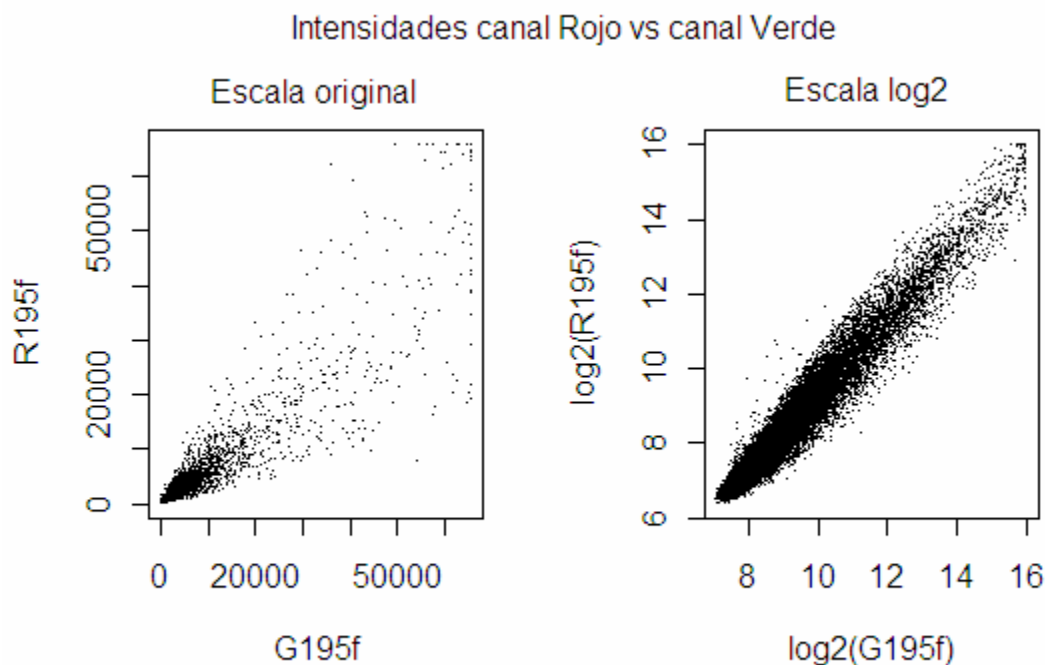
Vemos la distribución de las intensidades del canal rojo del primer arreglo en escala original y en escala logarítmica en base 2



Obtuvimos los gráficos en R con

```
> par(mfrow=c(1,2))
> hist(R195f)
> hist(log2(R195f))
```

5.5.3 Diagramas de Dispersión



```
> par(mfrow=c(1,2))
> plot(G195f,R195f, pch=".")
> plot(log2(G195f),log2(R195f), ",pch=".")
```

5.5.4 MAplot sin corrección por background

El gráfico más común es el denominado MA-plot. Se grafican los valores M (también conocido como log-ratio) en el eje vertical y los valores A en el horizontal.

¿Cómo se calculan dichos valores?

Se definen los valores M y A de la siguiente manera.

$$M = \log_2 \frac{R_f}{G_f}$$

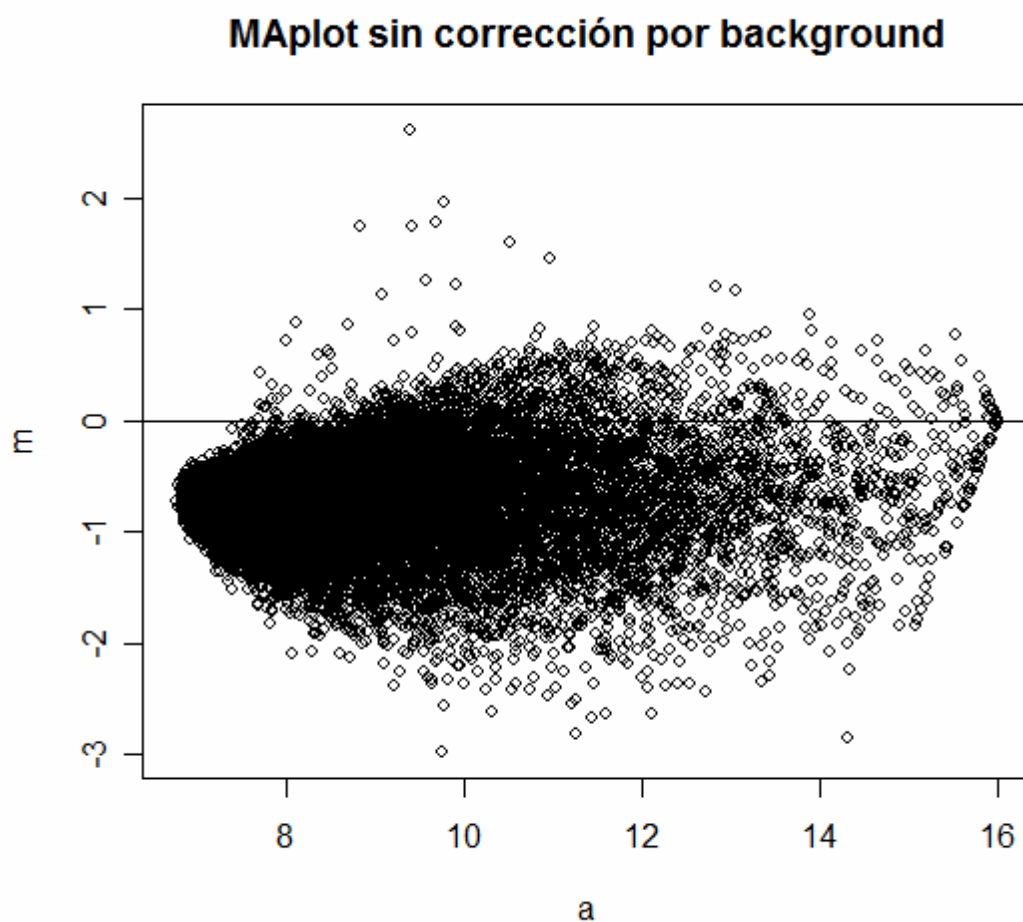
$$A = \frac{\log_2 R_f + \log_2 G_f}{2}$$

Para construir un gráfico MA-plot, para el primer arreglo de beta7, en R ingresamos los siguientes comandos

```
> m <- log2(R195f/G195f)
> a <- 0.5*log2(R195f*G195f)

> par(mfrow=c(1,1))
> plot(a,m, main="MAplot sin corrección por background")
> abline(0,0)
```

y obtenemos

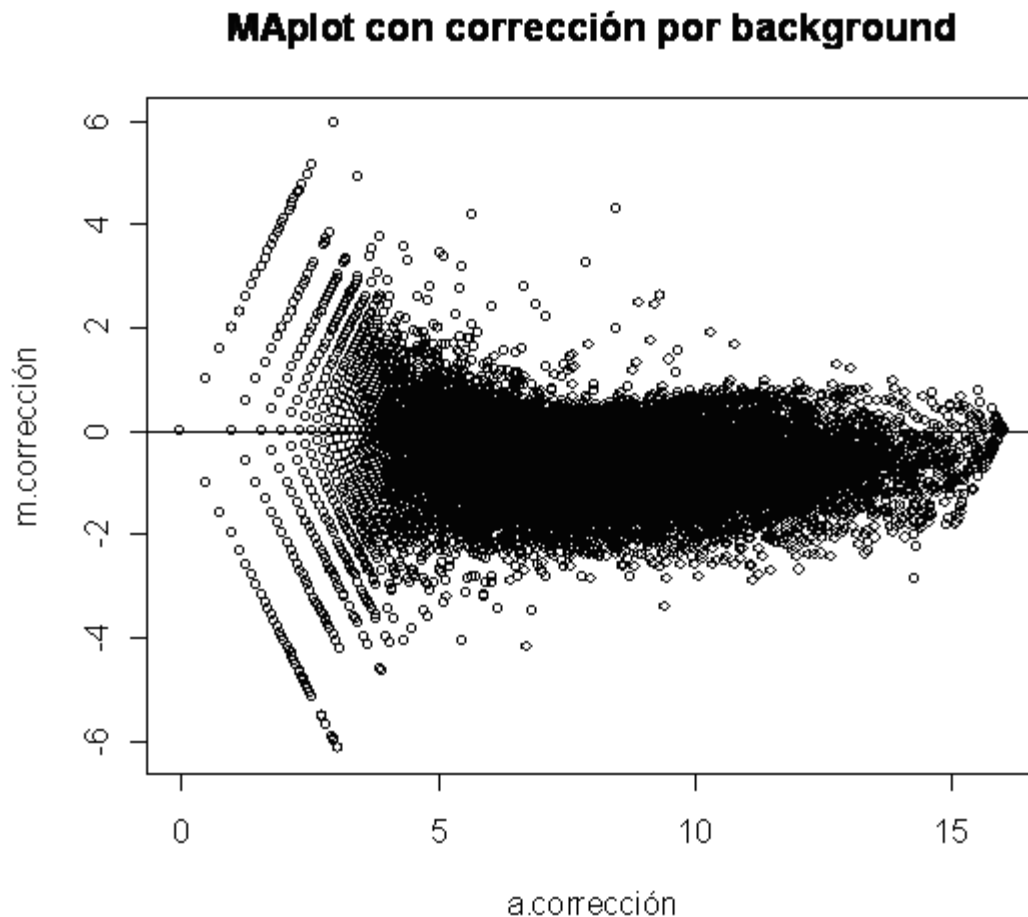


Un gráfico MA-Plot ideal es aquel en el cual toda la nube de puntos se distribuye en forma pareja alrededor del 0 del eje M.

Todas aquellas desviaciones con respecto al ideal se deben principalmente a errores sistemáticos vinculados con varios factores, como ser: diferencias en la eficiencia de la incorporación de los tintes, diferencias en la cantidad de mRNA en las muestras, diferencias en los parámetros de escaneado, diferencias entre los grupos de agujas, efectos espaciales y efectos debido al plato.

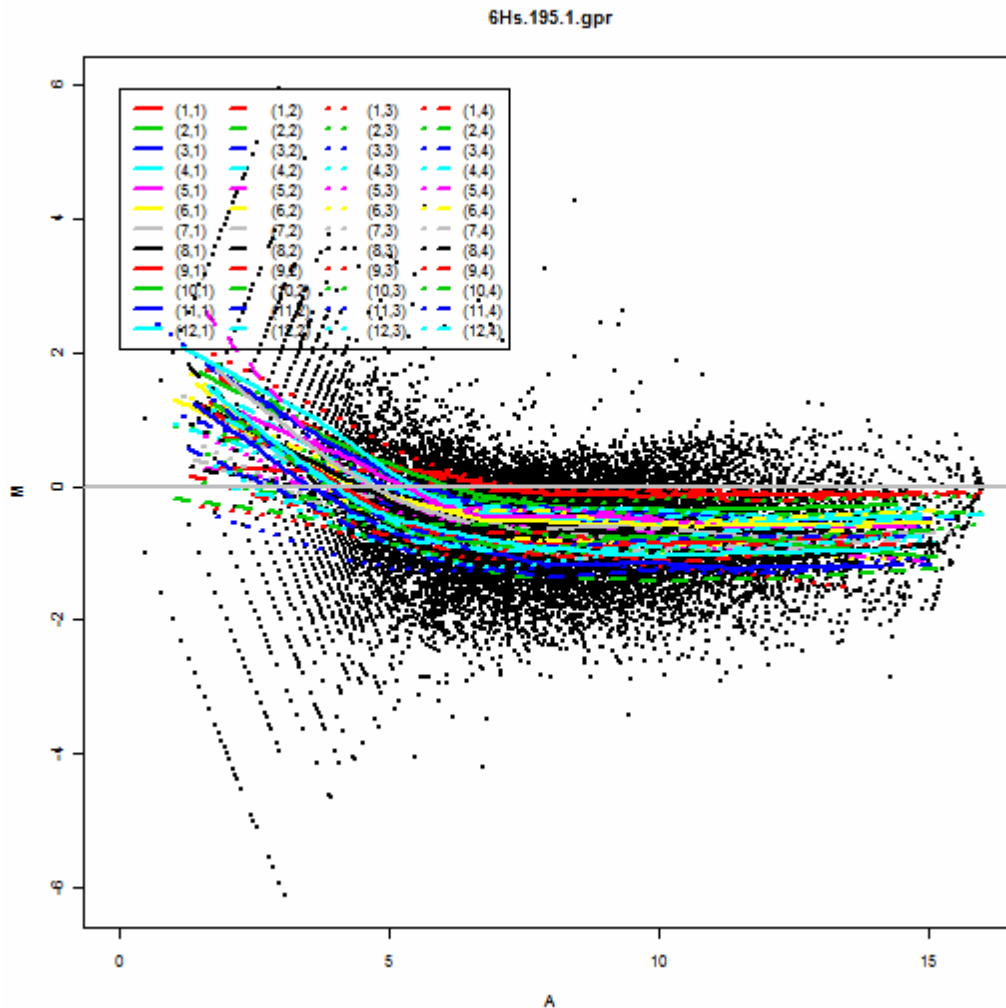
5.5.5 MApot con corrección por background

```
> par(mfrow=c(1,1))
> m.corrección <- log2((R195f-R195b)/(G195f-G195b))
Warning message:
NaNs produced in: log(x, base)
> a.corrección <- 0.5*log2((R195f-R195b)*(G195f-G195b))
Warning message:
NaNs produced in: log(x, base)
>
>
> plot(a.corrección,m.corrección, main="MAplot con corrección por background")
> abline(0,0)
```



5.5.5 Función maPlot

```
X11( pointsize =7)  
maPlot(beta7[,1])
```



¿Qué son las curvas que aparecen por defecto cuando se realiza un maPlot?

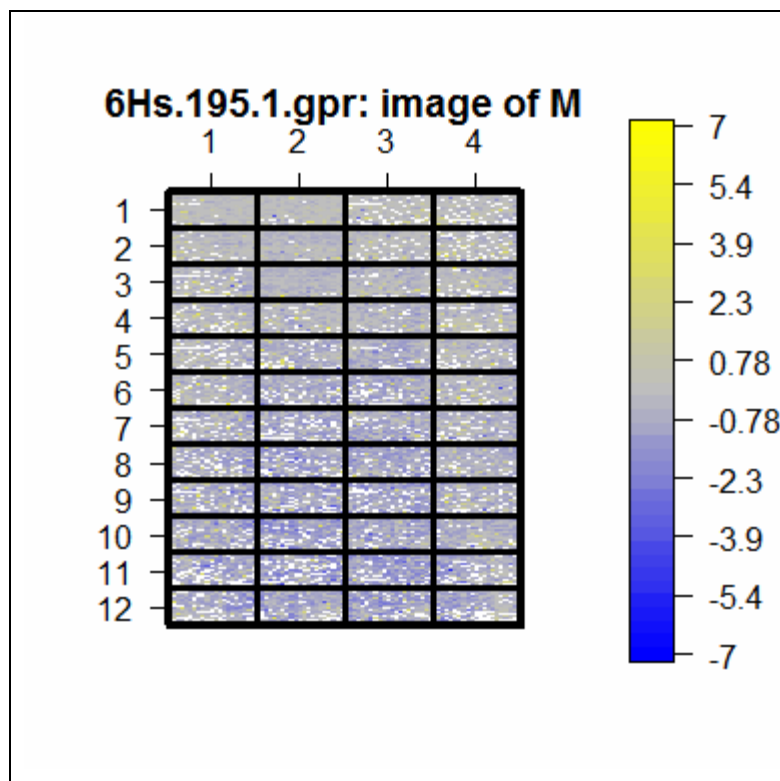
Son curvas Print tip Loess

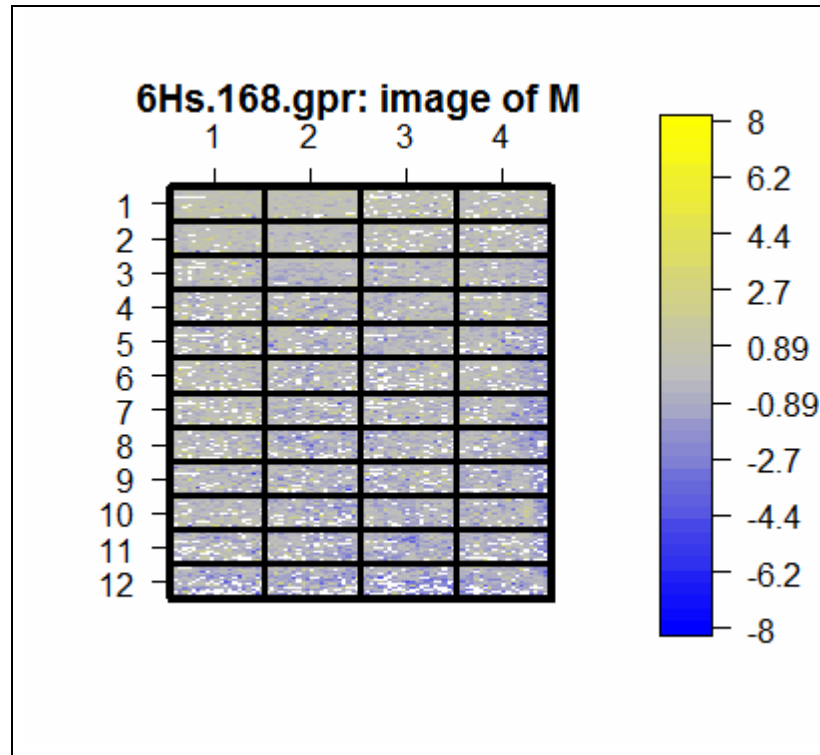
Veremos algo más sobre estas curvas en la próxima clase.

5.6 Más sobre gráficos espaciales

Por defecto, si no se utiliza el valor del parámetro **xvar** en la función `image`, se obtienen los gráficos espaciales para los valores de M . Los siguientes comandos permiten obtener los gráficos espaciales para los valores de M de los **dos primeros** arreglos de `beta7`, con `X11()` se abre una ventana para cada gráfico:

```
for(i in 1:2) {  
  X11()  
  image(beta7[,i])  
}
```

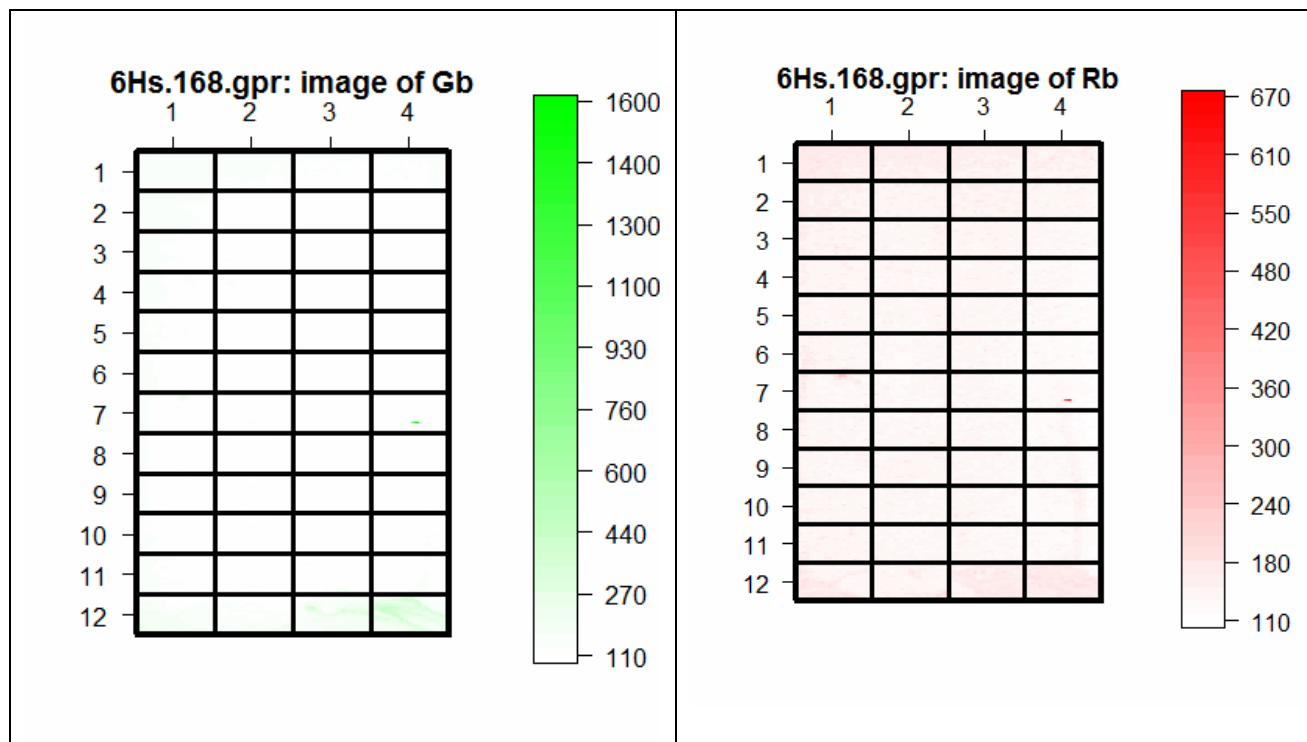




Como ya vimos que es posible modificar el valor por defecto utilizando el parámetro `xvar` Si se utiliza `xvar = "maRb"` se obtiene la el gráfico espacial para el background para el canal Cy5 (rojo). Por defecto se despliega la barra de escala de colores, **`bar=TRUE`**, que se puede eliminar con **`bar= FALSE`**.

```
for(i in 1:2) {  
  X11()  
  image(beta7[,i],xvar="maRb")  
}
```

```
for(i in 1:2) {  
  X11()  
  image(beta7[,i],xvar="maGb")  
}
```

5.7 Boxplots

El boxplot, es un gráfico propuesto por Tukey en 1977. Está basado en medidas resumen de la distribución de un conjunto de datos. Las medidas resumen son: la mediana, los cuartiles superior e inferior, el rango y posiblemente algunos valores extremos. La caja central del gráfico representa el *rango intercuartil (IQR)*, está determinada por el percentil del 25% y el del 75%, contiene el 50% central de los datos. La línea dentro de la caja representa a la mediana, una medida de posición del centro de los datos. Los valores que se encuentran más allá de **1.5 IQR** por encima del percentil del 75% y por debajo del percentil del 25% se grafican individualmente.

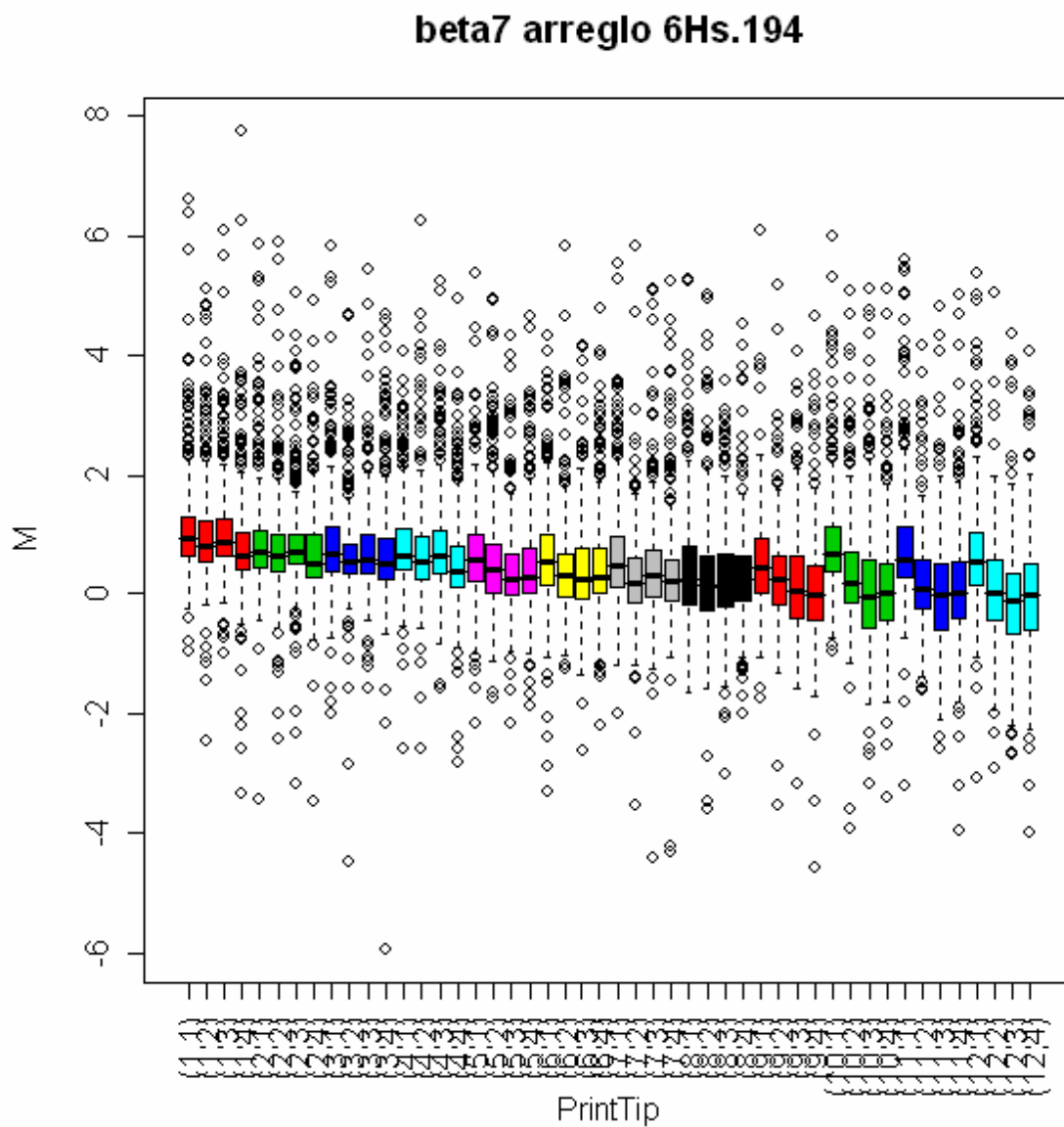
La función **boxplot** produce boxplots para objetos de clase **marrayRaw** y **marrayNorm** que veremos próximamente.

Los argumentos principales de la función **boxplot** son:

- x: Objeto de clase **marrayRaw** o **marrayNorm**.
- xvar: Nombre del método para estratificar los datos. Por ejemplo el nombre de un slot de un objeto "layout" como **maPlate** o un método como **maPrintTip**. Si xvar = NULL, no se estratifican los datos.
- yvar: Nombre del método para la obtención del estadístico de interés por

ejemplo maM.

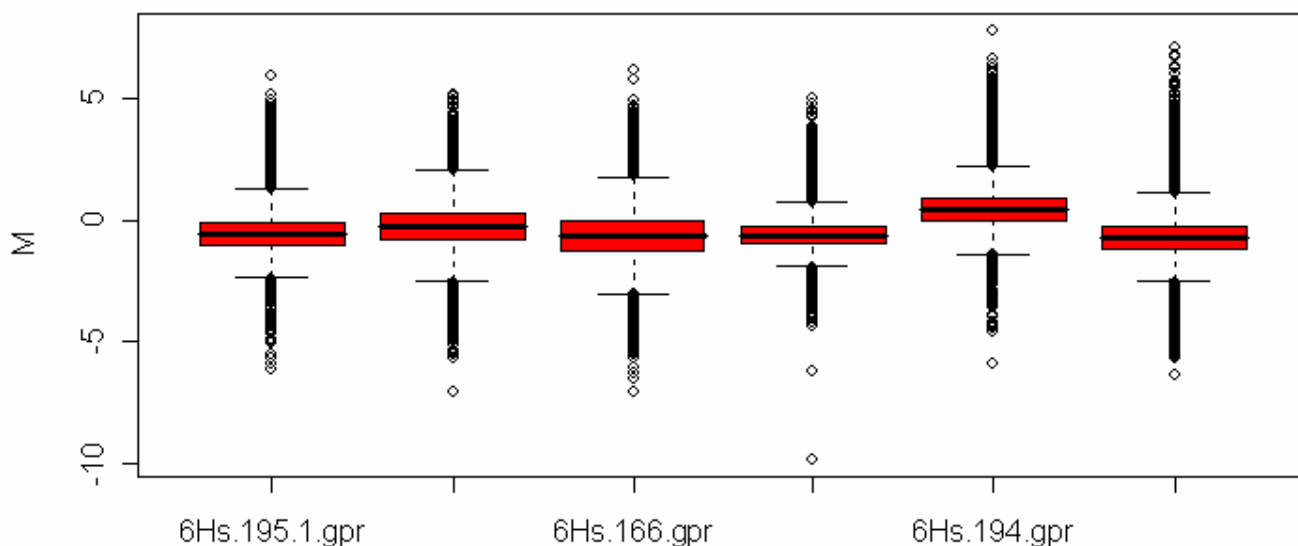
Con
`> boxplot(beta7[, 5], xvar = "maPrintTip", yvar = "maM", main="beta7 arreglo 6Hs.194")`
 obtenemos



Vemos una tendencia decreciente en los cocientes (ratios) de intensidades en relación al print-tip group en el arreglo 5. Los valores correspondientes a las agujas de la primera fila son superiores a los de la última. Este sesgo se corregirá mediante la normalización por grupo de aguja.

Por defecto la función boxplot produce un boxplot para cada arreglo:

```
> boxplot(beta7,yvar="maM")
```

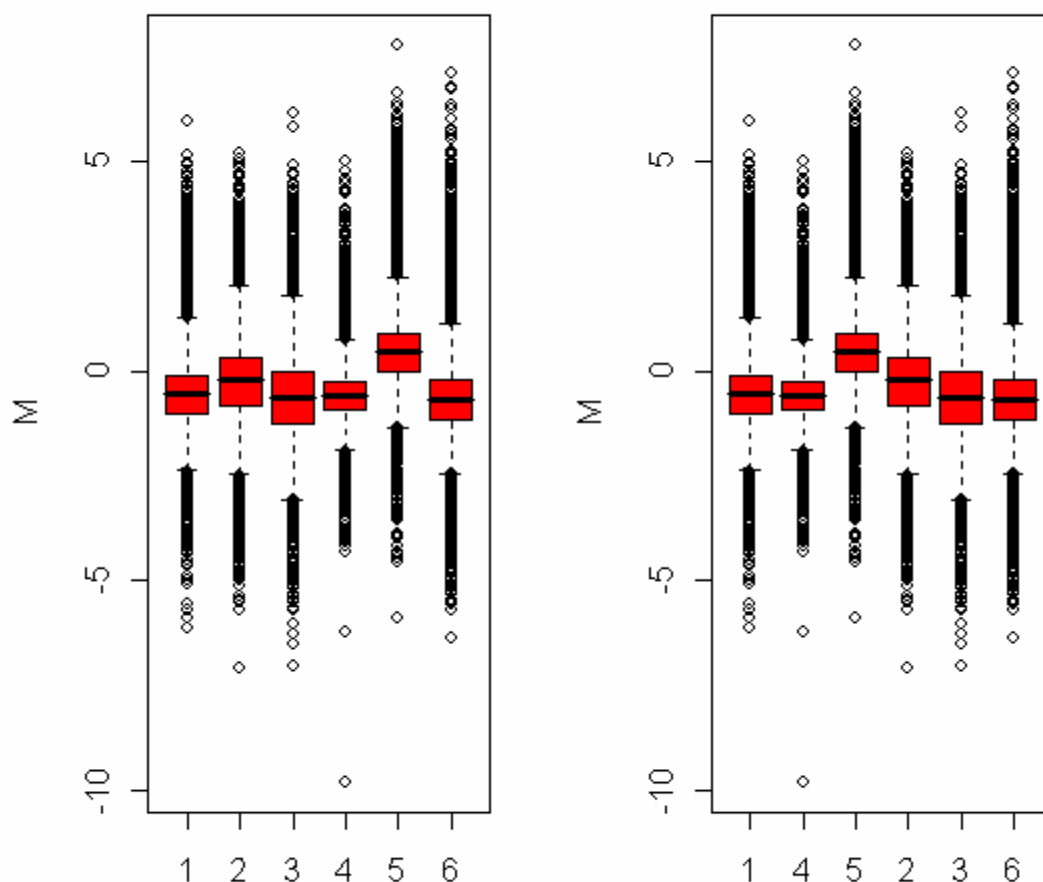


Recordemos

```
@maInfo
```

	FileNames	SubjectID	Cy3	Cy5	Date of Blood Draw
1	6Hs.195.1.gpr	1	b7 -	b7 +	2002.10.11
2	6Hs.168.gpr	3	b7 +	b7 -	2003.01.16
3	6Hs.166.gpr	4	b7 +	b7 -	2003.01.16
4	6Hs.187.1.gpr	6	b7 -	b7 +	2002.09.16
5	6Hs.194.gpr	8	b7 -	b7 +	2002.09.18
6	6Hs.243.1.gpr	11	b7 +	b7 -	2003.01.13

```
> par(mfrow=c(1,2))
> boxplot(beta7,yvar="maM",names=c("1","2","3","4","5","6"))
> boxplot(beta7[,c(1,4,5,2,3,6)], yvar = "maM",
names=c("1","4","5","2","3","6"))
```



5.8 Más opciones en los gráficos espaciales

Hemos visto que la función **image** crea un gráfico de dos dimensiones que representa espacialmente al microarreglo. Los gráficos pueden estar en escala de grises o de colores, en los que cada punto representa un estadístico del spot del microarreglo. Esta función puede ser utilizada para explorar si los datos presentan efectos espaciales. Los detalles de los argumentos de la función se encuentran con **?maImage**. Los estadísticos pueden estar dados por **maM** (intensidad del log Ratio), intensidad del background para el canal verde **maGb**, una medida de calidad del spot, etc.

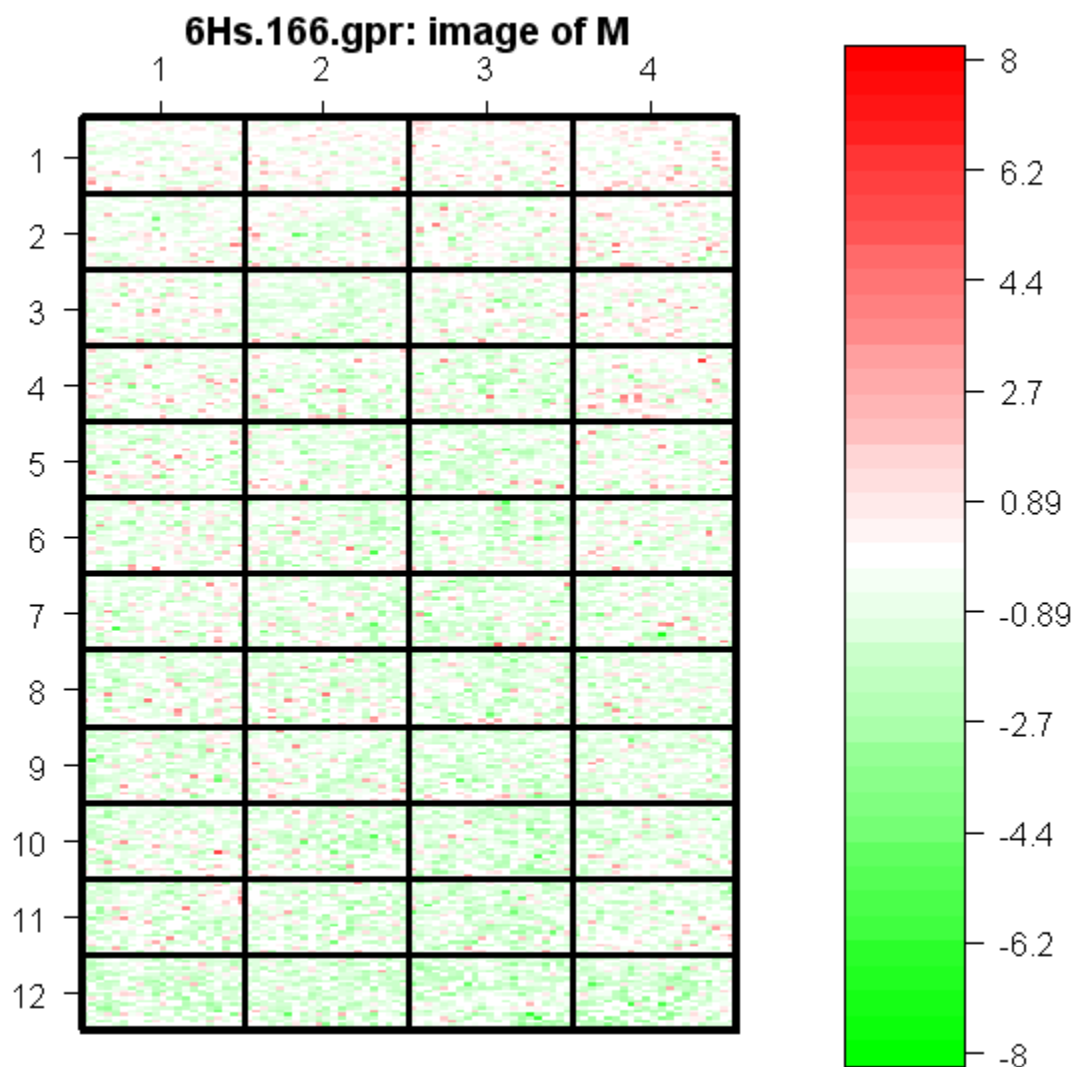
5.8.1 Modificación de la escala de colores

La función **maPalette** permite modificar la escala de colores

```
> colorG<- maPalette (low="white", high="green", k=50)
colorG: produce una escala del blanco al verde, , k=50 valor por defecto para la cantidad colores
```

```
> colorVR<- maPalette (low="green", mid="white", high="red")
colorVR: produce una escala de verde al rojo
```

```
> image (beta7[,3],xvar="maM",col=colorVR)
```



5.8.2 Gráficos de parámetros geométricos - de la construcción del arreglo - layout

5.8.2.1 Niveles de los parámetros

La función **image** puede utilizarse para graficar otros estadísticos por spot, además de las intensidades. Puede utilizarse para graficar los parámetros geométricos (layout) del arreglo:

maSpotRow, **maspotCol**, **maPrintTip**, **maPlate**.

```
> summary(factor(maPrintTip(beta7@maLayout) ))
  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16
483 483 483 483 483 483 483 483 483 483 483 483 483 483 483 483

 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
483 483 483 483 483 483 483 483 483 483 483 483 483 483 483 483

 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48
483 483 483 483 483 483 483 483 483 483 483 483 483 483 483 483

> 23*21
[1] 483
```

¿A qué corresponden los niveles? ¿Qué significa 483?

¿Cómo se interpreta el resumen siguiente?

```
> summary(beta7@maLayout@maPlate)
  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16
384 384 384 384 384 384 384 384 384 384 384 384 384 384 384 384

 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
384 384 384 384 384 384 384 384 384 384 384 384 384 384 384 384

 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48
384 384 384 384 384 384 384 384 384 384 384 384 384 384 384 384

 49 50 51 52 53 54 55 56 57 58 59 60 61
384 384 384 384 384 384 384 384 384 384 384 384 144
```

```
> summary(beta7@maLayout@maControls)
  Buffer      Empty Negative Positive   probes
    3      1328      225      204    21424
```

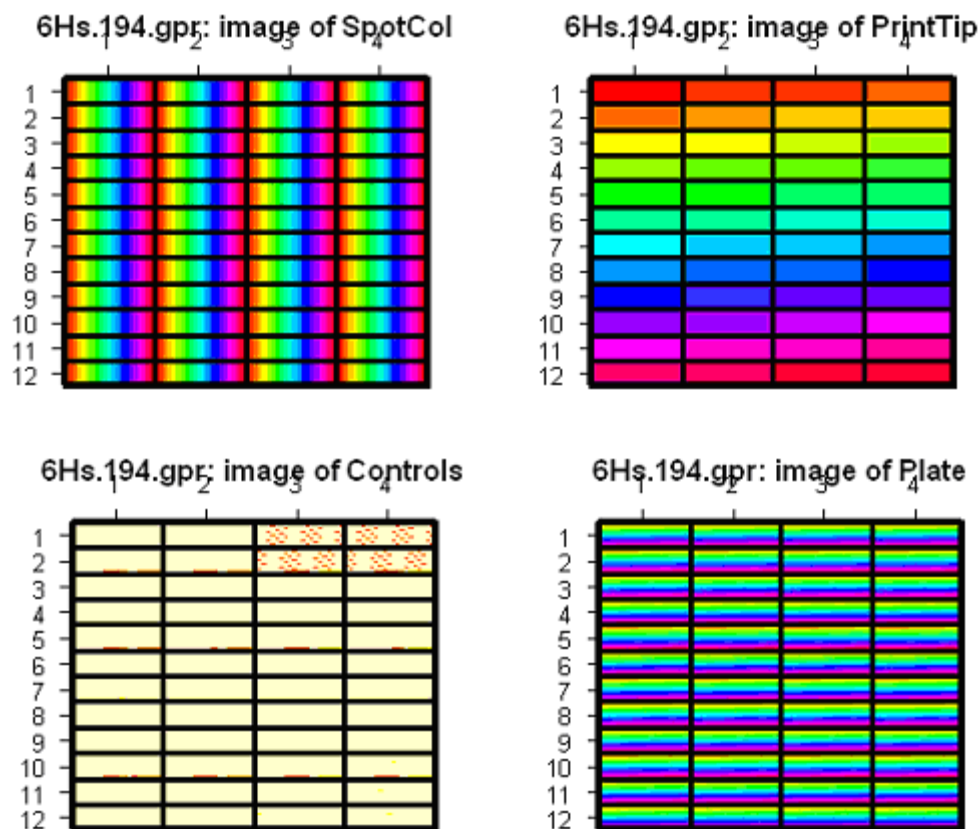
Verifique que utilizar la función **table**:

```
table(factor(maPrintTip(beta7@maLayout) ))  
table(beta7@maLayout@maPlate)  
table(beta7@maLayout@maControls)
```

produce las mismas salidas anteriores de **summary**

5.8.2.2 Obtención de los gráficos

```
par(mfrow=c(2,2))  
image (beta7[,5],xvar="maSpotCol", bar=FALSE)  
image (beta7[,5],xvar="maPrintTip", bar=FALSE)  
image (beta7[,5],xvar="maControls", col=heat.colors(10), bar=FALSE)  
image (beta7[,5],xvar="maPlate", bar=FALSE)
```



Si el objeto **marrayRaw** no contiene información en el slot **maPlate** debemos generarla utilizando la función **maCompPlate**. Esto ocurre en los datos de **swirl**:

```
> data(swirl)  
> maPlate(swirl) <- maCompPlate(swirl, n = 384)
```

n=384 indica la cantidad de cavidades (**wells**) que tiene el plato, generalmente es 384 ó 96.

para más información sobre los gráficos

> library (Biobase)

> openVignette()

13: marray - marrayPlots Overview