

Some useful graphic tools in R

Renée X. de Menezes

R Users Group - 7th February 2006

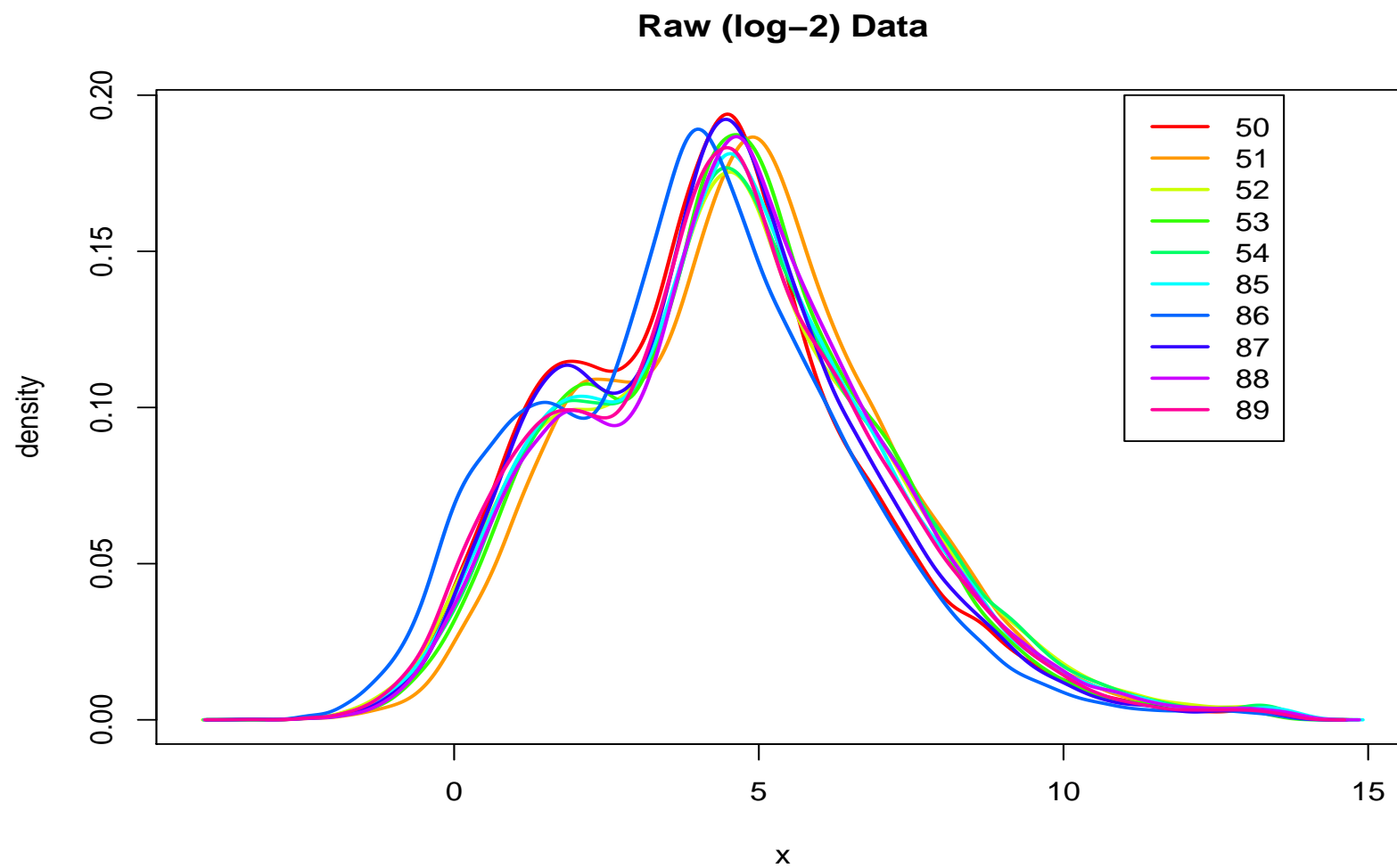
Example 1: Legend

- Legends can be included in all sorts of graphs
- Here we wish to include them in a densities graph of normalized chips

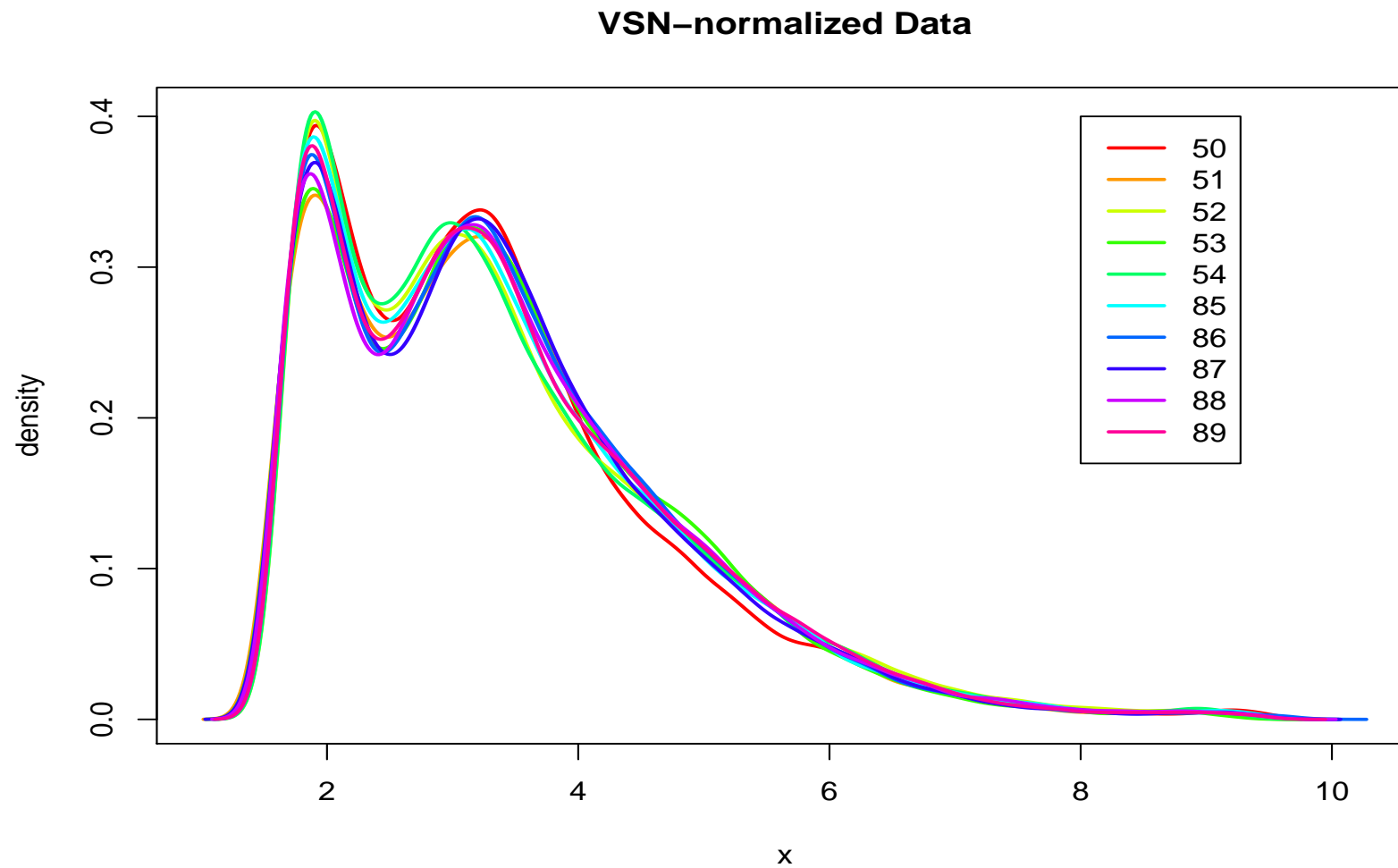
Example 1: Legend (continued)

```
all.data <- read.table("small_set_good.txt", sep="\t")  
  
library(vsn)  
  
vsndata <- vsn(all.data)  
  
vsndata <- exprs(vsndata)  
  
plotDensity(log2(all.data), col = rainbow(10), lwd = 2, lty  
= 1)  
  
legend(11, 0.2, legend = array.ids, lwd = 2, lty = 1,  
col =rainbow(10))
```

lograw 10goodarrays.pdf



vsn 10goodarrays.pdf



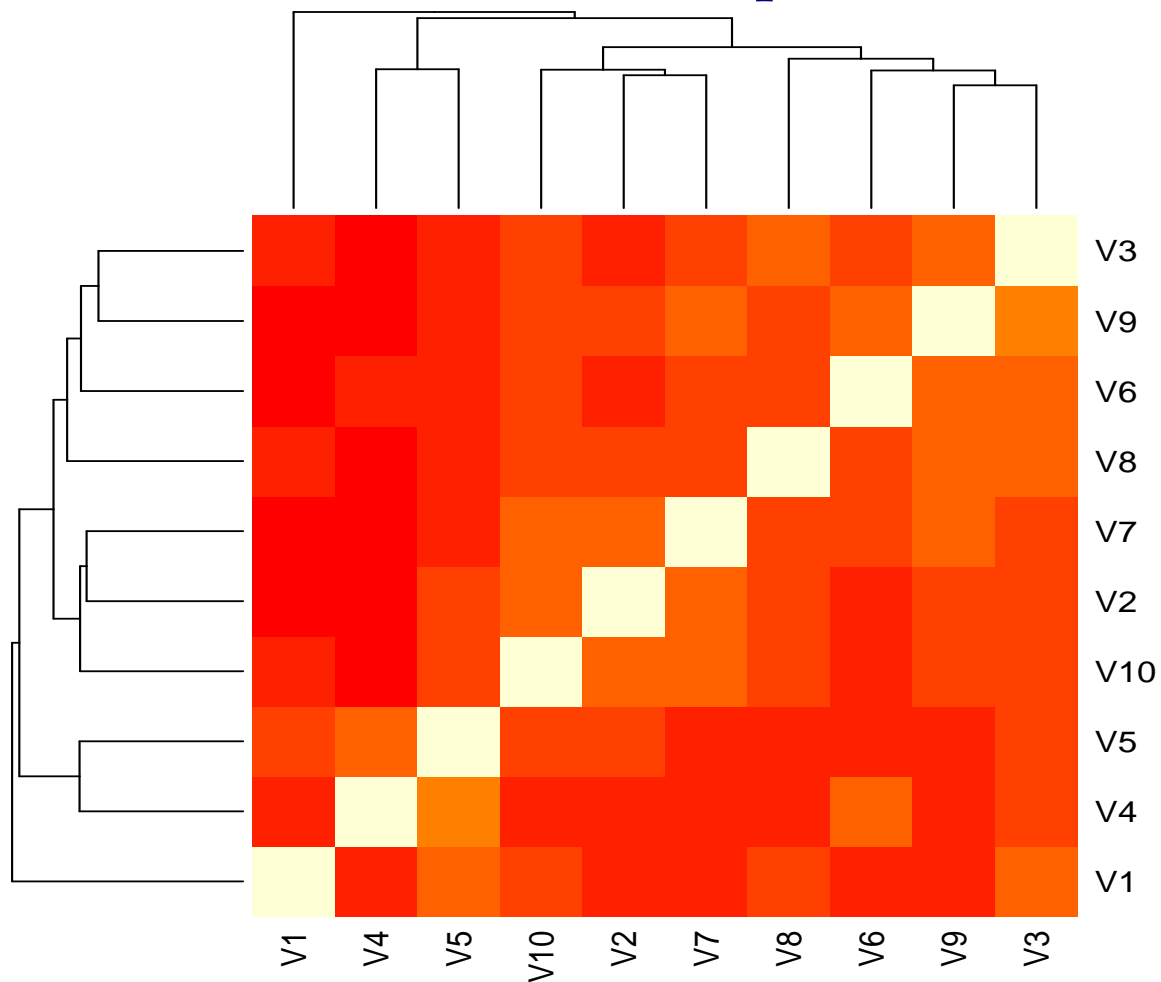
Example 2: Extras for heatmaps

- We can draw more informative heatmaps
- For example: correlation between chips
- Visualize to see if expected associations are present

```
cor.p <- cor(vsn.data,method="pearson")
```

```
heatmap(cor.p)
```

correlations.pdf



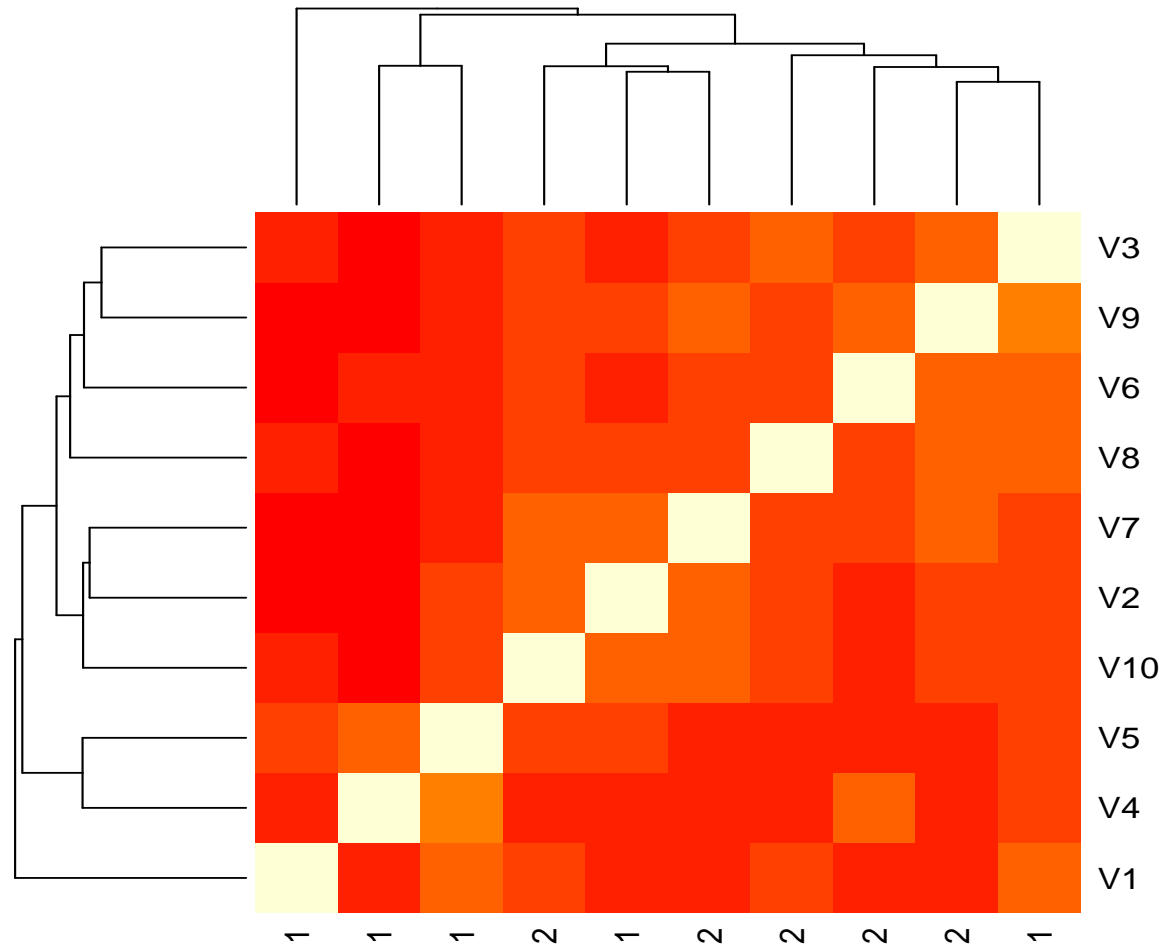
Example 2: Extras for heatmaps (continued)

replace column names by a variable

```
array.batch <- c(rep(1,5),rep(2,5))
```

```
colnames(cor.p) <- array.batch
```

```
heatmap(cor.p)
```

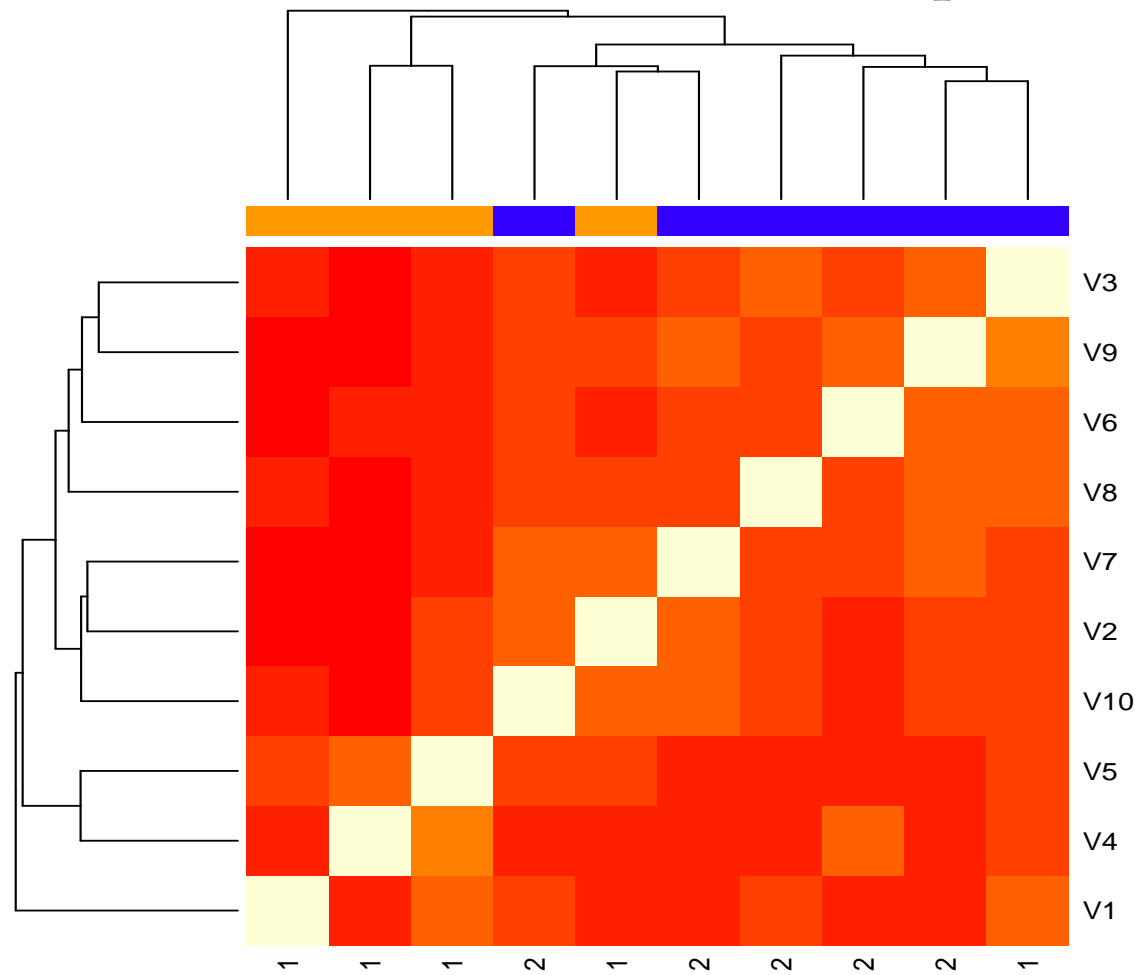



Example 2: Extras for heatmaps (continued)

add a colour bar to the graph

```
age.sample <- c(1,1,2,1,1,2,2,2,2,2)
f.age <- factor(age.sample)
vec.age <- rainbow(nlevels(f.age),start=.1,end=.7)
age.colour <- rep(0,length(f.age))
for(i in 1:length(f.age))
  age.colour[i] <- vec.age[ f.age[i]==levels(f.age) ]
heatmap(cor.p,ColSideColors = age.colour)
```

correlations colsidecolors.pdf

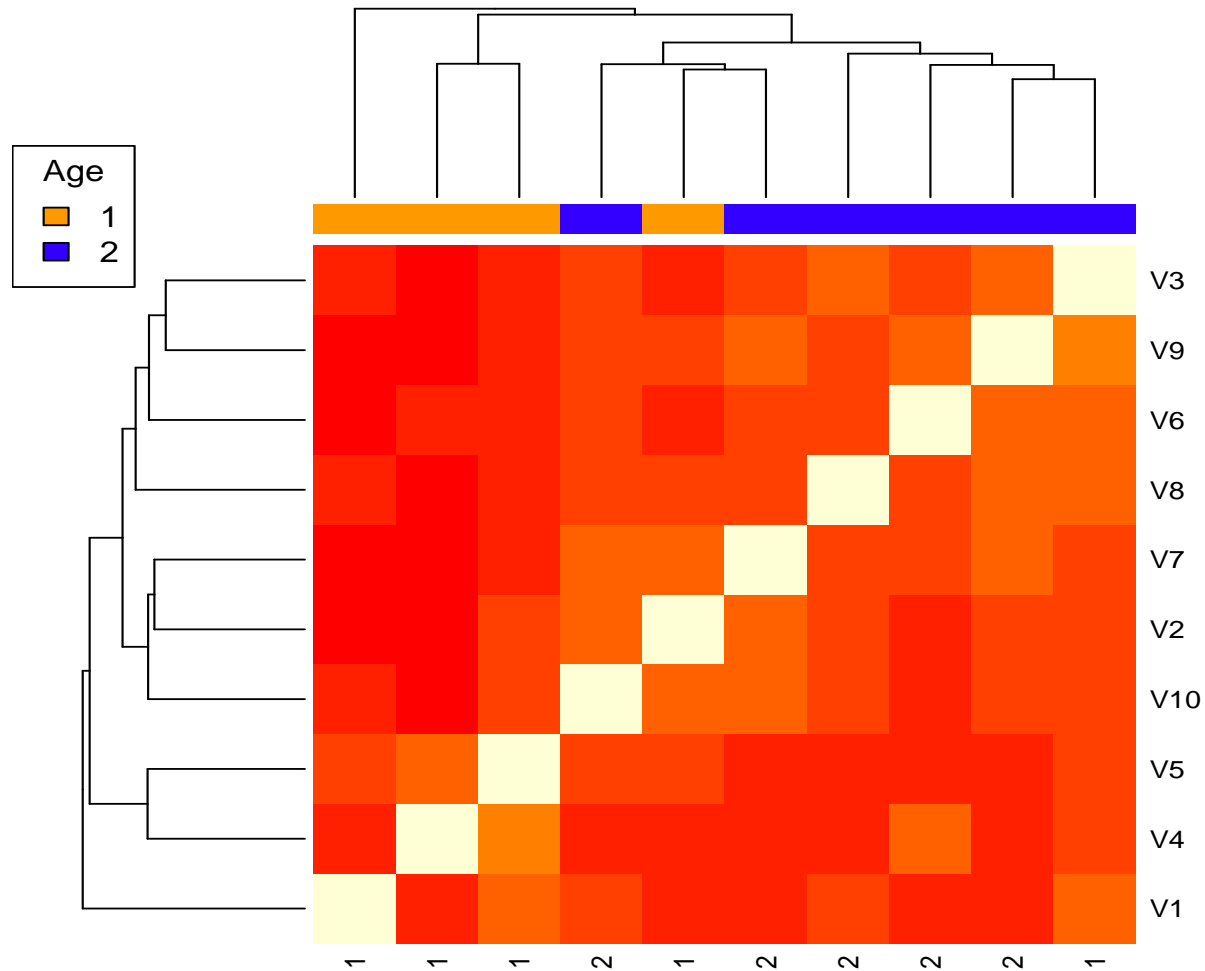


Example 2: Extras for heatmaps (continued)

add a legend explaining what the colours are

```
heatmap(cor.p, ColSideColors = age.colour)  
legend("topleft", legend=levels(f.age),  
fill=vec.age, title="Age")
```

correlations colsidecolors with legend topleft.pdf

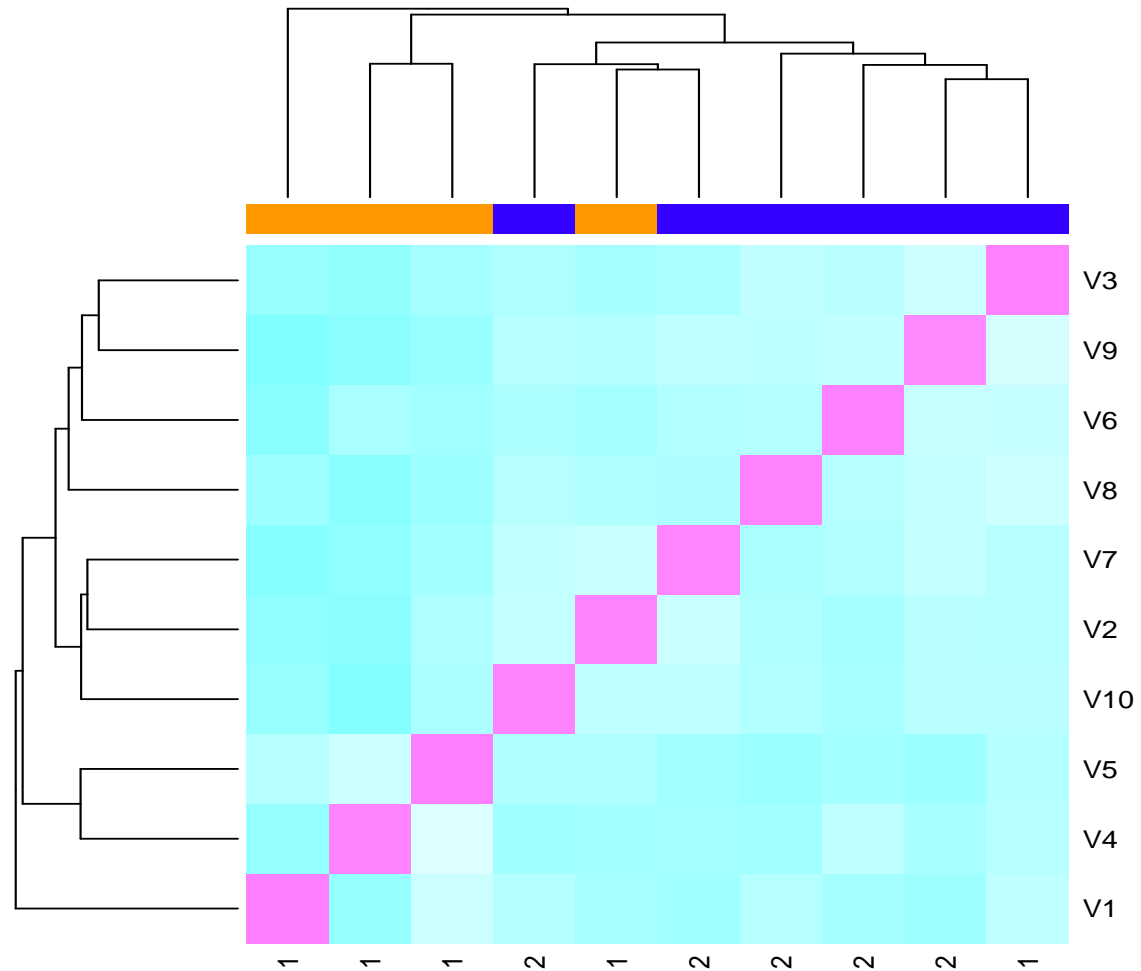


Example 3: Colour schemes for a heatmap

- default colours: heat colours
- many other options possible
- cyano-magenta colours

```
heatmap(cor.p, ColSideColors = age.colour, col=cm.colors(256))
```

correlations colsidecolors cm colors.pdf

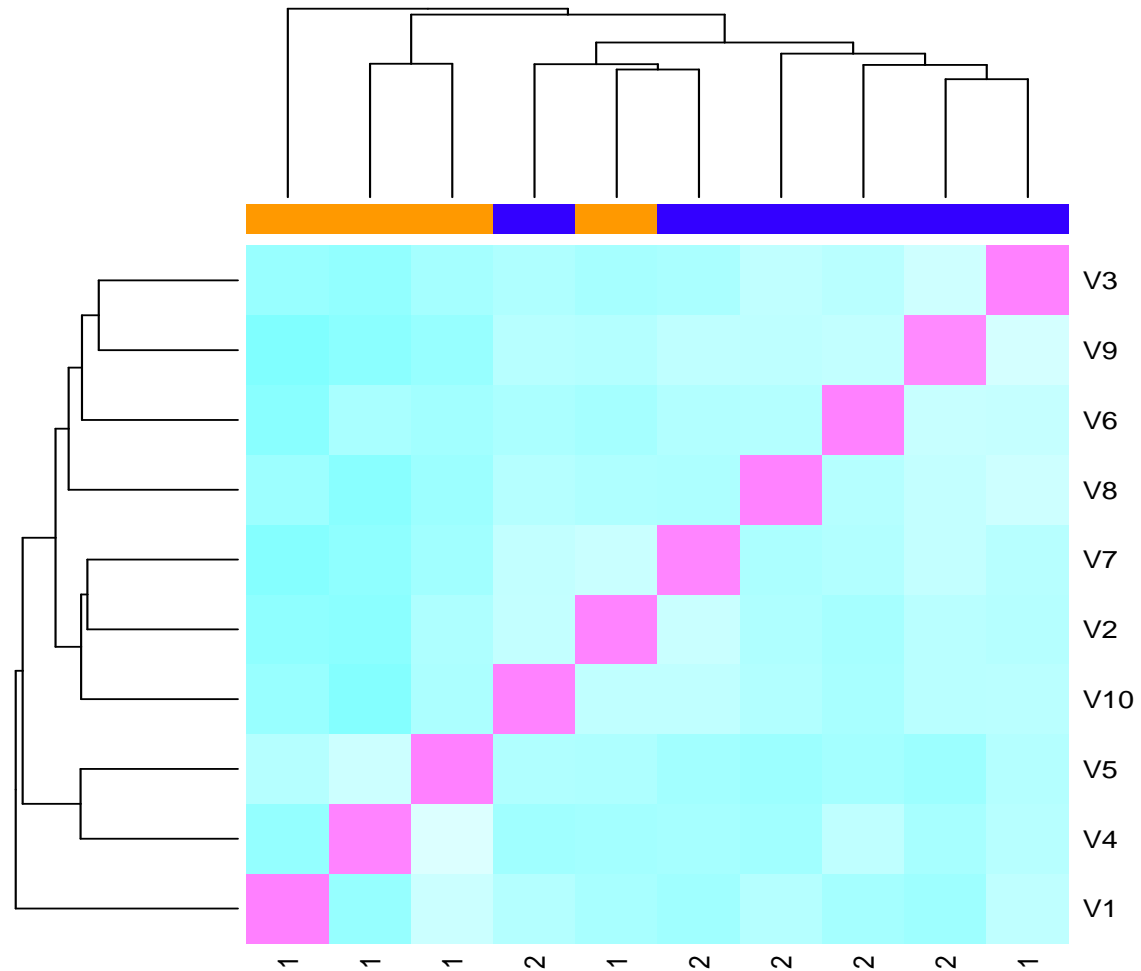


Example 3: Colour schemes for a heatmap (cont.)

topographical colours

```
heatmap(cor.p, ColSideColors = age.colour, col=topo.colors(256))
```


correlations colsidecolors cm colors.pdf



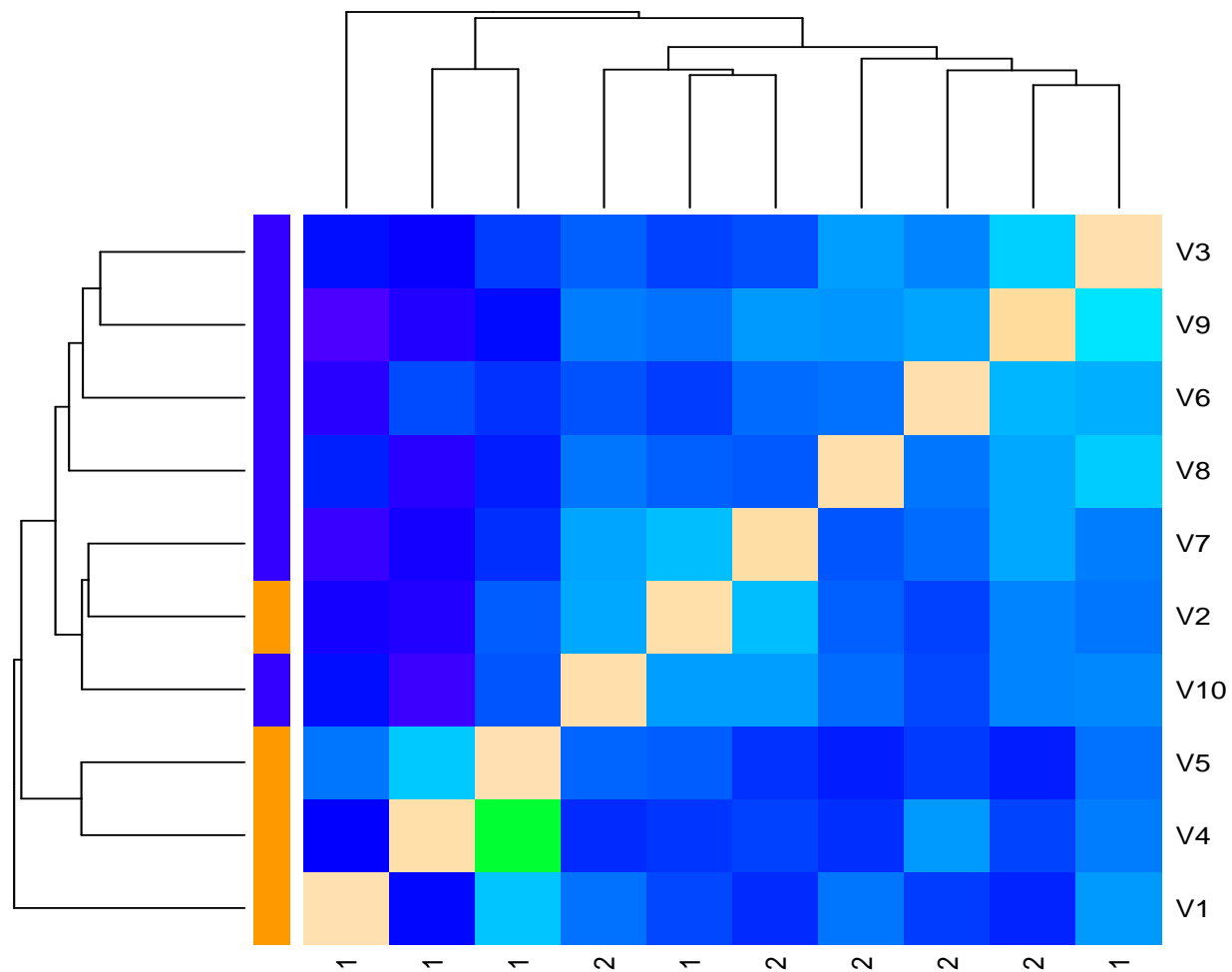
Example 3: Colour schemes for a heatmap (cont.)

- We can also put the colour bar on the vertical, and display it symmetrically
- However, there is a small bug: the rows get re-ordered, but the colour bar does not!
- Compare graphs without and with the symmetry command

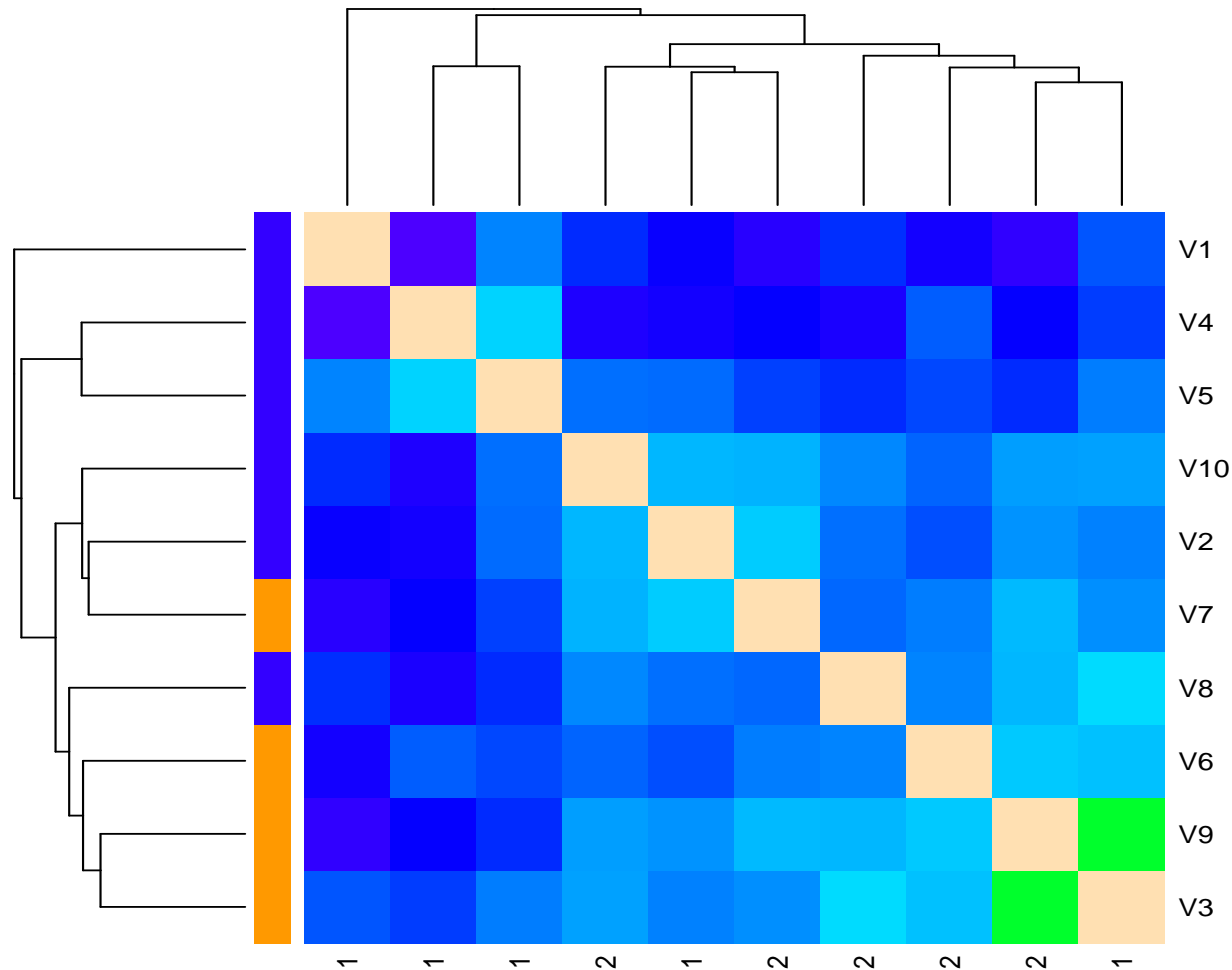
```
heatmap(cor.p, RowSideColors = age.colour, col=topo.colors(256))
```

```
heatmap(cor.p, symm=T, RowSideColors = age.colour, col=topo.colour
```

correlations colsidecolors cm colors rowsidebar.pdf



correlations colsidecolors cm colors rowsidebar symm.pdf



Example 4: Graphs for Wilcoxon/t test

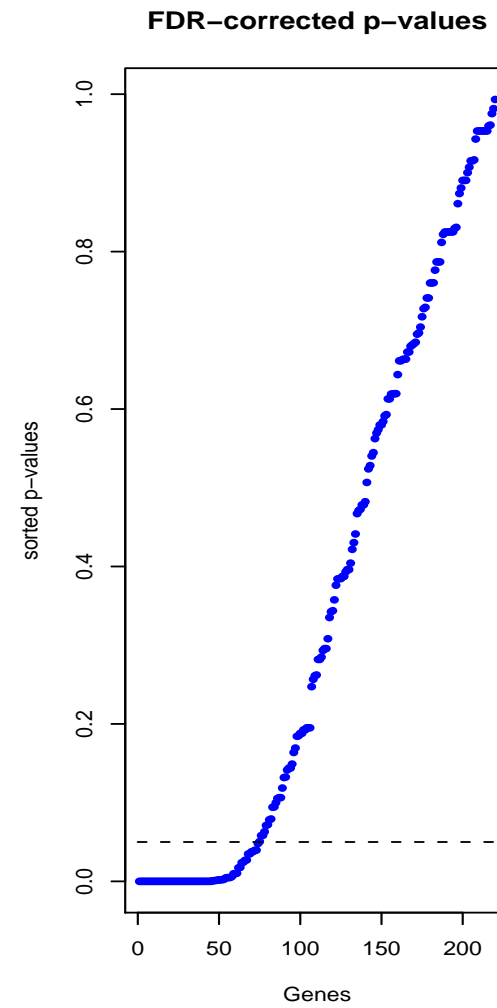
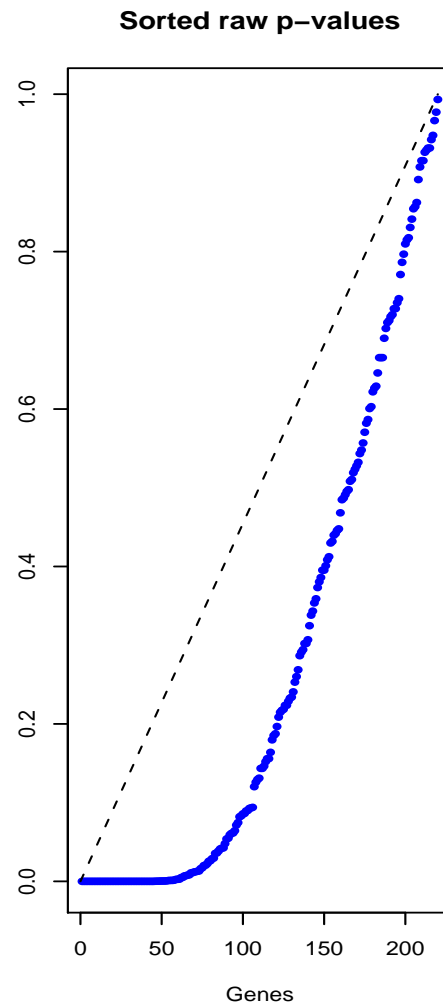
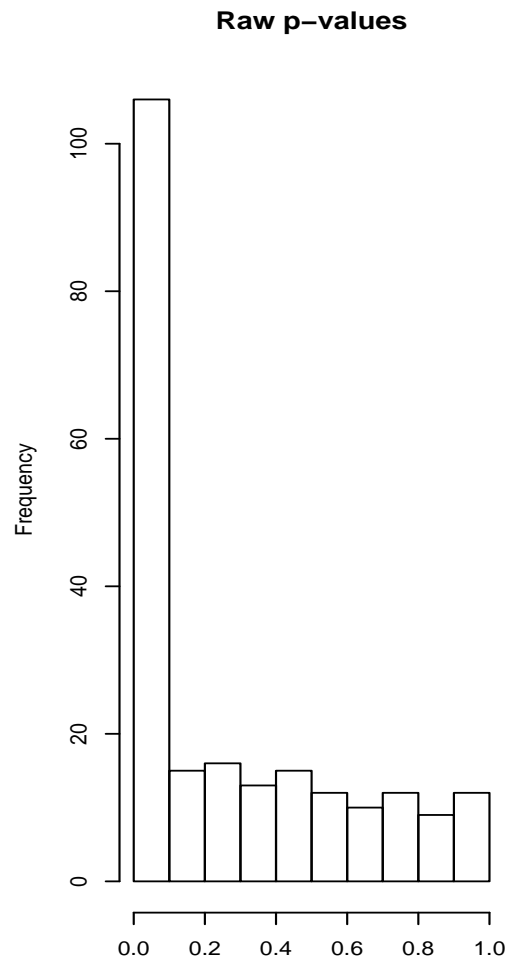
- Example: study involving 190 Affy chips
- Sample separated into two groups, factor "g"
- Selection: 220 probe sets for genes with same function
- Study: effect of group associated with expression levels

Example 4: Graphs for Wilcoxon/t test (cont.)

- Two analyses
- 1. compare expression levels in both groups with Wilcoxon test, per probe set
- 2. apply globaltest to test if expression profiles are associated with group

Example 4: Graphs for Wilcoxon/t test (cont.)

- Problem: studying the effect of a factor/variable
- Useful to have overview of the effect over all genes
- Graphs of raw p-values give a nice summary
- Histogram is flat, if no effect
- qq-plot nice in the middle
- I have included these graphs into a function



Example 5: Barplot for globaltest zscore

- Globaltest is applied to all probe sets simultaneously
- z-scores are saved
- probe sets with largest z-scores are selected
- corresponding gene names are extracted
- barplot of selected probe sets with gene names
- bar colour corresponds to gene

Example 5: Barplot for globaltest zscore (cont.)

```
gt.g <- globaltest(data.subset,f.group)
my.gt <- geneplot(gt.g)
zscore.g <- z.score(my.gt)
id.g <- zscore.g >=10 sum(id.g)
f.var <- factor(gene.names[id.g==T])
```

Example 5: Barplot for globaltest zscore (cont.)

```
col.vec <- rainbow(length(levels(f.var)))  
bar.col <- rep(0,length(f.var))  
for(i in 1:length(f.var))  
bar.col[i] <- col.vec[ f.var[i]==levels(f.var) ]  
zsco.g <- data.matrix(zscore.g[id.g==T])  
rownames(zsco.g) <- gene.names[id.g==T]  
barplot(sort(zsco.g[,1]),las=2,col=bar.col[order(zsco.g[,1])])  
title("Genes selected in global test")
```

Genes with most standardized influence in global testg

