



2017年度  
第5回バイオインフォマティクス実習

作図方法

先端医科学研究センター  
バイオインフォマティクス解析室 中林潤

- 大容量、高次元データを可視化
- 散布図、M-A plot、volcano plot、heatmap
- Rの作図機能を使用

# ファイルの取得

- 遺伝子発現プロファイル

GEO データベース accession number GSE10856

骨髄由来マクロファージをIgGまたはDcr3で処理

正規化した54675 x 4行列の遺伝子発現データから

5000遺伝子をランダムにサンプリングしたものを使用

課題配布フォルダのGSE10856\_sample.txtファイルを

各自のデスクトップへドラッグ アンド ドロップでコピー

# Rの設定

```
R _ □ X  
> Sys.setenv(http_proxy = "http://poxy.med.yokohama-cu.ac.jp:8080") ↵  
> source("http://bioconductor.org/biocLite.R") ↵  
> biocLite("gplots") ↵  
> library(gplots) ↵  
> biocLite("samr") ↵  
> library(samr) ↵
```

54675 x 4行列の遺伝子発現データ  
5000遺伝子をランダムにサンプリングしたものを使用  
課題配布フォルダのGSE10856\_sample.txtファイルを各自の  
デスクトップへドラッグ アンド ドロップでコピー

# scatter plot (散布図)

R



```
> plot(x[,5], x[,6], pch = 20, col = "gray", xlim = c(1,15), ylim = c(1,15), xlab = "lgG",  
+ ylab = "DcR3", main = "scatter plot")
```

## 散布図

データの分布を表示

Rでの行列の扱い

x[列,行]

plot(x軸, y軸, オプション)

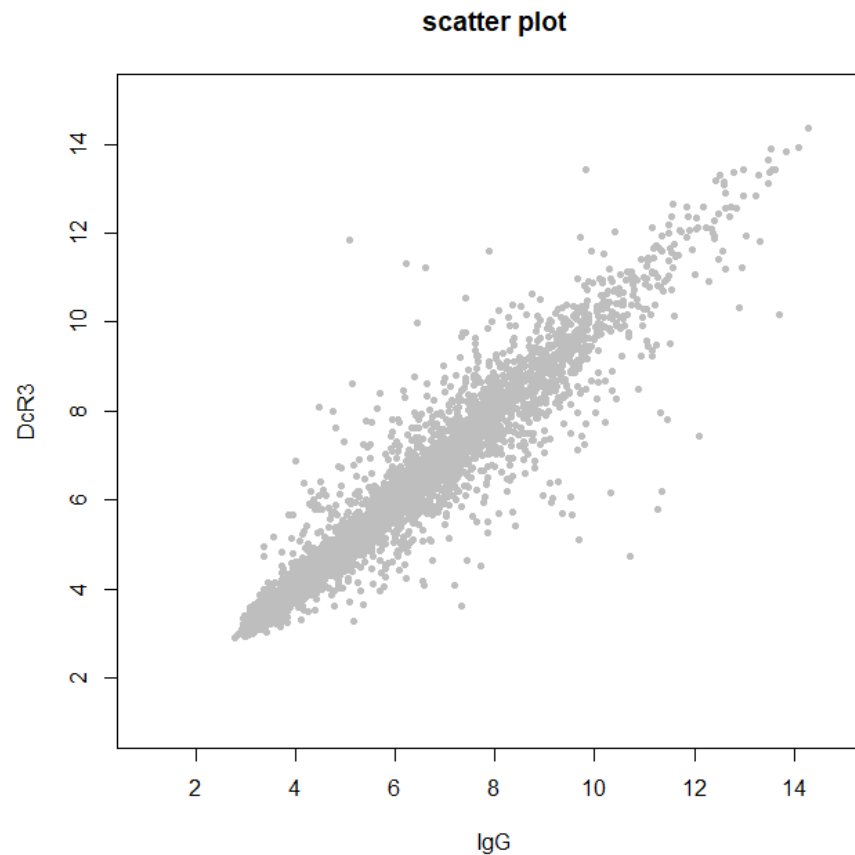
pch: 点のタイプ

col: 色

xlim, ylim: x軸、y軸の範囲

xlab, ylab: x軸、y軸のラベル

main: タイトル



# M-A plot

R



```
> plot((x[,5] + x[,6])/2, x[,5] - x[,6]), pch = 20, col = "gray", xlab = "A",  
+ ylab = "M", main = "M-A plot")
```

## 2群間比較

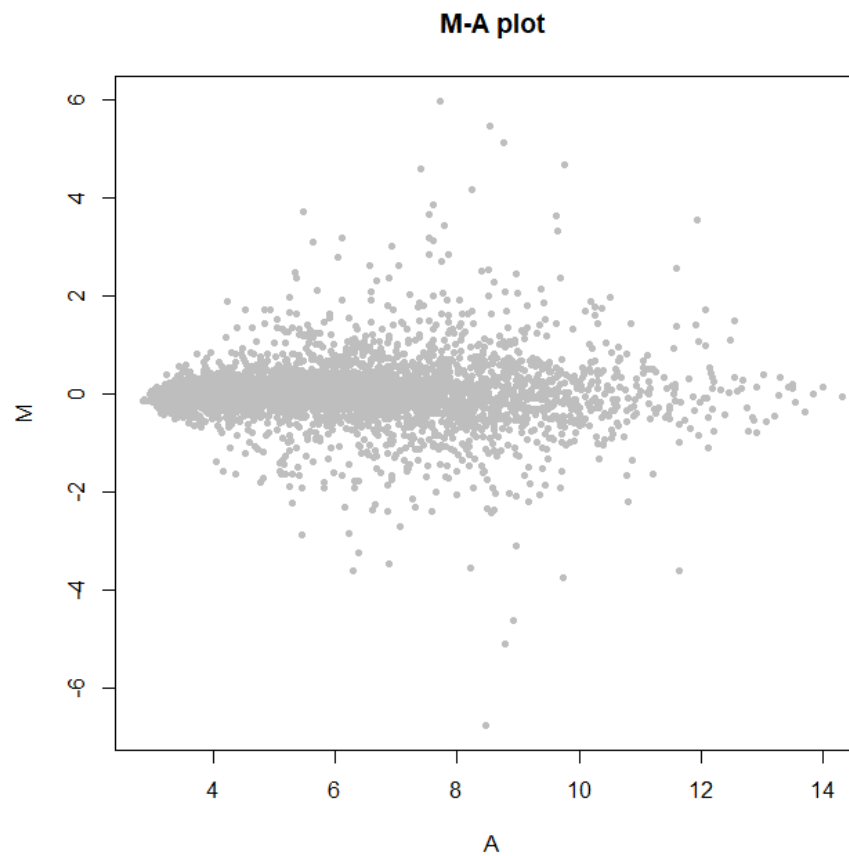
x軸: 平均発現量

$\log_2$  変換した値を使う

$(\log_2(A) + \log_2(B)) / 2$

y軸: 発現比

$\log_2(A/B)$



# volcano plot

R

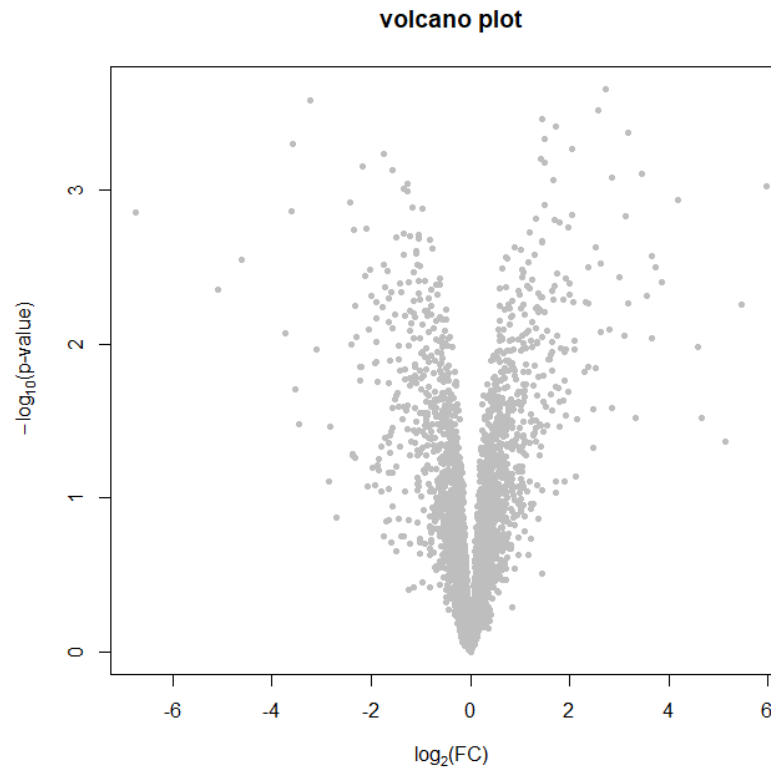


```
> data.tmp <- list(x = x[,1:4], y = factor(c("C", "T", "C", "T")), logged2 = TRUE) ↵  
> out <- samr(data.tmp, resp.type = "Two class unpaired", nperms = 20) ↵  
> p.value <- samr.pvalues.from.perms(out$tt, out$ttstar) ↵  
> plot(x[,5] - x[,6], -log10(p.value), xlab = expression(paste(log[2], "(FC)")), ↵  
+ ylab = expression(paste(log[10], "(p-value)")), col = "gray", main = "volcano plot") ↵
```

## 2群間比較

x軸: 発現比

y軸: p値



# heatmap

R



```
> heatmap.2(x[,1:4], col = greenred(75), cexRow = 0.1, cexCol = 0.8,  
+ trace = "none", density.info = "none", main = "heatmap")
```

散布図

データの分布を表示

Rでの行列の扱い

x[列,行]

gplots libraryのheatmap.2関数を使用

cexRow, cexCol:x軸、y軸のラベル文字の大きさ

col:greenred(階調)

