

クラスター解析デモ

FPKM (fragments per kilobase of exon per million reads mapped)
発現量の指標値、RNA-Seqのリードカウント数に対して、総リード数と転写産物長によっ
て補正した値

階層的クラスタリング

```
> data <- read.table("XXID_RNA_Seq_model_data.txt", header = TRUE)
> matrix_data <- as.matrix(data)
> head(matrix_data)
```

	Disruptant_induction	Disruptant	Wild_Type_induction
Wild_Type			
XXID_07421	0.00000	0	0
0			
XXID_07420	0.00000	0	0
0			
XXID_11561	0.00000	0	0
0			
XXID_07419	0.00000	0	0
0			
XXID_07418	0.13609	0	0
0			
XXID_11560	0.00000	0	0
0			

```
> heatmap(matrix_data, main = "HeatMap", Rowv = NA, Colv =
TRUE ,distfun = dist, hclustfun = hclust, col = heat.colors(256),
cexCol = 0.5)
```

```
cexRow = 0.5,
> heatmap(
```

非階層的クラスタリング

```
> data.km<-kmeans(data, 19, iter.max = 20) # クラスタリング
> Clstr_data <- data.km$cluster #クラスタリング結果の抽出
> Clstr_data #クラスタリング結果の表示
> data.clstr <- cbind(data, cluster = data.km$cluster) # データフレー
ムにマージ
> data.clstr_2 <- subset(data.clstr, cluster == 2,
c(Disruptant_induction,Disruptant,Wild_Type_induction,Wild_Type))
> write.table(data.clstr_2, "data.clstr_2.txt") #クラスタリング2結果の
出力
> clstr_2_axis <- 1:ncol(data.clstr_2) # data.clstr_2の列
> clstr_2_axis
> plot(0, 0, type = "l", xlim = range(clstr_2_axis), xaxp=c(1, 4,
3), ylim = range(data.clstr_2), xlab = "Condition", ylab = "FPKM")
> for (i in 1:nrow(data.clstr_2)) {lines(clstr_2_axis,
data.clstr_2[i,])}
```

```
> condition <- c(1, 2, 3, 4) # Median準備用X軸
> median <- c(median(data.clstr_2[, 1]), median(data.clstr_2[, 2]),
median(data.clstr_2[, 3]), median(data.clstr_2[, 4])) # Median計算
> median # Median表示
[1] 111.747 114.688 113.979 114.747
> lines(condition, median, type = "l", col="Red", xlim =
range(clstr_2_axis), xaxp=c(1, 4, 3), ylim = range(data.clstr_2),
xlab = "Condition", ylab = "FPKM")
```