

# マイクロアレイデータ解析（例）

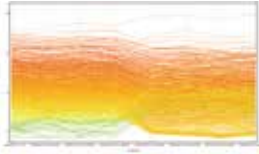
## ○データインポート

### ■Normalization

- 75 percentile shift 【デフォルト】  
各サンプルの 75% 番目の遺伝子の値を 0 に補正します。

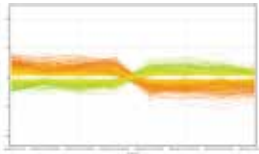
### ■Baseline Transformation

- ベースラインの補正をしない



ベースラインの補正をしないので発現量の情報は保持されていますが、視覚的にパターンを見つけにくくなります。

- Baseline to median of all samples



すべてのサンプルの中央値でそれぞれのサンプルの値を割ります。発現量の情報は失われますが、サンプル間の発現比は保持され、パターンを見つけやすくなります。クラスタリングをする際には基本的にこちらのベースライン補正を選択します。

- Baseline to median of control samples



control sample の値でそれぞれのサンプルの値を割るので、control sample との比較パターンを見つけやすくなります。ただし、Control sample の発現量の情報は失われるため、クラスタリングによるヒートマップ作成の際には適しません。

## ○有意差検定

### <基本納品データ>

- All Entities  
アレイに搭載されている Entity すべてのデータとなります。
- Reliable  
①のリストからコントロールスポットや発現の無い Entity を除去したデータとなります。
- T-Test or ANOVA=有意差検定リスト  
②のリストから T-test または ANOVA 等の有意差検定を行ったデータとなります。
- 2倍以上変動した Entity=Fold Change  
③のリストから 2倍以上変動した Entity のデータとなります。(2FoldUp と 2FoldDown)

有意差検定 (Statistical Analysis) 条件は、下記からご選択いただけます。

### ● Unpaired test

- 例：異なる Cell line 間 or 個体間で比較する場合
- T-test unpaired : 比較するグループが 2 群の場合
  - ANOVA : 比較するグループが 3 群以上の場合

### ● Paired test

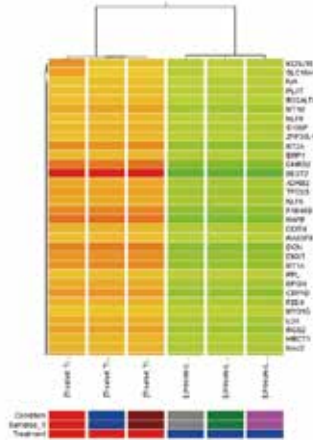
- 例：同一の Cell line 間 or 個体間で比較する場合
- T-test paired : 比較するグループが 2 群の場合
  - Repeated measures : 比較するグループが 3 群以上の場合

ANOVA は複数の条件を同時に比較することができ、どこか 1 条件でも変動していると発現差ありとみなされます。  
ある特定の比較ポイント (例えば Group1vs Group2, Group1vs Group3 など) で有意差がある遺伝子を抽出することも可能です (Post Hoc test)。

## ○クラスター解析

### <基本納品データ>

クラスター解析結果 (html 形式のファイル) とそれを開くためのファイル



左記のような Hierarchical データとなります。クラスター解析を行う際に使用するデータリストはデフォルトでは FoldChange のリストを用いますが、他のリスト (AllEntities, 有意差検定リストなど) でのご要望がございましたらご連絡ください。

## ○パスウェイ解析

### <基本納品データ>

パスウェイ解析結果 (html, Excel 形式のファイル) もしくはパスウェイに載っている遺伝子で解析対象遺伝子の発現値データ (Excel 形式)



アレイに搭載されている Entity はオレンジ色に表示され、解析対象となるデータリストにある Entity はさらにパッシングされて表示しています。発現変動につきまして Entity の横に棒グラフで示しております。ご注文いただく際には、解析対象となるデータリストをご指定ください。ご興味のあるパスウェイを解析される場合はそのパスウェイ (もしくはキーワード) をご指定ください。

## ○GO 解析

### <基本納品データ>

biological\_process のフォルダ内のエクセルファイル  
cellular\_component のフォルダ内のエクセルファイル  
molecular\_function のフォルダ内のエクセルファイル

GO (Gene Ontology project : <http://www.geneontology.org/>) は生物学的な語彙とそれに関連する遺伝子を収集し、遺伝子の機能や、局在、プロセスごとにカテゴライズされています  
GeneSpring の GO Analysis は、あるデータリストが GO で分類された遺伝子のリストと偶然一致する確率を計算し集計します。偶然一致する確率が低い (p-value が低い) ⇨ 関連が高いと解釈します。  
ご注文いただく際には、解析対象となるデータリストをご指定ください。もしくは、ご興味のある語彙だけを検索する場合はその語彙をご指定ください。