
AtCASTデータベース： 遺伝子発現プロファイルの相関から 植物の応答を推定するツール 使い方紹介 (4.0対応版)



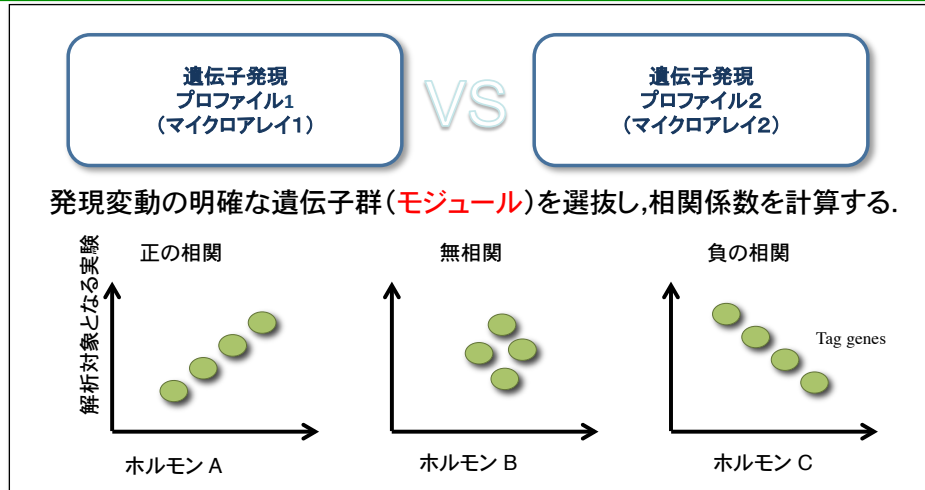
横浜市立大学 木原生物学研究所
筧雄介、嶋田幸久

目次

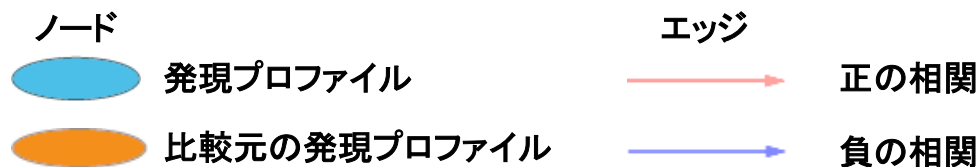
AtCASTの解析の仕組み	1 page
利用方法	
1. 公開データの発現プロファイルを検索	3 page
2. 準備したマイクロアレイデータを入力する	9 page
3. 準備したRNA-seqデータを入力する	12 page
備考	14 page

AtCASTによる解析の仕組み

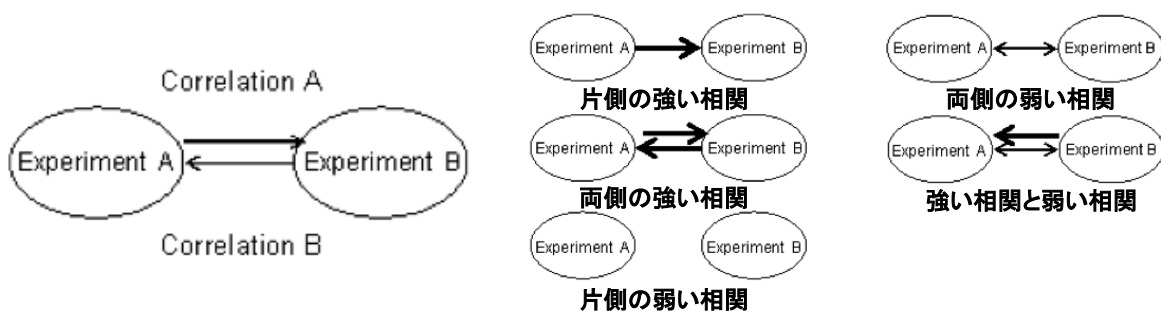
マイクロアレイ実験間の遺伝子発現相関解析



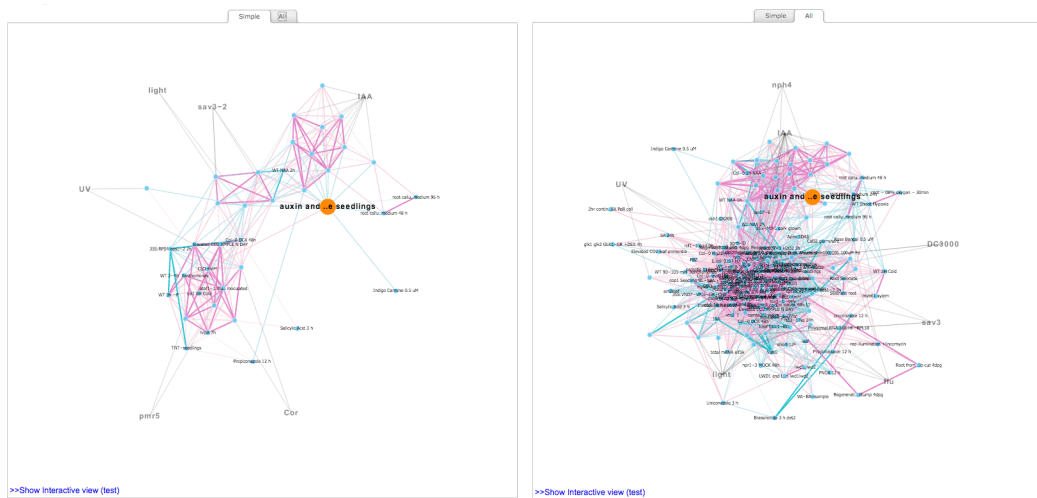
解析結果の表示



つながり方の例



モジュール相関ネットワーク例



同じ仕組みで植物が応答している実験が予想できます。

利用方法

- 使い方1
公開データの発現プロファイルを検索
- 使い方2
ユーザーが準備した遺伝子発現データを入力する

AtCAST 4.0: Home

Home Method Manual History

Welcome to AtCAST

[日本語マニュアル](#)

AtCAST is a web-based tool for analysis of transcriptome data and comparison of profiles with publicly available transcriptome data. AtCAST facilitates discovery of stimuli that induce similar gene expression responses and, therefore, are involved in similar biological processes. For example, one can search for experimental results demonstrating transcriptional changes that are similar or the opposite of those of a specific mutant in public transcriptome data.

Gene Set enrichment search ?

AGI code(s)

AtCAST will extract AGI codes from this text box.
Example: Two genes Aux/IAA1 (AT4G14560) and Aux/IAA19 (AT1G19220) were regulated by auxin.

Log₂ signal ratio ≥ 1.0

q-value ≤ 0.05

Sort by Number of DEGs in the list

Search

Select genes in a GO term.

Search GO term by text.

Select genes in pathway.

Search pathway name by text.

使い方1① →

使い方1② →

Gene Set correlation analysis

>Analyze Public data >Analyze your data

使い方1①

興味のある遺伝子セットのAGIコード入力
> Search

興味のある遺伝子情報の
GO termを入力(選択)すると、
左欄にAGIコードが表示される

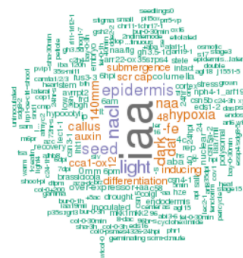
興味のある遺伝子セットの
pathwayを入力(選択)すると、
左欄にAGIコードが表示される

検索結果図

AtCAST 4.0: Search result

[Home](#) [Method](#) [Manual](#) [History](#)

Searching with AT4G14560 AT1G19220, Abs. Log2 Signal ratio >= 1.0, q-value <= 0.05
Please wait for a while. 0%>25%
.....>50%
.....>75%>100%
Number of hits = 221, Search time = 30.126449 sec
[Back](#)



検索結果から
関心のある実験名をクリック
ネットワーク図とリストのページへ

Number of genes in the list is 2 (AT4G14560, AT1G19220)

Experiment	p-value(Fisher)	Ave. signal ratio	Num. DEGs in list
IAA Treated	1.672e-06	2.471	2
Fei-0-1h IAA	4.709e-06	1.851	2
columbia auxin	4.069e-05	2.416	2

使い方1②

Find experimental condition by keyword ?

Example: Auxin, IAA, GSE10248

AND Search

検索したいマイクロアレイ実験のキーワード入力
 > 検索結果から関心のある実験名をクリック
 ネットワーク図とリストのページへ

検索結果図

検索結果から
 関心のある実験名をクリック
 ネットワーク図とリストのページへ

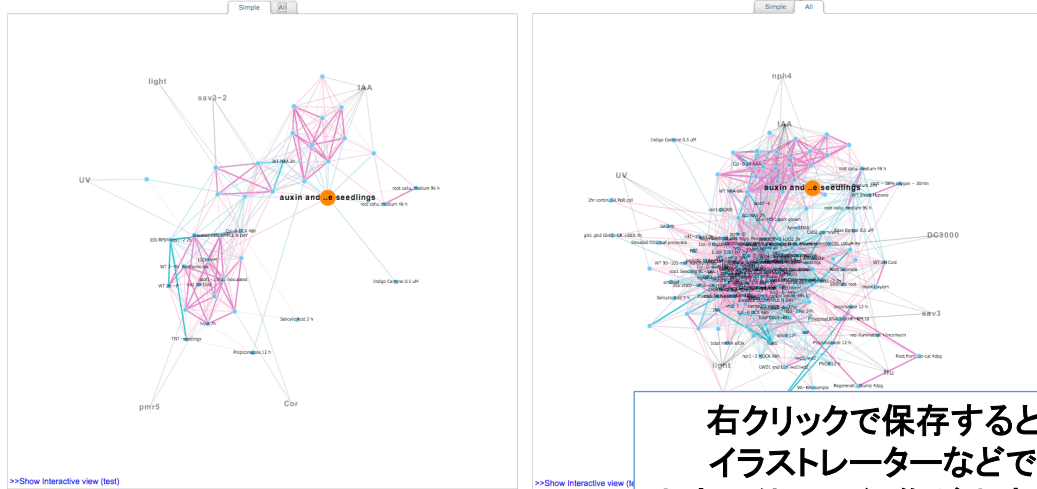
Back

Auxin

	Experiment	Genotype	Treatment / Tissue	Experimental Category
1	auxin and brassinazole seedlings		/	GEO bulk(microArray)
2	IAA/inhibitor B Treated		/	GEO bulk(microArray)
3	IAA/inhibitor A Treated		/	GEO bulk(microArray)
4	IAA Treated		/	GEO bulk(microArray)
5	inhibitor B Treated		/	GEO bulk(microArray)
6	inhibitor A Treated		/	GEO bulk(microArray)
7	miR393 flg22 2h		/ Leaf tissue	GEO bulk(microArray)
8	miR393 flg22 1h		/ Leaf tissue	GEO bulk(microArray)
9	miR393 0h		/ Leaf tissue	GEO bulk(microArray)
10	WT flg22 2h		/ Leaf tissue	GEO bulk(microArray)
11	WT flg22 1h		/ Leaf tissue	GEO bulk(microArray)
12	WT NAA/flg 2h		/ Leaf tissue	GEO bulk(microArray)

モジュール相関ネットワーク図

タブで Simple, All を
選択可能



右クリックで保存すると
イラストレーターなどで
文字や位置の編集が出来ます

相関の強さなどのリスト

List of experiments correlated with "auxin and brassinazole seedlings" ?

[Download transcriptome of top 30](#)

Score	Experiment	Correlation from auxin and brassinazole seedlings	Correlation to auxin and brassinazole seedlings	Control Experiment	Genotype	Treatment	Tissue	Link to original data	Experimental Category	Data set Details
1	auxin and brassinazole seedlings	1.00 >More info	1.00 >More info	auxin seedlings				Original data	GEO bulk	link
2	Bay-0-1h IAA	-0.52 >More info	-0.72 >More info	Bay-0-0h IAA			whole plant	Original data	GEO bulk	link
3	root inducing medium	-0.56 >More info	-0.65 >More info	root callus induction medium		48 hours after incubating on callus inducing medium	root explants	Original data	GEO bulk	link
4	L-AOPP+IAA	-0.50 >More info	-0.71 >More info	untreated		4 weeks old at time of treatment	whole plant	Original data	GEO bulk	link
5	sav3-2 Low R:FR light	0	0					Original data	GEO bulk	link
6	sav3-2 shade	0	0				seedlings	Original data	GEO bulk	link
7	IAA Treated	-0	-0					Original data	GEO bulk	link
8	CLO 4mM	0.43 >More info	0.74 >More info	untreated		4 weeks old at time of treatment	whole leaf	Original data	GEO bulk	link
9	root callus inducing medium 96 h	-0.52 >More info	-0.52 >More info	root callus		96 hours after		Original data	GEO bulk	link
10	Cor 24h	0.46 >More info	0.46 >More info					Original data	GEO bulk	link

他の実験の解析結果へのリンク

オリジナルの登録情報へのリンク
(主に外部サイト)

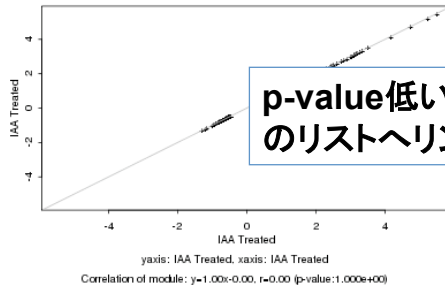
実験間反復などの統計情報

モジュールの情報へのリンク

IAA Treated

In comparison with IAA Treated [help](#)[Home](#) [Method](#) [Manual](#) [History](#) [close](#)

Scatter plot of selected modules.



p-value低い順100位までのリストへリンク

Trend of Up regulated genes including module

GO ID	Term	Ontology	p-value
GO:0009733	response to auxin	BP	1.506e-34
GO:0009733	response to hormone	BP	2.066e-20
GO:0009733	response to endogenous stimulus	BP	5.612e-20
...	...	...	...
View Hierarchical networks of GOs (BP MF CC)			

Trend of Down regulated genes including module

GO ID	Term	Ontology	p-value
GO:0030001	metabolic process	BP	1.506e-34
GO:0006661	phosphatidylinositol metabolic process	BP	2.066e-20
GO:0046488	phosphatidylinositol metabolic process	BP	5.612e-20
...	...	...	...
View Hierarchical networks of GOs (BP MF CC)			

GOE階層ネットワーク図へ

List of genes in module

...up regulated genes ...down regulated genes

比較元の実験における発現

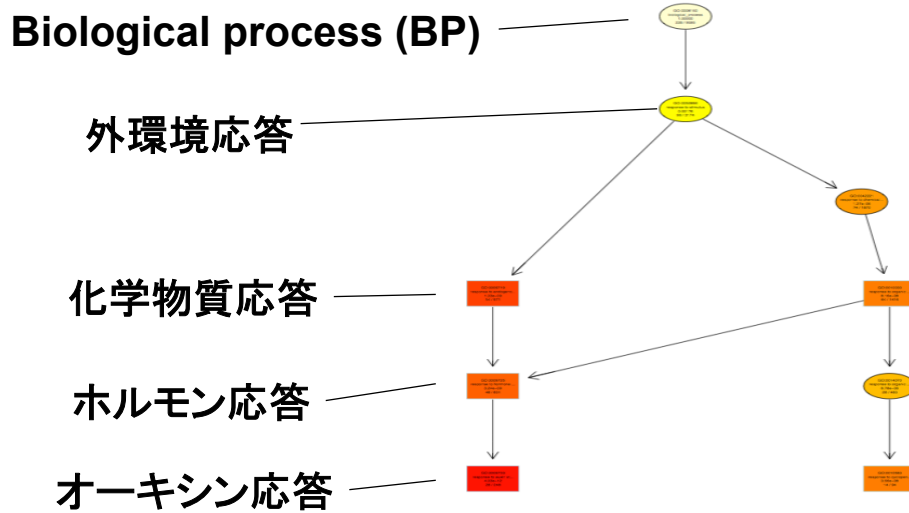
Downstream gene expressions
Downstream sequences

	probeID	AGICode	Annotation	Log ₂ signal ratio slr-1 2h NAA	Log ₂ signal ratio IAA Treated
1	260097_at	AT1G73220	AtOCT1, organic cation/carnitine transporter1, OCT1, organic cation/carnitine transporter1	3.715	-0.013
2	249046_at	AT5G44400	[FAD-binding Berberine family protein]	2.758	-0.219
3	264940_at	AT1G60470	AtGolS4, galactinol synthase 4, GolS4, galactinol synthase 4	2.431	0.076
4	252202_at	AT3G50300	[HXXXD-type acyl-transferase family protein]	2.261	-0.725
5	256788_at	AT3G13730	CYP90D1, cytochrome P450, family 90, subfamily D, polypeptide 1	2.244	-0.359
6	256626_at	AT3G20015	[Eukaryotic aspartyl protease family protein]	2.197	-0.664
7	251858_at	AT3G54820	PIP2:5, plasma membrane intrinsic protein 2:5, PIP2D, PLASMA MEMBRANE INTRINSIC PROTEIN 2D	2.147	-0.048
8	252414_at	AT3G47420	AtG3Pp1, Glycerol-3-phosphate permease 1, ATPS3, phosphate starvation-induced gene 3, G3Pp1, Glycerol-3-phosphate permease 1, PS3, phosphate starvation-induced gene 3	2.080	0.225
9	257540_at	AT3G21520	AtDMP1, Arabidopsis thaliana DUF679 domain membrane protein 1, DMP1, DUF679 domain membrane protein 1	2.018	-0.086
10	258241_at	AT3G27650	LBD25, LOB domain-containing protein 25	1.944	-0.435
11	250002_at	AT5G18690	AGP25, arabinogalactan protein 25, ATAGP25, ARABIDOPSIS THALIANA ARABINOGALACTAN PROTEINS 25	1.780	-0.506
12	253017_at	AT4G37970	ATCAD6, CAD6, cinnamyl alcohol dehydrogenase 6	1.730	-0.160
13	257690_at	AT3G12830	SAUR72, SMALL AUXIN UPREGULATED 72	1.716	-0.434
14	255160_at	AT4G07820	[CAP (Cysteine-rich secretory protein superfamily protein)]	1.619	-0.303
15	251317_at	AT3G61490	[Pectin lyase-like superfamily protein]	1.619	-0.303

比較先の実験における発現

GOE階層ネットワーク図

Ontology termには階層があり、一番下の階層の物が最も詳しい(注目すべき)



使い方2. ユーザーが準備した遺伝子発現データを入力する

 **AtCAST 4.0: Home**

Home Method Manual History

Welcome to AtCAST

[日本語マニュアル](#)

AtCAST is a web-based tool for analysis of transcriptome data and comparison of profiles with publicly available transcriptome data. AtCAST facilitates discovery of stimuli that induce similar gene expression responses and, therefore, are involved in similar biological processes. For example, one can search for experimental results demonstrating transcriptional changes that are similar or the opposite of those of a specific mutant in public transcriptome data.

Gene Set enrichment search ?

AGI code(s)

AtCAST will extract AGI codes from this text box.
Example: Two genes Aux/IAA1 (AT4G14560) and Aux/IAA19 (AT1G19220) were regulated by auxin.

Log₂ signal ratio ≥ 1.0

q-value ≤ 0.05

Sort by Number of DEGs in the list

Search

Select genes in a GO term.

Search GO term by text.

Select genes in pathway.

Search pathway name by text.

Gene Set correlation network analysis

>Analyze Public data

>Analyze your data

使い方2

AtCAST 4.0:Analyze

[Home](#) [Method](#) [Manual](#) [History](#)



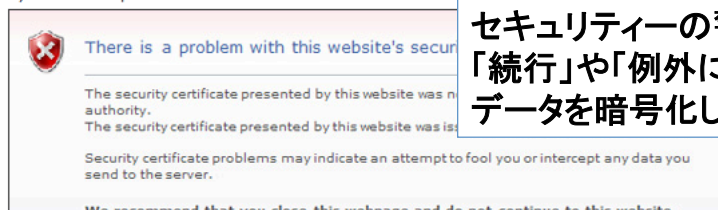
>Analysis of microarray data

>Analysis of RNA-seq data

まずはマイクロアレイデータの入力について

To upload user data to the AtCAST database in a secure way, the website encrypts this session with self-signed certificate. If you get warning message (as shown in the following screen shot), please confirm and click "continue to this website" button.

1) Internet Explorer 7 for Windows XP



データ入力を初めて行う時は図のようなセキュリティの警告が出ます。「続行」や「例外に追加」を選んでください。データを暗号化してやり取りするためです。

AtCAST 4.0:Data Upload

[Data Upload](#) > [Confirm the entry](#) > [Start the analysis](#)

It takes about an hour to finish all the calculations. URL to your result page will be sent to your E-mail address.

[See a sample of the data format](#)

E-mail

* Required fields

Files

*4 data files are required in total Control and Treatment.

Control ☐ use artificial control data (platform:ATH1)

Treatment

選択... ファイルが選択されていません。
選択... ファイルが選択されていません。
選択... ファイルが選択されていません。
選択... ファイルが選択されていません。
選択... ファイルが選択されていません。
選択... ファイルが選択されていません。

選択... ファイルが選択されていません。
選択... ファイルが選択されていません。
選択... ファイルが選択されていません。
選択... ファイルが選択されていません。
選択... ファイルが選択されていません。
選択... ファイルが選択されていません。

Platform

Genotype, Treatment, Tissue, Control Exp

Genotype

Treatment

Tissue

*印がついているところは必須項目です必ず入力してください。1アレイにつき1ファイル入力になります。Platformについては、MAS5で計算したデータ以外はAnalyze using AGI codeを選んでください。

入力データの準備

Header*	(1)probeID*	(2)Signal*	(3)Detection*	(4)Detection p-value*
	probeID	Signal	Detection	p-value
	AFFX-BioB-5_at	456.3	P	0.030976
	AFFX-BioB-M_at	707.7	P	0.004484
	AFFX-BioB-3_at	346.6	A	0.131361
	AFFX-BioC-5_at	1546.7	P	0.000052
	AFFX-BioC-3_at	699.9	P	0.00141
	AFFX-BioDn-5_at	1439.8	P	0.000195
	AFFX-BioDn-3_at	7844.7	P	0.000052
	AFFX-CreX-5_at	16712.1	P	0.000052
	AFFX-CreX-3_at	21738.6	P	0.000044

Affimetrixのアレイチップを使ってMAS5でシグナル値を計算している場合図のような4列の表を用意してください。それ以外の場合は左の2列をAGIコードとシグナル値になるように用意してください。Excelなどで作って“タブ区切りテキストまたは .xls, .xlsx“で保存してください。

Rを使ったデータの準備 (affyパッケージ)

```
setwd("where_your_cel_files_are") #Modify this line
library(affy)
data <- ReadAffy()
data.signals <- mas5(data)
data.mas5calls = mas5calls(data)
data.mas5pvals <- assayData(data.mas5calls)[["se.exprs"]]
data.file.list <- dir()
n = 0
a <- as.list(NULL)
for(data.file in data.file.list) {
  if(any(i <- grep("CEL|cel", data.file))) {
    n = n+1
    a[n] <- gsub("\\.CEL.gz|\\.cel.gz|\\.CEL|\\.cel", "\\..txt",
data.file)
    sa <- exprs(data.signals)[,n]
    sb <- exprs(data.mas5calls)[,n]
    sc <- data.mas5pvals[,n]
    sd <- cbind(sa, sb, sc)
  }
}
```

sep="\t"

同じフォルダに集めたCELファイルから全ての入力ファイルを一括で準備するためのRスクリプトも入力ページ「See a sample format」のリンク先にのせてあります。

AtCAST 4.0:Analyze

[Home](#) [Method](#) [Manual](#) [History](#)



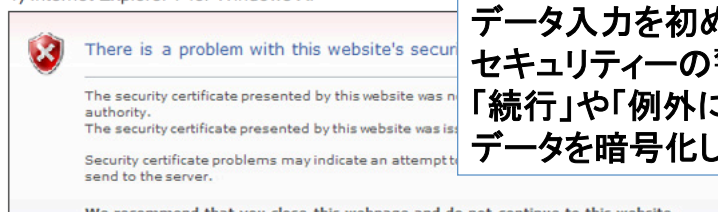
>Analysis of microarray data

>Analysis of RNA-seq data

RNA-seqデータの入力について

To upload user data to the AtCAST database in a secure way, the website encrypts this session with self-signed certificate. If you get warning message (as shown in the following screen shot), please confirm and click "continue to this website" button.

1) Internet Explorer 7 for Windows XP



データ入力を初めて行う時は図のようなセキュリティの警告が出ます。「続行」や「例外に追加」を選んでください。データを暗号化してやり取りするためです。

AtCAST 4.0:Data Upload

[Data Upload](#) > [Confirm](#) > [Start](#)
the entry the analysis

It takes about an hour to finish all the calculations. URL to your result page will be sent to your E-mail address.

[See a sample of the data format](#)

E-mail

* Required fields

Input File (Counts)

選択... ファイルが選択されていません。 *

ID Control column number(s)

1 -- -- -- *

Treatment column number(s)

-- -- -- -- *

Platform

RNA-seq *

Genotype, Treatment, Tissue, Control Experiment

Genotype

Treatment

*印がついているところに最低限入力してください
遺伝子ごとのリード数の表を1ファイル入力になります。
それぞれの実験区がどの列に相当するか選んでください。

入力データの準備

Header

(1)Gene locus (2)Gene expression counts

tracking_id	q1_count	q2_count	q3_count	q4_count
AT1G01010	797.318	779.892	660.844	563.1
AT1G01020	306.281	356.935	360.256	319.1
AT1G01030	55.3282	60.8646	99.0703	89.3
AT1G01040	1734.93	1679.45	1696.58	1648.1
AT1G01046	13.832	16.5057	13.5096	11.27
AT1G01050	1805.08	1770.23	2106.37	2160.1

左の1列をAGIコード、それ以外がリードカウントになるように用意してください。FPKMなどを発現値として入力しても解析できますが、精度が多少落ちます。Excelなどで作って“タブ区切りテキスト“で保存してください。

AT1G01090	4088.36	4193.47	4469.47	4893.1
-----------	---------	---------	---------	--------

備考: データの違いについて

- 計算方法の違い
(MAS5, RMA, GCRMA, FARMS, DFW)
- RNAseqとマイクロアレイデータの違い
(ダイナミックレンジ、検出感度)

ほぼ影響なく比較可能