

---

# 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 (3.01対応版)



横浜市立大学 木原生物学研究所  
筧雄介、嶋田幸久

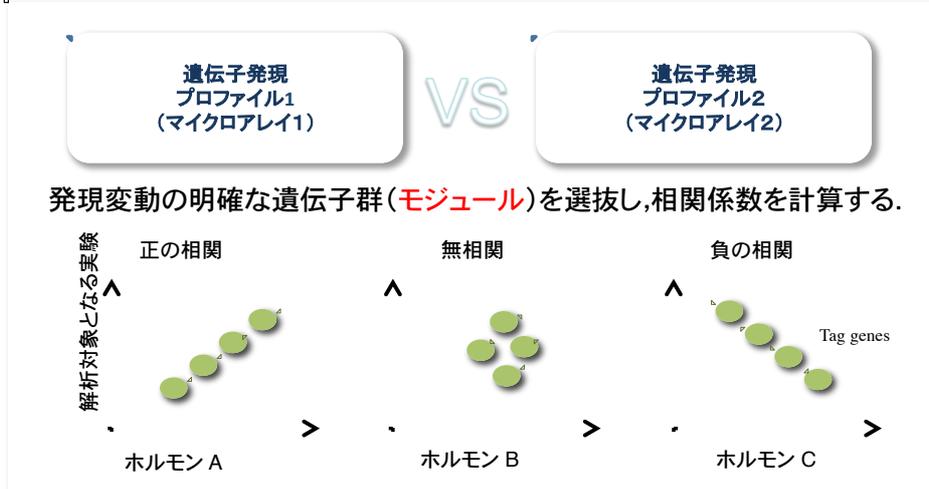
---

## 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、RNA-seqデータを 入力する	9 page
備考	14 page

# AtCASTによる解析の仕組み

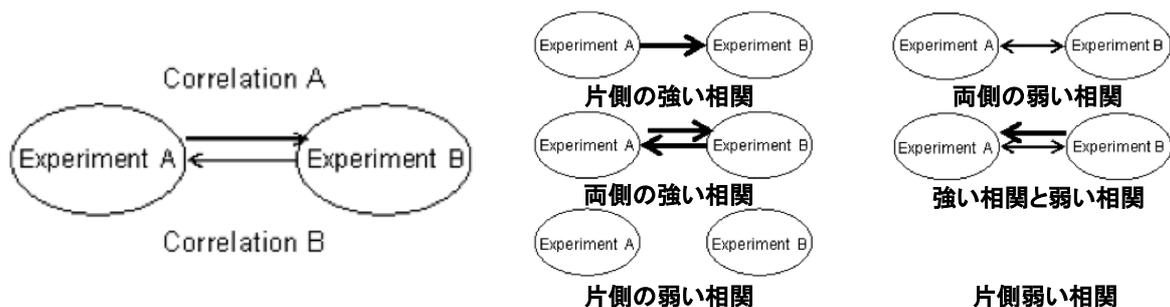
## マイクロアレイ実験間の遺伝子発現相関解析



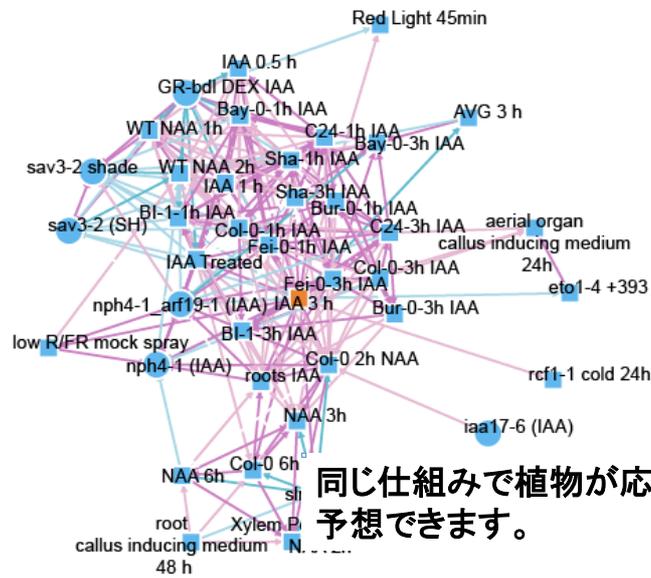
## 解析結果の表示



### 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  
準備したマイクロアレイデータを入力する

## AtCAST3:Home

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

### Welcome to AtCAST

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.

#### What's new

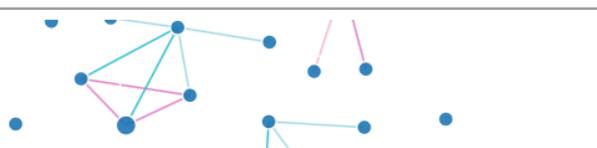
- More than 7000 microarray data were newly included.
- Gene Ontology term Enrichment (GOE) analysis is introduced to show the characteristics of the genes in a selected module.
- AtCAST3.0 supports "AG" (8k array), "ATH1" (23k array) and "Arabidopsis gene 1.1 ST array" (28k array) .
- Other microarray data can be input using AGI code.
- Input of RNA-seq data is supported.

These new features will facilitate analysis of associations between data obtained using different platforms.

>Analyze Public data

>Analyze your data

↑  
使い方1



# AtCAST3:Analyze public data

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

Search experiments by keyword 

Example: Auxin, IAA, GSE10248

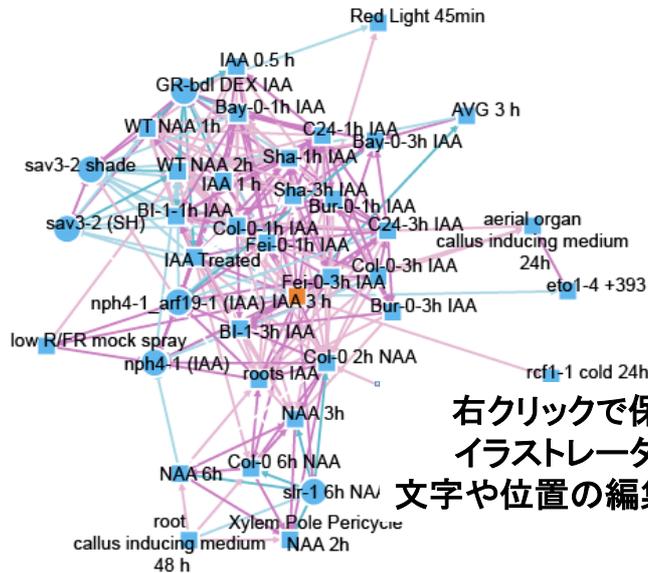
-----

↑  
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

Contact : [atpbsmd@yokohama-cu.ac.jp](mailto:atpbsmd@yokohama-cu.ac.jp)  
( When you send E-mail please change \* to @ )

[AtGenExpress JPN](#) --+--+ Version 3

# モジュール 相関ネットワーク図



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 "IAA 3 h" ?

Score	Experiment	Correlation from IAA 3 h	Correlation to IAA 3 h	Control Experiment	Genotype	Treatment	Tissue	Link to original data	Experimental Category	Data set Details
1	IAA 3 h	1.00 >More info	1.00 >More info	Mock 3 h	Col-0	1 -micro-M IAA	seedling	Original data	Basic hormone treatment of seedlings	link
2	Col-0-3h IAA	0.90 >More info	0.69 >More info	Col-0-0h IAA			whole plant	Original data	GEO bulk	link
3	Fei-0-3h IAA	0.88 >More info	0.81 >More info	Fei-0-0h IAA				Original data	GEO bulk	link
4	C24-3h IAA	0.86 >More info	0.80 >More info	C24-0h IAA				Original data	GEO bulk	link
5	BI-1-3h						whole plant	Original data	GEO bulk	link
6	Bur-0-						whole plant	Original data	GEO bulk	link
7	IAA 1 h	0.83 >More info	0.84 >More info	Mock 1 h	Col-0	1 -micro-M IAA	seedling	Original data	Basic hormone treatment of seedlings	link
8	Sha-3h IAA	0.83 >More info					ant	Original data	GEO bulk	link
9	Fei-0-1h IAA	0.83 >More info					ant	Original data	GEO bulk	link
10	C24-1h IAA	0.82 >More info					ant	Original data	GEO bulk	link
11	Bay-0-3h IAA	0.82 >More info					ant	Original data	GEO bulk	link
12	Col-0 6h NAA	0.81 >More info	0.76 >More info	Col-0 6h MOCK				Original data	GEO bulk	link
13	Mutant iaa17-6 (IAA)	0.79 >More info	0.13 >More info	Col (IAA)	iaa17-6	5 -micro-M IAA	seedling	Original data	GEO bulk	link
14	Bur-0-1h IAA	0.78 >More info	0.59 >More info	Bur-0h IAA			whole plant	Original data	GEO bulk	link
15	GR-bdf DEX IAA	-0.77 >More info	-0.63 >More info	GR-bdf IAA				Original data	GEO bulk	link
16	IAA Treated	0.77 >More info	0.81 >More info	Negative Control				Original data	GEO bulk	link

実験間反復などの統計情報

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)

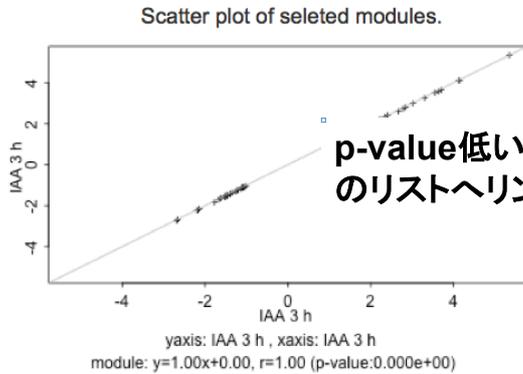
↑ モジュールの情報へのリンク

# IAA 3 h

Comparison with IAA 3 h [help](#)

モジュールの情報ページ

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



p-value低い順100位までのリストへリンク

GO enrichment(Up regulatd genes)



GOID	Term	Ontology	p-value
GO:0009733	response to auxin stimulus	BP	4.030e-12
GO:0009746	response to auxin stimulus	BP	4.025e-09
GO:0009745	response to auxin stimulus	BP	4.025e-09
GO:0009744	response to auxin stimulus	BP	4.025e-09
GO:0009743	response to auxin stimulus	BP	4.025e-09
GO:0009742	response to auxin stimulus	BP	4.025e-09
GO:0009741	response to auxin stimulus	BP	4.025e-09
GO:0009740	response to auxin stimulus	BP	4.025e-09
GO:0009739	response to auxin stimulus	BP	4.025e-09
GO:0009738	response to auxin stimulus	BP	4.025e-09
GO:0009737	response to auxin stimulus	BP	4.025e-09
GO:0009736	response to auxin stimulus	BP	4.025e-09
GO:0009735	response to auxin stimulus	BP	4.025e-09
GO:0009734	response to auxin stimulus	BP	4.025e-09
GO:0009733	response to auxin stimulus	BP	4.030e-12
GO:0044221	response to chemical stimulus	BP	1.271e-06
...	...	...	...

GO階層ネットワーク図へのリンク

GO enrichment(Down regulatd genes)



GOID	Term	Ontology	p-value
GO:0044765	single-organism transport	BP	4.019e-07
GO:0006790	sulfur compound metabolic process	BP	6.698e-07
GO:0044550	secondary metabolite biosynthetic process	BP	7.788e-07
GO:0019748	secondary metabolic process	BP	9.058e-07
GO:0044710	single-organism metabolic process	BP	1.968e-06

List of genes in module

up regulated genes down regulated genes

比較元の実験における発現

[download](#)

	probeID	AGICode	Annotation	Log <sub>2</sub> signal ratio IAA 3 h	Log <sub>2</sub> signal ratio GR-bdi DEX IAA
1	245076_at	AT2G23170	Auxin-responsive GH3 family protein	5.350	-1.832
2	253423_at	AT4G32280	indole-3-acetic acid inducible 29	4.150	-2.132
3	253066_at	AT4G37770	1-amino-cyclopropane-1-carboxylate synthase 8	4.124	-1.810
4	253908_at	AT4G27260	Auxin-responsive GH3 family protein	3.710	-1.426
5	258399_at	AT3G15540	indole-3-acetic acid inducible 19	3.636	-2.480
6	244911_at	ATMG00820	Reverse transcriptase (RNA-dependent DNA polymerase)	3.555	-0.711
7	263436_at	AT2G28690	Protein of unknown function (DUF1635)	3.305	-0.140
8	257766_at	AT3G23030	indole-3-acetic acid inducible 2	3.018	-1.388
9	252103_at	AT3G51410	Arabidopsis protein of unknown function (DUF241)	2.846	-1.053
10	248801_at	AT5G47370	Homeobox-leucine zipper protein 4 (HB-4) / HD-ZIP protein	2.785	-1.720
11	248163_at	AT5G54510	Auxin-responsive GH3 family protein	2.668	-1.220
12	245397_at	AT4G14560	indole-3-acetic acid inducible	2.415	-2.507
13	266364_at	AT2G41230	unknown protein; Has 75 Blast hits to 75 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 75; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	2.352	-0.159
14	266830_at	AT2G22810	1-aminocyclopropane-1-carboxylate synthase 4	2.224	-4.046
15	259773_at	AT1G29500	SAUR-like auxin-responsive protein family	1.986	-3.216
16	248282_at	AT5G52900	unknown protein; Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 173; Other Eukaryotes - 2996 (source: NCBI BLINK).	1.745	-0.159
17	245593_at	AT4G14550	indole-3-acetic acid inducible 14	1.745	-0.159
			unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein		

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)

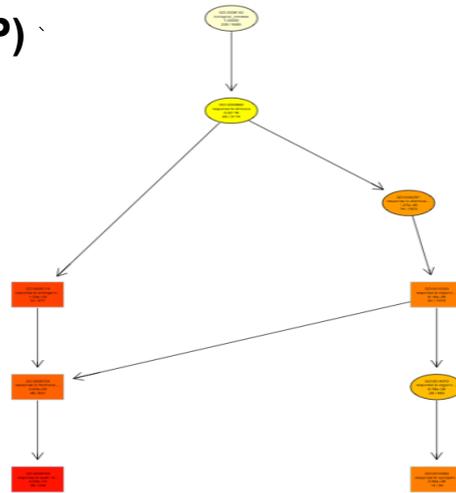
Biological process (BP)

外環境応答

化学物質応答

ホルモン応答

オーキシン応答



## 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

### AtCAST3:Home

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

#### Welcome to AtCAST

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.

#### What's new

- More than 7000 microarray data were newly included.
- Gene Ontology term Enrichment (GOE) analysis is introduced to show the characteristics of the genes in a selected module.
- AtCAST3.0 supports "AG" (8k array), "ATH1" (23k array) and "Arabidopsis gene 1.1 ST array" (28k array) .
- Other microarray data can be input using AGI code.
- Input of RNA-seq data is supported.

These new features will facilitate analysis of associations between data obtained using different platforms.

[>Analyze Public data](#)

[>Analyze your data](#)



# AtCAST3: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。

## AtCAST3: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

Control

ファイルが選択されていません。\*

ファイルが選択されていません。\*

ファイルが選択されていません。

ファイルが選択されていません。

ファイルが選択されていません。

Treatment

ファイルが選択されていません。\*

ファイルが選択されていません。\*

Platform

\*

GI Affymetrix ATH1-121501 Control Expen

Affymetrix Arabidopsis gene 1.1 ST

Affymetrix AG

Analyze using AGI code

Tissue

Control Experiment

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-BloB-5_at	456.3	P	0.030976
	AFFX-BloB-M_at	707.7	P	0.004484
	AFFX-BloB-3_at	346.6	A	0.131361
	AFFX-BloC-5_at	1546.7	P	0.000052
	AFFX-BloC-3_at	699.9	P	0.00141
	AFFX-BloDn-5_at	1439.8	P	0.000195
	AFFX-BloDn-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。

# AtCAST3: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。

## AtCAST3: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

Input File (Counts)

選択... ファイルが選択されていません。 \*

Control column number(s)

 \*  \*  \*  \*

Treatment column number(s)

 \*  \*  \*  \*  \*

Platform

RNA-seq \*

Genotype, Treatment, Tissue, Control Ex

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	4469.42	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)

ほとんど影響なく比較可能