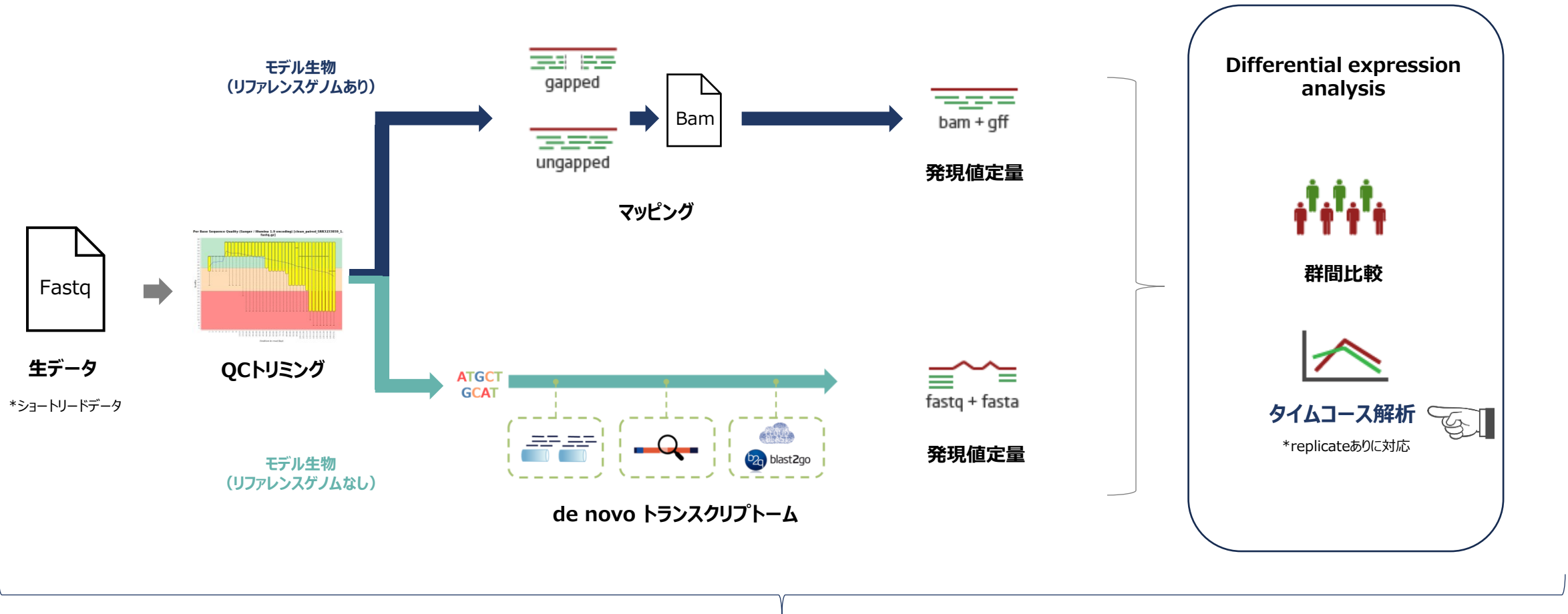




経時データのRNA-Seq解析

フィルジェン株式会社
バイオインフォマティクス部

RNA-Seqデータ解析



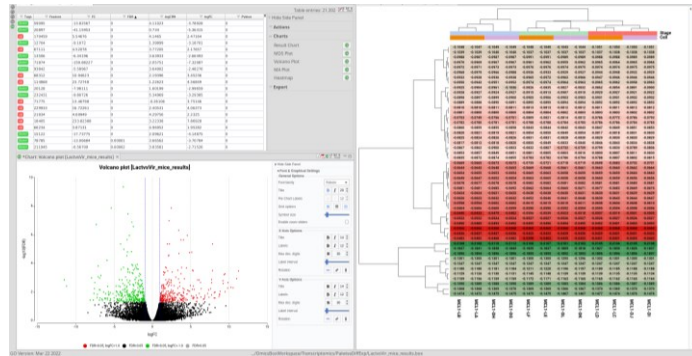
生データから解析可能
マウス操作
高スペックPCは不要

*受託サービスなどで得られた発現量に関するファイルにも対応



群間比較

発現変動のある遺伝子リスト、発現差を視覚化したボルケーノプロット、ヒートマップが得られます。



特定の条件(時点)における発現値の違い

例) 遺伝子Aは12hのテスト群で高発現

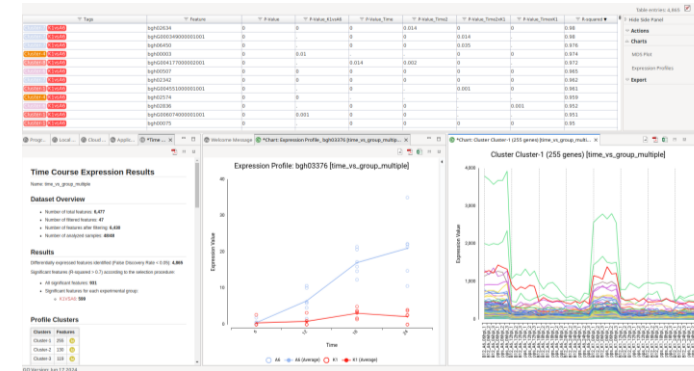


特定の条件間の発現値をもとにした
統計的差異にフォーカスした解析に適している



タイムコース解析

個々の遺伝子変動ではなく、各遺伝子発現パターンリスト、時系列プロットやクラスター図が得られます。



特定の時点ではなく、発現プロファイルの傾向の違い

例) 時間の経過とともに一貫して発現が増加する遺伝子群を含むクラスター1



遺伝子発現の動的変化や傾向に興味がある場合に適している

- 2つの異なる系統のArabidopsis thaliana (シロイヌナズナ)

病原体Bghに耐性
病原体Bghに感受性



- 各系統には2つの異なる病原体Bghの分離株 (A6 と K1) を接種



- 植物系統と病原体の組み合わせごとに、4つの異なるタイムポイント (6時間、12時間、18時間、24時間)



- 3回の反復

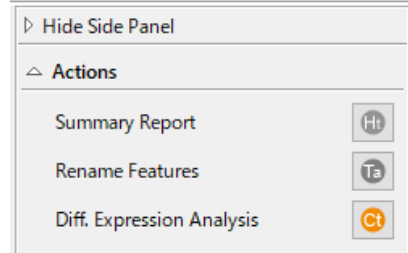
N=3

合計 48 のサンプル



事前にこれらの情報をエクセルなどの表計算ツール上にまとめた
実験設計データを作成し、タブ区切りテキストファイルとして保存しておく

	A	B	C	D	E
1	Sample	Time	pathogen	MLA1	replicate
2	B12_A6_6hpi_1	6 A6	resistant	r1	
3	B12_A6_6hpi_2	6 A6	resistant	r2	
4	B12_A6_6hpi_3	6 A6	resistant	r3	
5	B12_A6_12hpi_1	12 A6	resistant	r1	
6	B12_A6_12hpi_2	12 A6	resistant	r2	
7	B12_A6_12hpi_3	12 A6	resistant	r3	
8	B12_A6_18hpi_1	18 A6	resistant	r1	
9	B12_A6_18hpi_2	18 A6	resistant	r2	
10	B12_A6_18hpi_3	18 A6	resistant	r3	
11	B12_A6_24hpi_1	24 A6	resistant	r1	
12	B12_A6_24hpi_2	24 A6	resistant	r2	
13	B12_A6_24hpi_3	24 A6	resistant	r3	



発現値定量結果のサイドパネルからDifferential expression analysisをクリックし
Time course expression analysisを選択します。

*発現値テーブルから解析する場合は、Transcriptomics ModuleアイコンをクリックしLoad count tableよりデータをロードする。

💡 遺伝子フィルタリングは結果の信頼性を確保するために極めて重要です。一般的に、多くの遺伝子を除外することなく、フィルタリングのしきい値を可能な限り高く設定することが望ましいです。

今回の分析では、遺伝子フィルタリングしきい値を 2 CPM (Counts Per Million) に設定し、少なくとも 10 個のサンプルがこのしきい値を満たすことを要求しました。

サンプルしきい値は、最小の実験グループのサイズに対応することを推奨。
今回のデータセットは、このグループサイズは 12 (植物系統と病原体分離株の組み合わせ) です。堅牢性と包括性のバランスをとるため、しきい値を 10 に設定しました。

遺伝子フィルタリングに関する設定画面

Time Course Expression Analysis (counts_athaliana_timecourse)

Configuration 2

Experimental Design File

Experimental Design

Sample	Time	pathogen	MLA1	replicate
B12_A6_6hpi_1	6	A6	resistant	r1
B12_A6_6hpi_2	6	A6	resistant	r2
B12_A6_6hpi_3	6	A6	resistant	r3
B12_A6_12hpi_1	12	A6	resistant	r1
B12_A6_12hpi_2	12	A6	resistant	r2
B12_A6_12hpi_3	12	A6	resistant	r3
B12_A6_18hpi_1	18	A6	resistant	r1
B12_A6_18hpi_2	18	A6	resistant	r2
B12_A6_18hpi_3	18	A6	resistant	r3
B12_A6_24hpi_1	24	A6	resistant	r1
B12_A6_24hpi_2	24	A6	resistant	r2
B12_A6_24hpi_3	24	A6	resistant	r3

Default < Back **Next >** Run Cancel

事前に作成した実験設計データを入力

Time Course Expression Analysis (counts_athaliana_timecourse)

Configuration 3

The number of samples reaching CPM filter (10) has been set to a lower value than the size of the smallest sample group (24). You may consider updating this filtering parameter to the minimum sample group size.

Design Type

☐ Single Series Time Course

☒ Multiple Series Time Course

Targets

Time Factor: Time

Experimental Factor: MLA1

Control Condition: susceptible

Statistical Settings

Significance Level (Alpha): 0.05

R-squared Cutoff: 0.7

Visualization of Results

Number of Clusters: 9

Clustering Method: Hierarchical Clustering

Version Details:

- maSigPro 1.66.0
- R 4.1.2
- BiocManager 1.30.9

Please Cite:

- Conesa A., Nueda MJ., Ferrer A. and Talon M. (2006). maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. Bioinformatics (Oxford, England), 22(9), 1096-102.
- Nueda MJ., Tarazona S. and Conesa A. (2014). Next maSigPro: updating maSigPro bioconductor package for RNA-seq time series. Bioinformatics (Oxford, England), 30(18), 2598-602.

Default < Back Next > **Run** Cancel

分析オプション設定画面

「Multiple Series」を選択。

*時間条件のみ解析に使用する場合は「single Series」

Time factorには実験デザインの時系列の項目名を指定。
Experimental factorには解析に考慮したい項目と条件を入力。

今回は病原体に耐性か感受性であるかの条件を含めました。

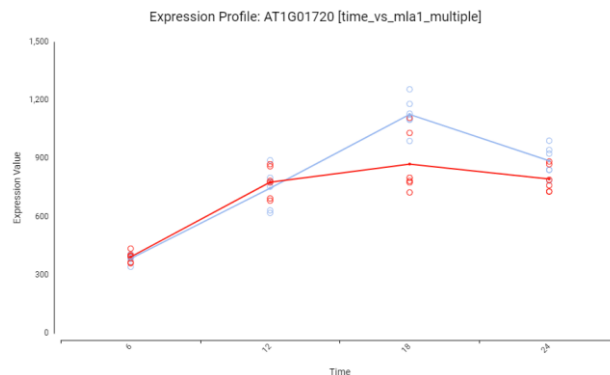
統計的しきい値、クラスタリング パラメータなどを指定。
解析結果を見てクラスター数を増減させることができます。

今回はデフォルト値のまま実行します。

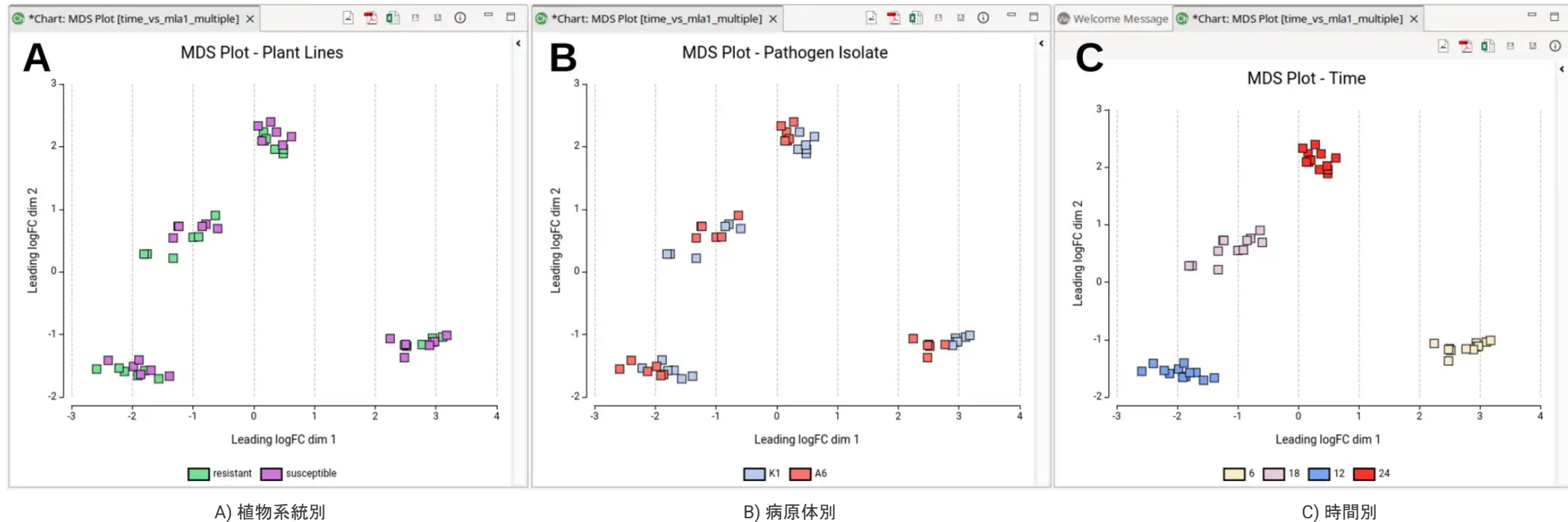
Tags	Feature	P-Value	P-Value_resistantvsusce...	P-Value_Time	P-Value_Time2	P-Value_Time2xresistant	P-Value_Timexresistant	R-squared
	AT2G24250	0	.	0.001	0	.	.	0.261
Cluster-5	AT5G51370	0	.	0	.	.	.	0.775
Cluster-4 resistantvsusceptible	AT1G01720	0	.	0	0	.	0.003	0.907
Cluster-6	AT4G38500	0	.	0	0	.	.	0.731
	AT2G25580	0	.	0	0	.	.	0.329
Cluster-5 resistantvsusceptible	AT5G13090	0	0.039	0	0	.	.	0.797
Cluster-3	AT4G26530	0	.	0	0	.	.	0.729
	AT2G24290	0.001	.	0.001	0.001	.	.	0.251
	AT2G37560	0	.	0	0.002	.	.	0.487
Cluster-2	AT5G63370	0	.	0	0	.	.	0.757
	AT5G49060	0	.	0	.	.	.	0.246

経時的に変動が見られた遺伝子がクラスターに分けられタグがつけられます。
また、条件間で遺伝子発現傾向に違いがある場合は、条件タグが追加されます。

まずは、サイドパネルからChartsを展開し、各種グラフを作成し、今回の設定が適切であったか評価していきます。



右クリックで、
個々の遺伝子の発現プロットを作図できます。



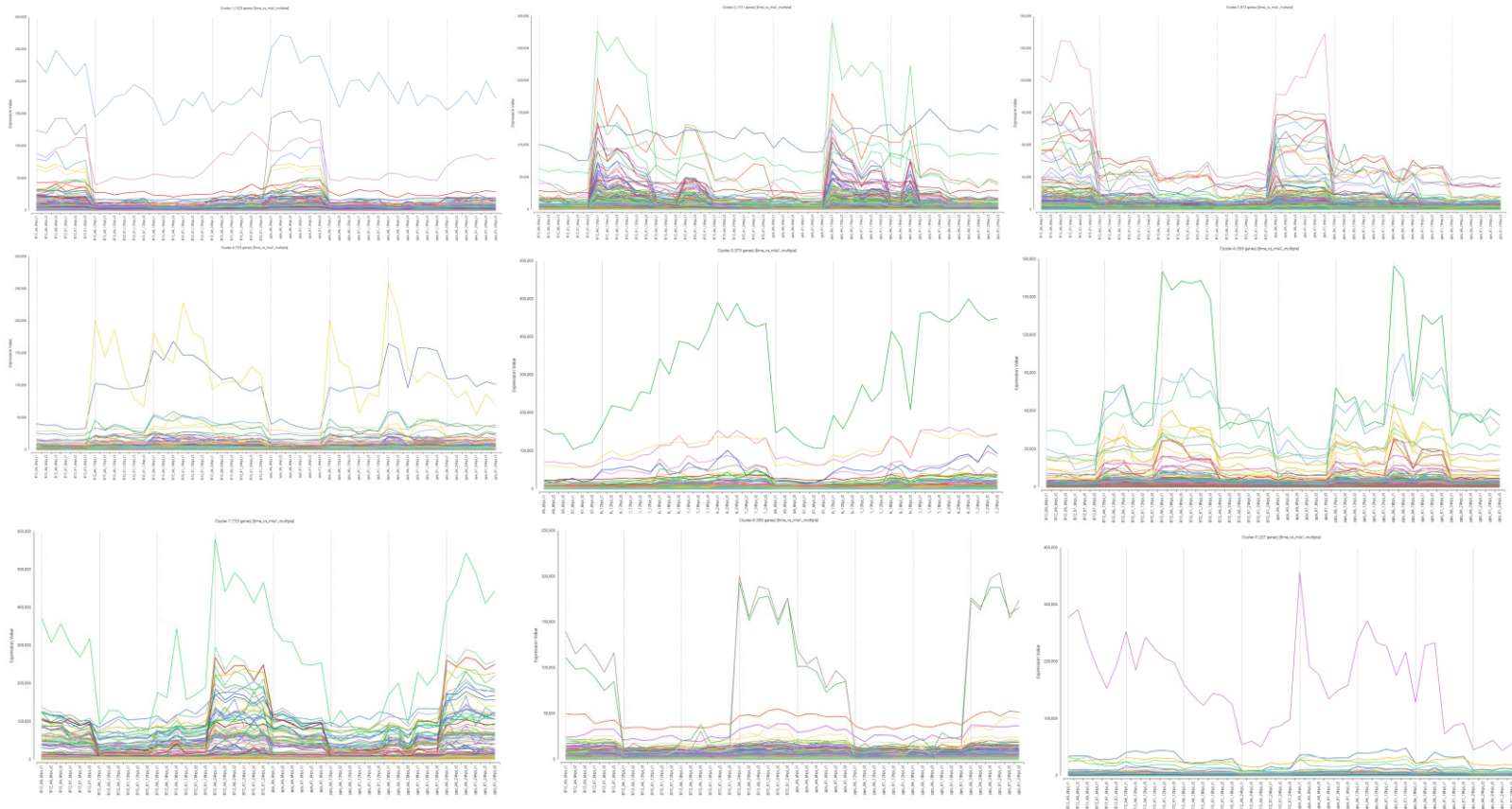
異なる実験条件に従ってMDSプロットを色分けすることで、全体的な遺伝子発現プロファイルへの影響を評価することができます。

時間がサンプル間の変動の主な原因であることが明らかである（図C）。
各時間クラスター内では、異なる植物系統のサンプルが混在しているため、
植物系統が全体的な遺伝子発現に与える影響は少ないようである（図A,C）。

しかし、病原体の分離株は遺伝子発現に影響を与えるようで、この因子によって内部的に分離された時間クラスターもある（図B,C）。

結論として、このデータセットでは、接種後の時間と病原体の分離が2つの主要な変動要因であるため、「Experimental Factor」の設定にて、「病原体」を選択し、タイムコース解析を再実行することにより、興味深い結果が得られる可能性があります。

実験全体の発現プロファイルプロットには、各クラスター内の遺伝子のサンプル全体の発現レベルが表示されます。



クラスターの数にデータに適しているか評価することに役立ちます。

明確なクラスタープロファイル:
各クラスターが明確なプロファイルを示している場合、クラスターの数適切であることを示します。

あいまいなクラスター:
クラスターが明確に定義されていない場合、クラスターの数不十分である可能性があります。

クラスター間の類似したプロファイル:
クラスターが類似したプロファイルを示している場合、クラスターの数多すぎること示している可能性があります。

概要発現プロファイルは、異なる時点における各クラスター内の遺伝子の発現レベルの中央値の概要を示します。

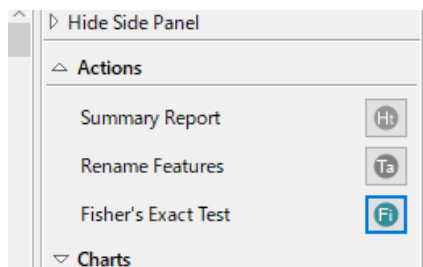
データ内の傾向とパターンを識別するのに役立ちます。

クラスター 5,7:
このクラスター内の遺伝子は、時間の経過とともに一貫して発現が増加しています。

クラスター 3:
このクラスター内の遺伝子は、時間の経過とともに発現が減少しています。

この観察結果は、前ページの実験全体の発現プロファイルプロットで見られるものと一致しています。

これらの結果をさらに解釈するには、特定のクラスターの遺伝子が特定の生物学的機能に関与しているかどうかを評価することが有用です。これは、OmicsBox のエンリッチメント解析 (Fisher's Exact Test) によって実現できます。



タイムコース解析の結果のサイドパネルから
Fisher's Exact Testを選択します。

*Functional Analysisモジュールの機能

Fisher's Exact Test (time_vs_mla1_multiple)

Fisher's Exact Test Input

Test Set

Groups: Cluster-8

Reference Set

Remaining Genes: ☒

Groups: Cluster-2

Functional Annotation

Note:
Only those genes that are included in the test or reference sets, and are also present in the functional annotation, will be utilized for the analysis.

Reference Annotation: Browse...

Default < Back **Next >** Run Cancel

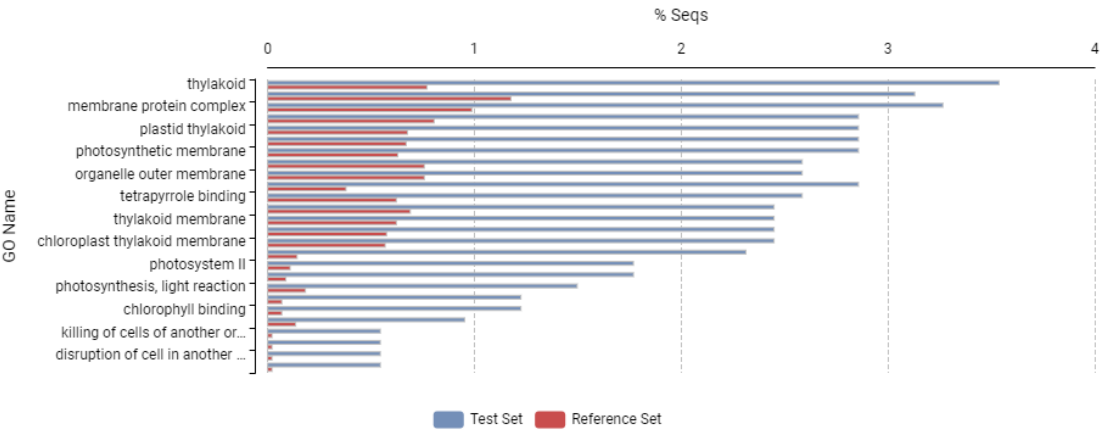
解析の対象となるクラスターを指定

リファレンスアノテーションを指定
解析結果の遺伝子名と一致する参照アノテーションデータを別途ダウンロードしておく必要があります。

今回はOmicsBoxに搭載されたBioMartからのデータダウンロード機能を使用し得たファイルを使用 (functional analysis > Load > Load Data from BioMart)。

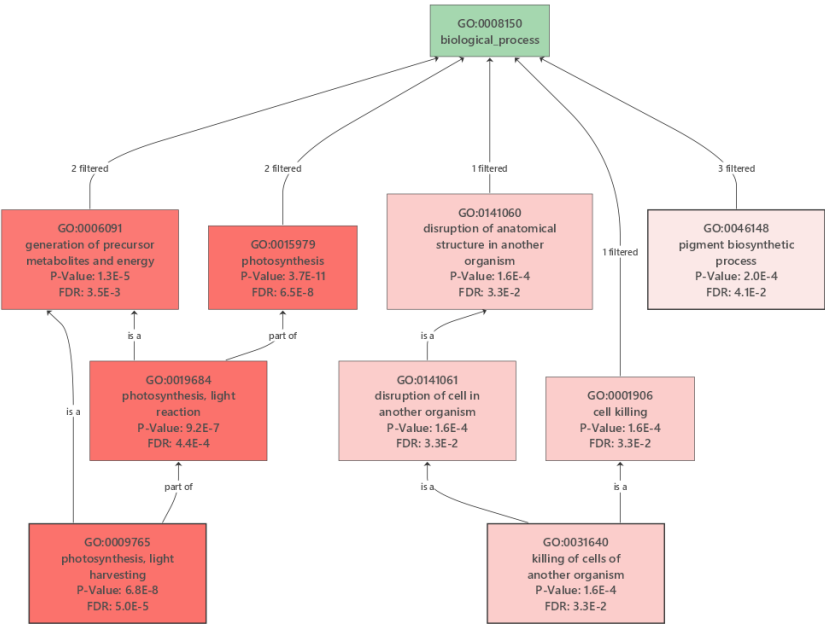
Tag	GO Term	GO Name	GO Category	Adj. P-value ▲	P-value	Nr Test	Nr Reference	Not Annot Test	Not Annot Ref
OVER	GO:0009521	photosystem	CELLULAR_COMPONENT	1.500753E-9	2.834818E-13	17	21	718	14772
OVER	GO:0009522	photosystem I	CELLULAR_COMPONENT	6.542519E-8	3.169557E-11	13	13	722	14780
OVER	GO:0015979	photosynthesis	BIOLOGICAL_PROCESS	6.542519E-8	3.70751E-11	21	56	714	14737
OVER	GO:0009523	photosystem II	CELLULAR_COMPONENT	2.398276E-7	1.812071E-10	13	16	722	14777
OVER	GO:0009579	thylakoid	CELLULAR_COMPONENT	1.920665E-6	1.814002E-9	26	114	709	14679
OVER	GO:0009765	photosynthesis, light harvest...	BIOLOGICAL_PROCESS	5.039688E-5	6.843737E-8	9	10	726	14783
OVER	GO:0016168	chlorophyll binding	MOLECULAR_FUNCTION	5.039688E-5	6.843737E-8	9	10	726	14783
OVER	GO:0034357	photosynthetic membrane	CELLULAR_COMPONENT	5.039688E-5	7.615698E-8	21	93	714	14700
OVER	GO:0009534	chloroplast thylakoid	CELLULAR_COMPONENT	1.116745E-4	1.898508E-7	21	99	714	14694

Enriched Bar Chart



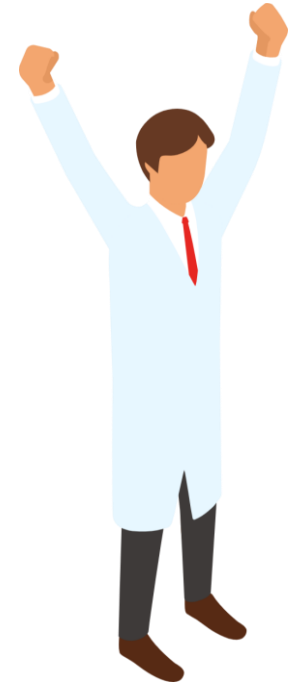
特定のクラスターにどのような機能が含まれているか調べることができます。

pigment metabolic process
nucleoside phosphate catabolic process
ribonucleoside diphosphate catabolic process
organic hydroxy compound metabolic process
ribonucleoside diphosphate metabolic process
purine-containing compound catabolic process
purine nucleoside diphosphate catabolic process
purine nucleoside diphosphate metabolic process
purine nucleoside triphosphate metabolic process
purine ribonucleoside triphosphate metabolic process
disruption of anatomical structure in another organism
regulation of nucleobase-containing compound metabolic process
nucleotide catabolic process
terpenoid metabolic process
carotenoid metabolic process
photosynthesis, light reaction
flavonoid biosynthetic process
terpenoid biosynthetic process
photosynthesis, light harvesting
ribonucleotide catabolic process
tetraterpenoid metabolic process
purine nucleotide catabolic process
organophosphate catabolic process
phenylpropanoid metabolic process
regulation of RNA metabolic process
regulation of RNA biosynthetic process



OmicsBox の経時データRNA-seq

- 生データ、あるいは受託サービスで得られたデータを使用
- 初心者でも解析できるインターフェース
- 7日間無料のデモライセンス→ [詳細\(PDF\)](#)



お問い合わせ先：フィルジェン株式会社

TEL 052-624-4388 (9:00～17：00)

FAX 052-624-4389

E-mail: support@filgen.jp