

DIA 定量プロテオーム解析ソフトウェア Scaffold DIA

Volcano plot

図の補足説明

→ P.34

日本語マニュアルの該当ページ
×対応するスライド番号

スライド番号→

1

1

Scaffold DIA 何ができるか

- DIAデータのタンパク質の同定、定量
- **検索対象**: 4種の検索に対応
 - FASTA
 - BLIB (DDA-ライブラリ),
 - **DLIB** (prositライブラリ、配列から理論DDAライブラリ作成)
 - ELIB (DIA解析結果ライブラリ)
- 定量解析/検定とグラフ表示
- Viewerによる結果ファイルのシェア

対応フォーマット: ProteoWizard Msconvert を使っている

各社メーカーのファイルフォーマット [読み込みなどがよく確認されているもの]

装置メーカー

ファイルフォーマット

SCIEX

*.wiff (.wiff.scanも同じフォルダに置く事)

Agilent

*.d (ディレクトリ)

Thermo

*.raw

オープンフォーマット (原理的には mzML フォーマットになっていればデータの読み込みが可能)

HUPO Proteomics Standards Initiative mzML

*.mzML

* mzMLはほぼすべてのメーカーで変換可能なフォーマット

Samples 画面

同定されたタンパク質の一覧並びに定量値をはじめとするタンパク質の関連情報が表示

解析結果の概要を確認する上で主体となる画面

タンパク質

GO情報

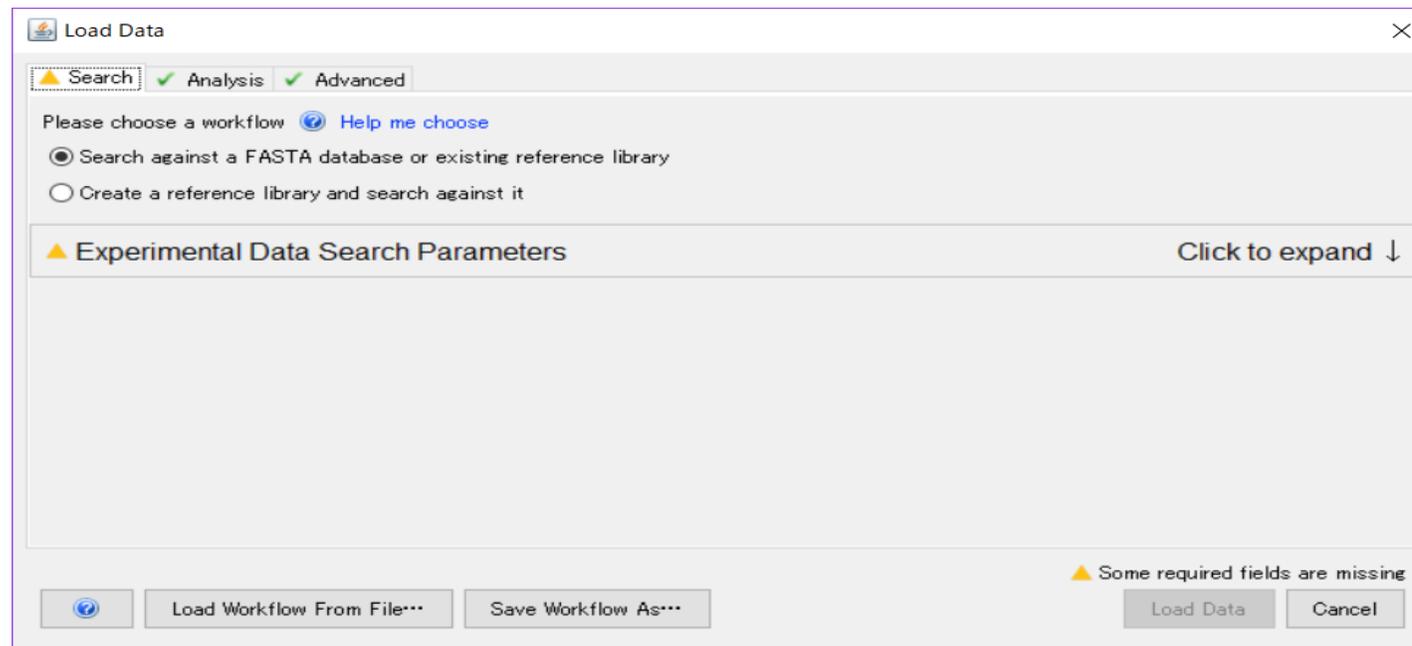
定量値

#	Variable Star	Protein Name	Molecular Weight	Protein Group Score	Identified Peptide Count	Exclusivity	t-test CL: Treatment	Taxonomy	Biological Process	Cellular Component	Molecular Function	Control	Insulin
1	✓	sp P49792 RBP2_HUMAN E3 SUMO-protein ligase Ra...	358 kDa	1	54	100%	0.26	Homo sapiens	biological adhesion, biological regulation, cellular process, developmental process, enzyme regulator activity, establishment of localization, growth, cytoplasm, extracellular region, membrane, nucleus, ribosome, binding, catalytic activity, structural molecule activity			2.45E8	2.24E8
2	✓	sp P15924 DESP_HUMAN Desmoplakin OS=Homo sa...	332 kDa	1	53	100%	0.94	Homo sapiens				1.32E8	1.39E8
3	✓	sp O43707 ACTN4_HUMAN Alpha-actinin-4 OS=Hom...	105 kDa	1	51	69%	0.48	Homo sapiens				6.17E8	5.81E8
4	✓	sp Q5T4S7 UBRR4_HUMAN E3 ubiquitin-protein ligase ...	574 kDa	1	49	100%	0.27	Homo sapiens				8.93E7	8.06E7
5	✓	sp P11388 TOP2A_HUMAN DNA topoisomerase 2- alp...	174 kDa	1	44	100%	0.036	Homo sapiens				4.16E8	4.49E8
6	✓	sp P18206 VINC_HUMAN Vinculin OS=Homo sapiens ...	124 kDa	1	43	100%	0.31	Homo sapiens				6.58E8	5.44E8
7	✓	sp P12814 ACTN1_HUMAN Alpha-actinin-1 OS=Hom...	103 kDa	1	43	63%	0.49	Homo sapiens				1.77E8	1.68E8
8	✓	sp Q9NR30 DDX21_HUMAN Nucleolar RNA helicase 2...	87 kDa	1	41	100%	0.23	Homo sapiens				5.71E8	5.91E8
9	✓	sp Q9P2E9 RRBP1_HUMAN Ribosome-binding protei...	152 kDa	1	41	100%	0.090	Homo sapiens				1.22E8	1.04E8
10	✓	sp Q86UP2 KTN1_HUMAN Kinectin OS=Homo sapien...	156 kDa	1	41	100%	0.18	Homo sapiens				1.24E8	1.09E8
11	✓	sp O75533 SF3B1_HUMAN Splicing factor 3B subunit ...	146 kDa	1	40	100%	0.14	Homo sapiens				2.66E8	2.42E8
12	✓	sp P34932 HSP74_HUMAN Heat shock 70 kDa protein...	94 kDa	1	38	100%	0.65	Homo sapiens				5.61E8	5.55E8
13	✓	sp P26639 SYTC_HUMAN Threonine--tRNA ligase, cyt...	83 kDa	1	38	100%	0.51	Homo sapiens				8.78E8	9.20E8
14	✓	sp Q14152 EIF3A_HUMAN Eukaryotic translation initi...	167 kDa	1	38	100%	0.83	Homo sapiens				4.83E8	4.82E8
15	✓	sp P22102 PUR2_HUMAN Trifunctional purine biosyn...	108 kDa	1	37	100%	0.65	Homo sapiens				5.69E8	5.46E8
16	✓	sp P50990 TCPO_HUMAN T-complex protein 1 subun...	60 kDa	1	36	100%	0.25	Homo sapiens				1.20E9	1.27E9
17	✓	sp P78371 TCPB_HUMAN T-complex protein 1 subuni...	57 kDa	1	35	100%	0.39	Homo sapiens				7.59E8	8.82E8
18	✓	sp Q04637 IF4G1_HUMAN Eukaryotic translation initi...	175 kDa	1	34	100%	0.51	Homo sapiens				4.85E8	4.79E8
19	✓	sp P14625 ENPL_HUMAN Endoplasmic OS=Homo sa...	92 kDa	1	34	100%	0.89	Homo sapiens				1.19E9	1.17E9
20	✓	sp P08729 K2C7_HUMAN Keratin, type II cytoskeletal ...	51 kDa	1	33	100%	0.61	Homo sapiens				2.17E9	2.18E9
21	✓	sp P00558 PGK1_HUMAN Phosphoglycerate kinase 1 ...	45 kDa	1	30	100%	0.52	Homo sapiens				4.58E9	4.82E9
22	✓	sp P06733 ENO4_HUMAN Alpha-enolase OS=Homo ...	47 kDa	1	30	100%	0.13	Homo sapiens				7.42E9	1.01E10
23	✓	sp Q5UIP0 RIF1_HUMAN Telomere-associated protein...	274 kDa	1	25	100%	0.71	Homo sapiens				3.07E7	2.96E7

データ取り込み時の画面

データ取り込み開始

- 以下のいずれかの操作を実施
 - メニューのFile -> New
 - Ctrl + N
 - メニューバー下にあるアイコン  をクリック



① workflow : elib作成と連動した2段階検索かそれ以外か

Search a reference library : 直接ライブラリー検索を実施

Create a chromatogram library and search against it : 2段階検索

② ライブラリ、タンパク質配列の指定

直接検索対象とするファイル(elib, dlib, blib, fasta)と、それに対応する配列データベース(fasta)を選択

③ parameter [ライブラリ検索の時には意味のない設定もあり]

Instrument Type :

推奨パラメーターセットの提案

Fragmentation :

CID/HCD/ETDから選択。考慮するイオンシリーズ

Precursor / Fragment / Library Fragment Tolerance :

理論値と実測スペクトルデータとの許容誤差範囲

Digestion Enzyme : ペプチド切断方法

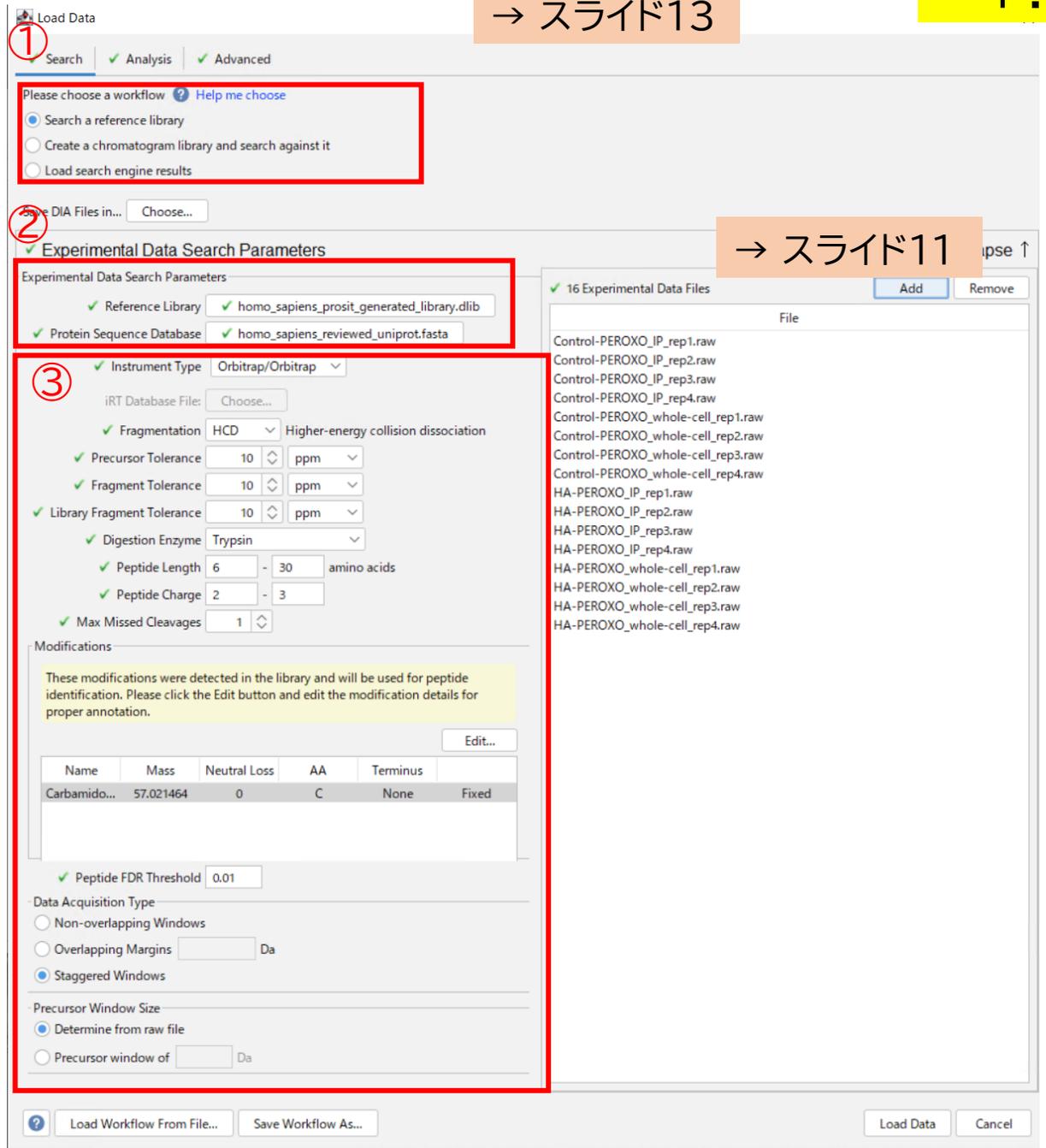
Peptide Length : 検索対象とするペプチドの長さ

Peptide Charge : 検索対象とするペプチドの電荷

Max Missed Cleavages :

理論ペプチド作成の際、切断対象のアミノ酸を何回無視したペプチドを作成するかの設定

Modifications: 考慮する修飾



④ Peptide FDR Threshold :

ペプチドの同定基準値、FDRの値

⑤ Data Acquisition Type

DIAの測定でのPrecursor Isolation Window タイプについて
[次、次々スライドに補足説明図]

- Non-overlapping Windows :

両端のオーバーラップ領域がない

- Overlapping Margins :

両端のオーバーラップ領域がある (値も指定)

- Staggered Windows :

window領域が半分ずつずれた測定を実施

⑥ Precursor Window Size

DIAの測定Window に対する設定。通常はrawファイルに書かれた情報をそのまま利用

⑦ Experimental Data File

検索対象となるrawファイルを選択

The screenshot shows the 'Load Data' software interface. At the top right, there is a yellow box with the text '→ P.17'. The interface has a top navigation bar with 'Search', 'Analysis', and 'Advanced' tabs. Below this, there are options to 'Please choose a workflow' with 'Search a reference library' selected. The main area is titled 'Experimental Data Search Parameters' and contains various settings: Reference Library (saccharomyces_cerevisiae_prosit_generated_library.dlib), Protein Sequence Database (saccharomyces_cerevisiae_reviewed_uniprot.fasta), Fragmentation (HCD), Precursor Tolerance (10 ppm), Fragment Tolerance (10 ppm), Library Fragment Tolerance (10 ppm), Digestion Enzyme (Trypsin), Peptide Length (6-30 amino acids), Peptide Charge (2-3), and Max Missed Cleavages (1). A table under 'Modifications' shows 'Carbamidom...' with mass 57.021464, neutral loss 0, AA C, terminus None, and fixed. On the right, a list of '1 Experimental Data File' shows '90min_DDA_HEK293_200ng_1.raw'. At the bottom, there are sections for 'Peptide FDR Threshold' (0.01), 'Data Acquisition Type' (Non-overlapping Windows selected), and 'Precursor Window Size' (Determine from raw file selected). A red box highlights these bottom sections. A red circle with the number '7' is next to the data file list. An orange box at the bottom right contains the text '→ スライド 9,10'. At the bottom of the interface are buttons for 'Load Workflow From File...', 'Save Workflow As...', 'Load Data', and 'Cancel'.

④

⑤

⑥

補足説明：対応するDIA解析 -- marginあり

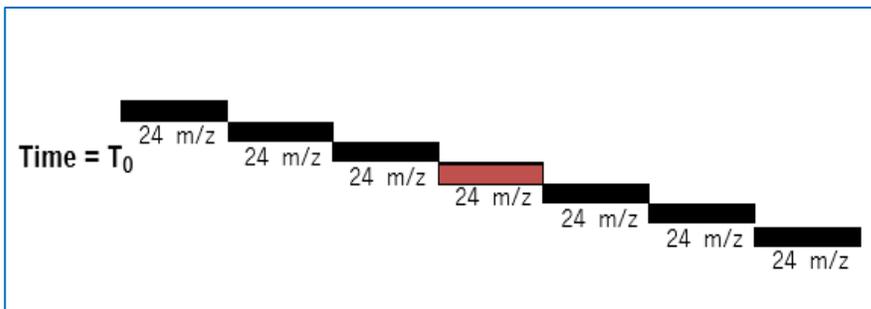
Data Acquisition Type

Non-overlapping Windows

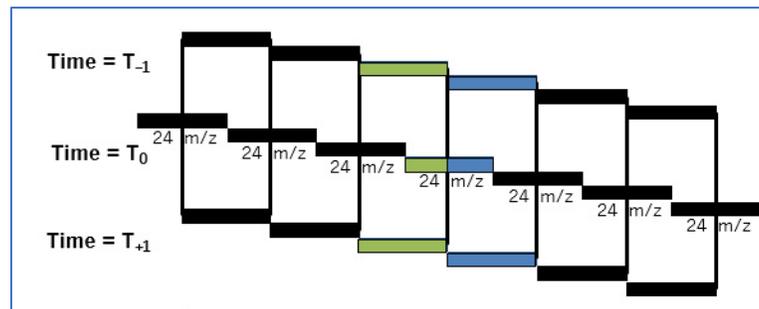
Overlapping Margins Da

Staggered Windows

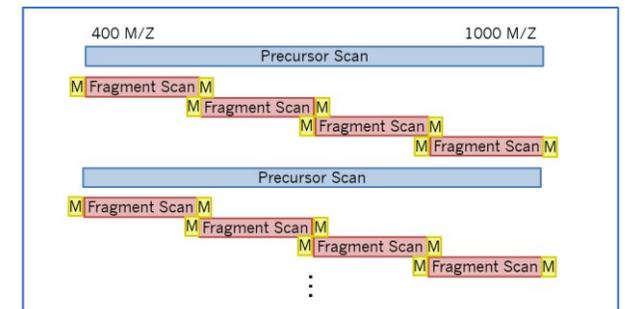
Non-overlapping



Overlapping

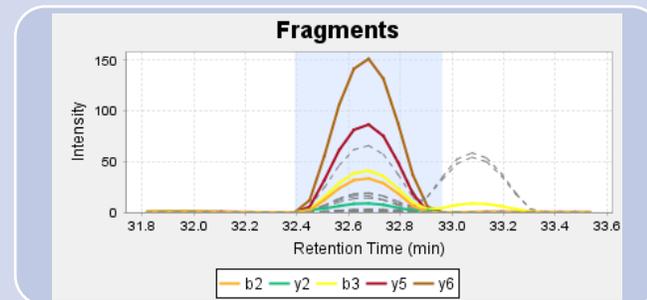
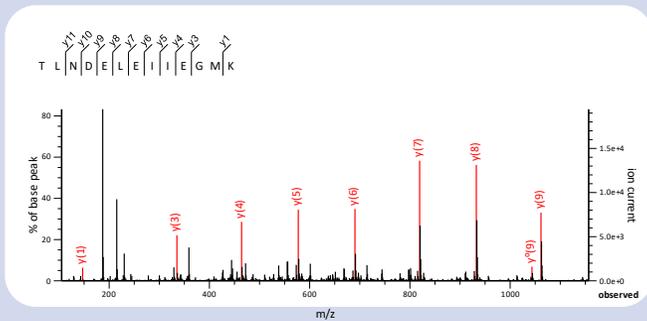


Staggered



補足説明：検索対象 FASTA、BLIB、ELIB (+DLIB)

```
MAFKDGTGKTPVEPEVAIHRIRITLTSRNVKSLKVKCADLIRGAKEKNLKVKGVPVMPKTLRITTF  
QMRIPKRLIDLHSPSEIVKQITSISIEPGVEVEVTIADA  
>sp|A2BDB0|ACTG_XENLA Actin, cytoplasmic 2 OS=Xenopus laevis GN=fac  
MEEELAALVINDGSGMCKAGFAGDDAPRAVFPISVGRPRHQGVMMGQKDSYVGDEAQSQRGILT  
DMEKIQWHTFYNELRVAPEEHPVLLTEAPLNPKANREKMTQIMFETFTNPAMYVAIQAVLSLYASC  
HTVPVIEGYALPHAILRLDLAGRDLTDYLMKILTERGYSFTTTAEREIVRDIKEKLCYVALDFEQL  
ELPDGQVITIGNERFRCPALFQPSFLGMESCGIHTTFNSIMKCDVDIRKDLYANTVLSGGTTM  
APSTMKIKIAPPERRKYSVWIGGSLASLSTFQMMWISKQDEYDESGPSIVHRKCF  
>sp|A2Q0Z0|EF1A1_HORSE Elongation factor 1-alpha 1 OS=Equus caball  
MGKEKTHINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAEMGKGSFKYAWVLDKLEAE  
ETSKYVVTIIDAPGHRDFIKNMIITGTSQADCAVLIVAAGVGEFEAGISKNGOTREHALLAYTLGV  
PYSQKRYEEIVKEVSTYIKKIGYNPDTVAFPVPSGWNQDNMLEPSANMPWFKGMKVTTRKQGNASGT  
PTDKPLRPLQDYYKIGGIGTVPVGRVETGVLKPGMVTFAFYNVTEVKSVMHHEALSEALPGC  
RRGNVAGDSKNDPPEAAGFTAQVITLNIHPGQISAGYAPVPCHTAHIAKFAELKEKIDRRSGK  
IVDMVPGKPMCVESFSDYPLGRFAVRDMRQTVAVGVIAKAVDKKAAGAGKVTKSAGKAKAK
```



FASTA

- 配列ファイル

*配列から測定条件などを考慮して計算されたスペクトルデータ (DLIB) も使用できる。

BLIB

- (DDA) マススペクトルピークリスト

ELIB

- (DIA) RT + m/zのピーク位置

*DLIBフォーマットはELIBとほぼ同じ

補足説明：Prositで作成されたライブラリ、DLIB

Library Creation Parameters

The following parameters were used in the conversion process and thus should match your instrumentation settings. If they vary dramatically from your settings please contact us.

Parameter	Setting
Charge Range	2 - 3
Maximum Missed Cleavages	1
m/z Range	396.4 - 1002.7
Default NCE	33
Default Charge	3

Featured Libraries

Coronavirus reference

Download Coronavirus only DLIB (1.9 MB)

Download Coronavirus plus Human (pan human) DLIB (259 MB)

Download Coronavirus only FASTA - 13 entries

Download Coronavirus plus Human FASTA - 20,350 entries

Available Libraries

Arabidopsis thaliana

Download Arabidopsis thaliana DLIB (1.2 GB)

Download FASTA file accessed 10/22/19 - 15,896 entries

Caenorhabditis elegans

Download Caenorhabditis elegans DLIB (380 MB)

Download FASTA file accessed 10/23/19 - 4,089 entries

Danio rerio

Download Danio rerio DLIB (250 MB)

Download FASTA file accessed 10/22/19 - 3,125 entries

Drosophila melanogaster

Download Drosophila melanogaster DLIB (365 MB)

Download FASTA file accessed 10/22/19 - 3,586 entries

Escherichia coli (strain K-12)

<https://support.proteomesoftware.com/hc/en-us/articles/360035151172-Prosit-Derived-Spectral-Libraries-for-Scaffold-DIA-Searches>

Prosit

- ・配列からdeep Neural Networkアルゴリズムで計算された理論スペクトルで、保持時間やintensityも予測

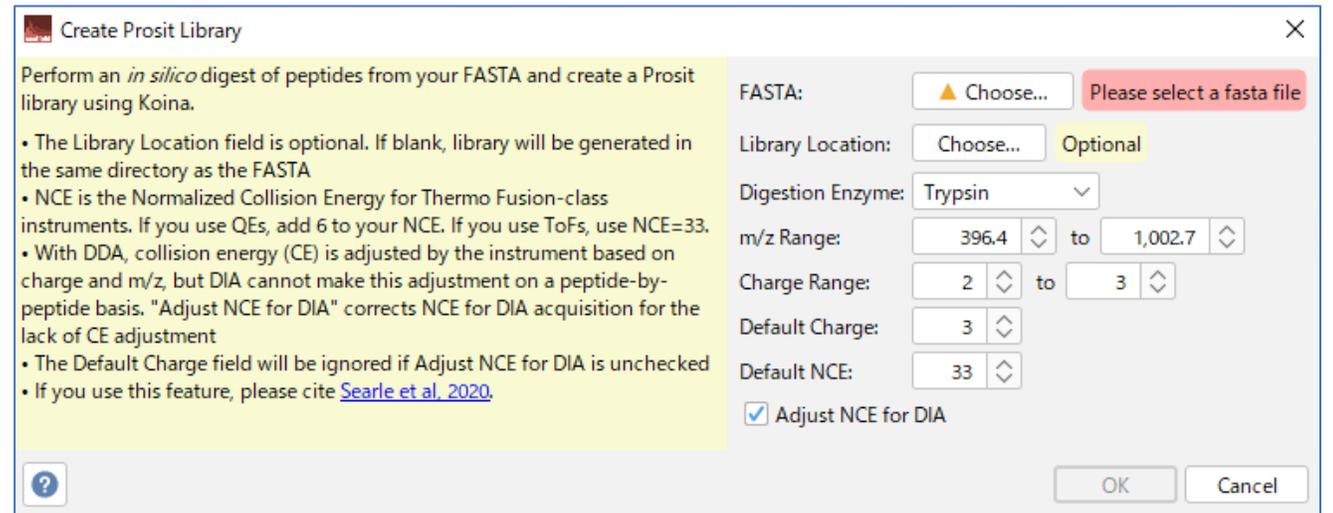
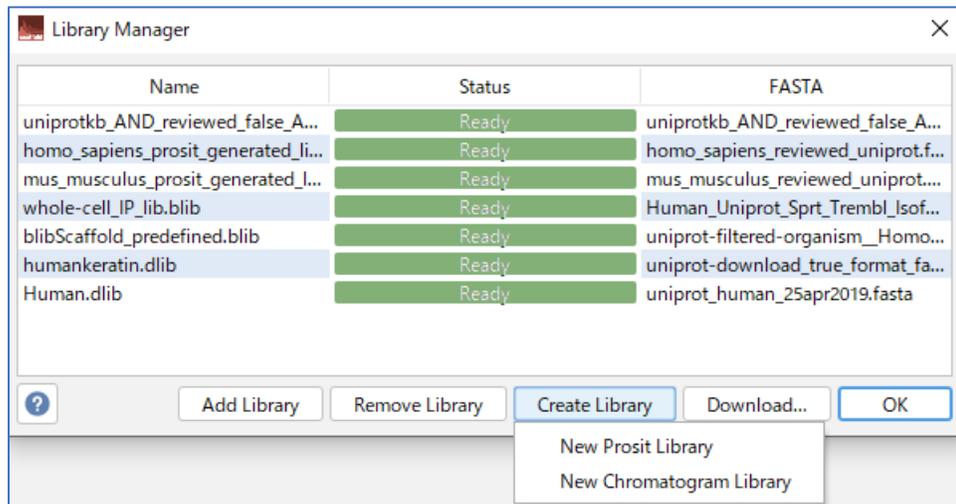
- ・Proteome Software社側で準備したDLIBを公開している

- ・多くのケースでBLIBの代わりになる、手軽に使える検索対象として利用してください

- ・Scaffold DIA上にFASTAからprositのインプットファイルを作成するツールあり

- ・修飾が苦手

補足説明2 : FASTAからprositライブラリを作成(4.0)



1.(menu) File :open library manager
-> Create Library
-> New Prosit Library

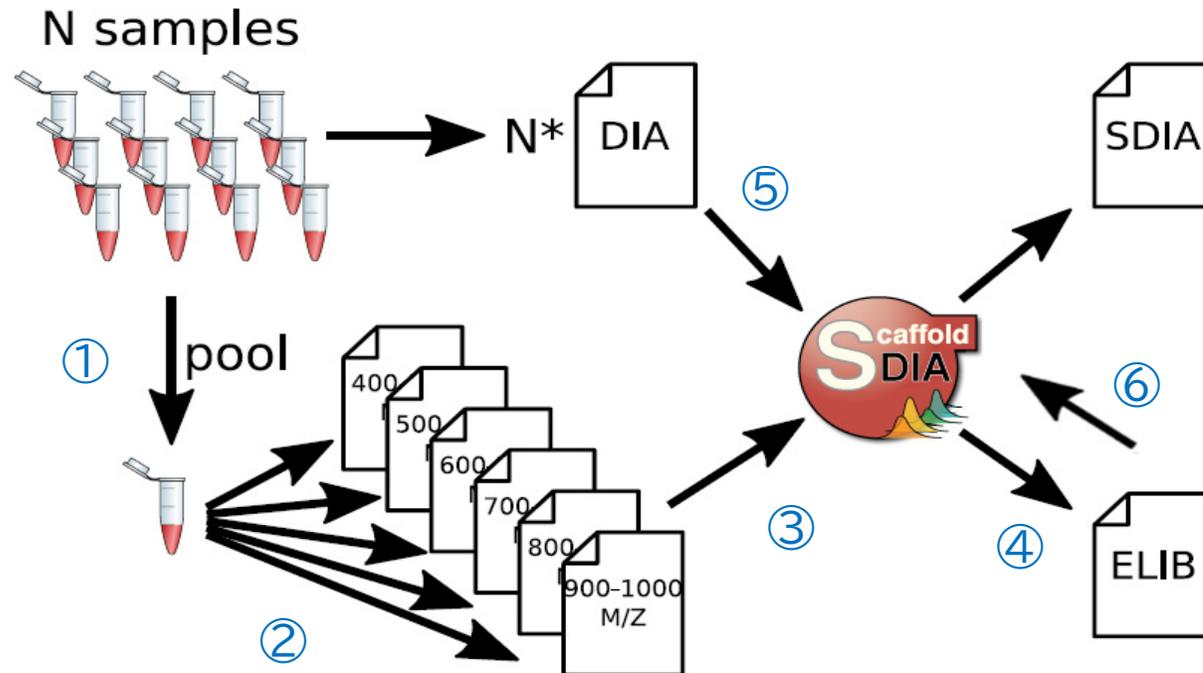
2.(dlib)library作成のためのパラメータを入力し、
”OK” (よくわからなければデフォルト値)

補足説明：ELIBと「2段階」検索

「pool」データへの検索でELIBを作成し、そのELIBに対して元のデータを検索するプロトコル例

The Combined Workflow

For the highest quality identifications, Scaffold DIA can create a library. A pooled sample is searched with XCorDIA and the resulting ELIB is used to extract and quantify peptides from experimental samples.



Analysisタブ、Advanced タブ、取り込み開始

Load Data

Search Analysis Advanced

Shared Evidence Clustering

- Perfect shared evidence protein groups
- Moderate shared evidence clusters
- Any shared evidence clusters

Target Protein FDR 1.0% FDR

Minimum Number of Peptides 2

Some required fields are missing

Load Data Cancel

Analysis
→ P.18

Load Data

Search Analysis Advanced

Processing Directory

C:\temp Help

- Create a subfolder for intermediate files and remove it upon completion
- Create a subfolder for intermediate files and retain it
- Write intermediate files directly to this location

Minimum Number Of Quant Ions 3

Maximum Number Of Quant Ions 5

Percolator Training Set Size 500000

Percolator Training Set Threshold 0.001

Filter RT - min

Advanced
→ P.20

すべての項目を埋めてから「Load Data」ボタンを押す

Load Data Cancel

ペプチド・タンパク質の定性(同定)

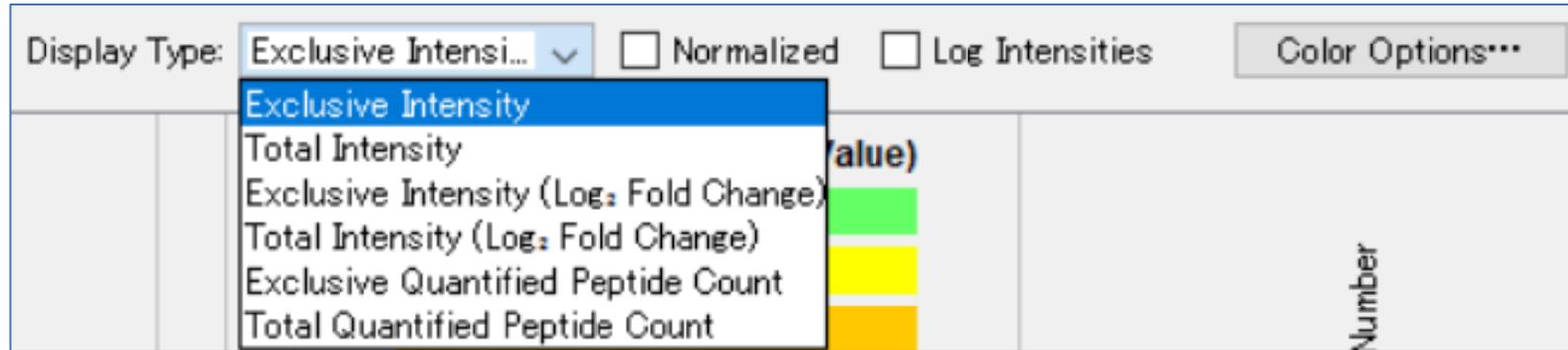
• ペプチド

- encyclopeDIA スコア評価
- peptide FDR 設定値 (percolator も使用)を満たす

• タンパク質

- 同定ペプチド N(デフォルト2) 以上アサインされている
- protein FDR 設定値を満たす
- clustering ルールでグループ化

表示される「定量値」 [Display Type]



Exclusive : ユニークペプチドのみ
Total : シェアペプチド含む

• Exclusive/total Intensity

ペプチド定量値(フラグメントXICのピーク強度から算出) の和

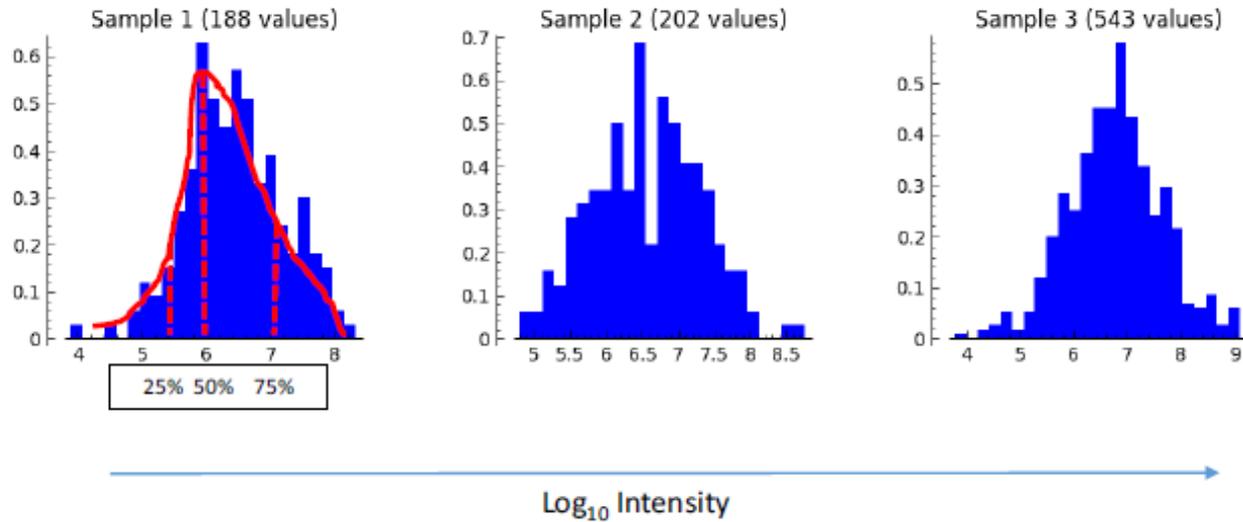
• Exclusive/total Intensity (Log₂ Fold Change)

ペプチド定量値について、Referenceとの比を取りその数字に対して2を底にするLogに変換

• Exclusive/total Quantified Peptide Count

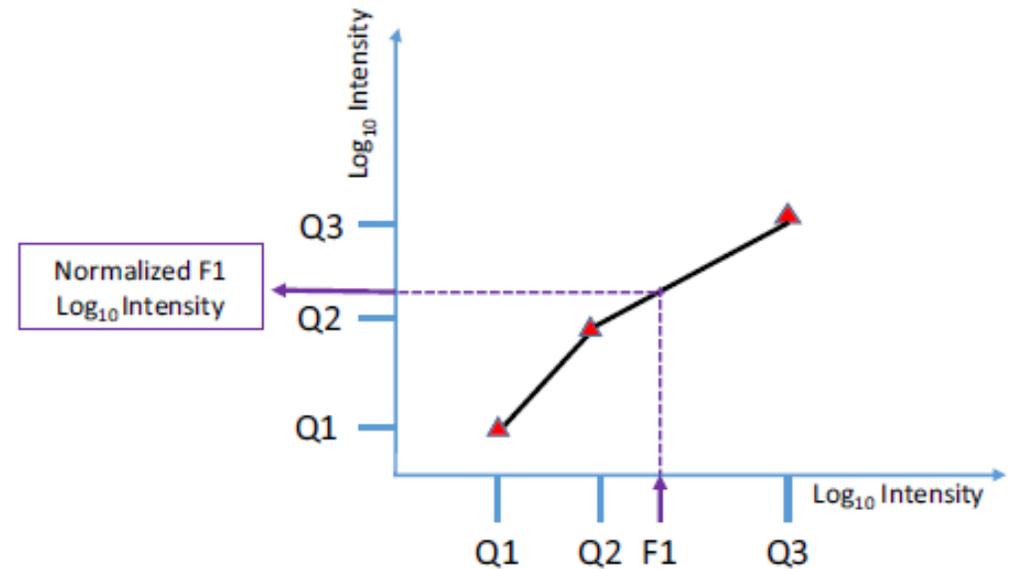
タンパク質あるいは表示グループにアサインされたペプチドで、かつ定量に使われたフラグメントを持つペプチドに限定し数え上げた数(標準化処理も実施)。

Normalization



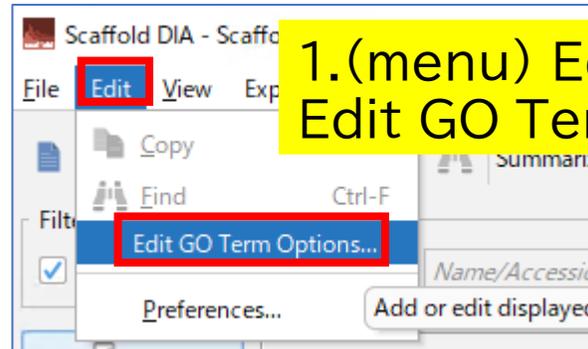
サンプル強度分布によるNormalization

- 全データ vs 各sample、それぞれ強度分布を作成
- 各サンプルにおいて、25%、50%、75%にあたる定量値を、全データ平均にそろえる
- 各四分位数の間は点間を結ぶ直線に従って変換する

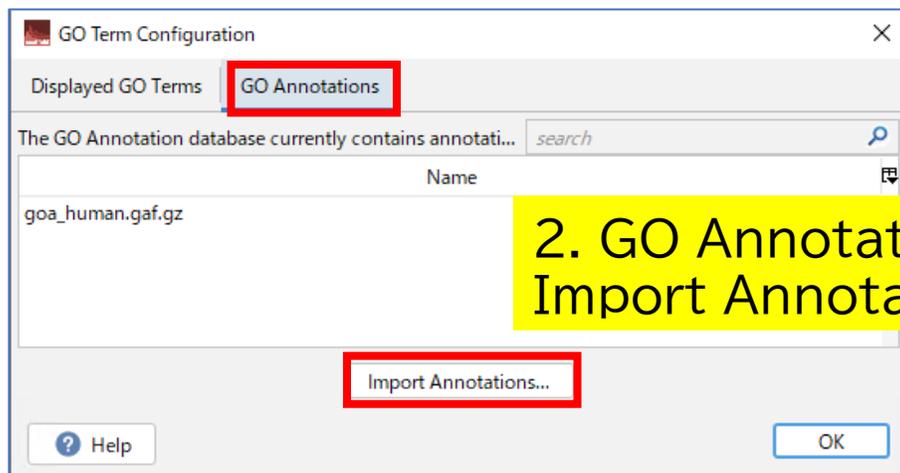


GOファイルのセット方法

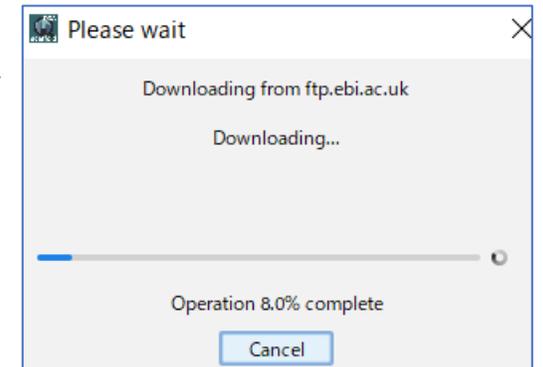
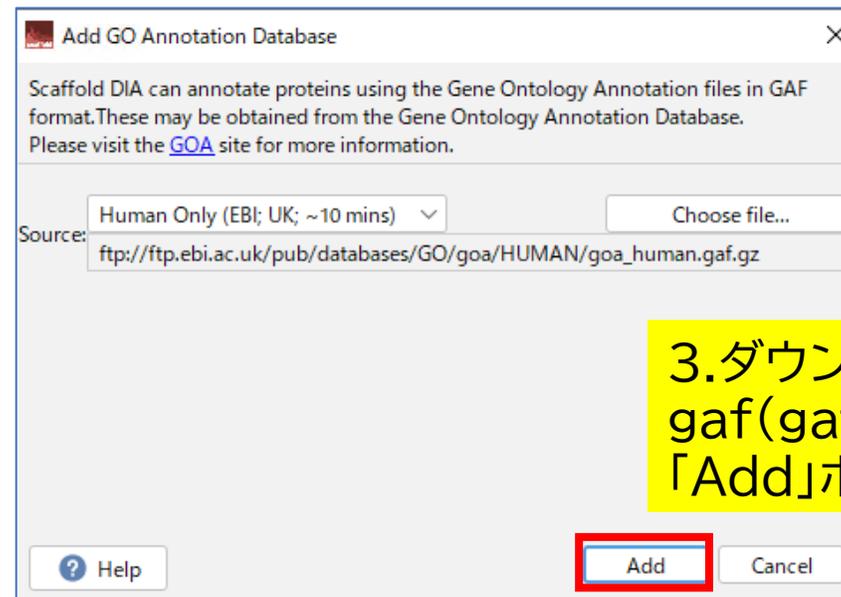
1. (menu) Edit ->
Edit GO Term Options



2. GO Annotations タブ、
Import Annotations

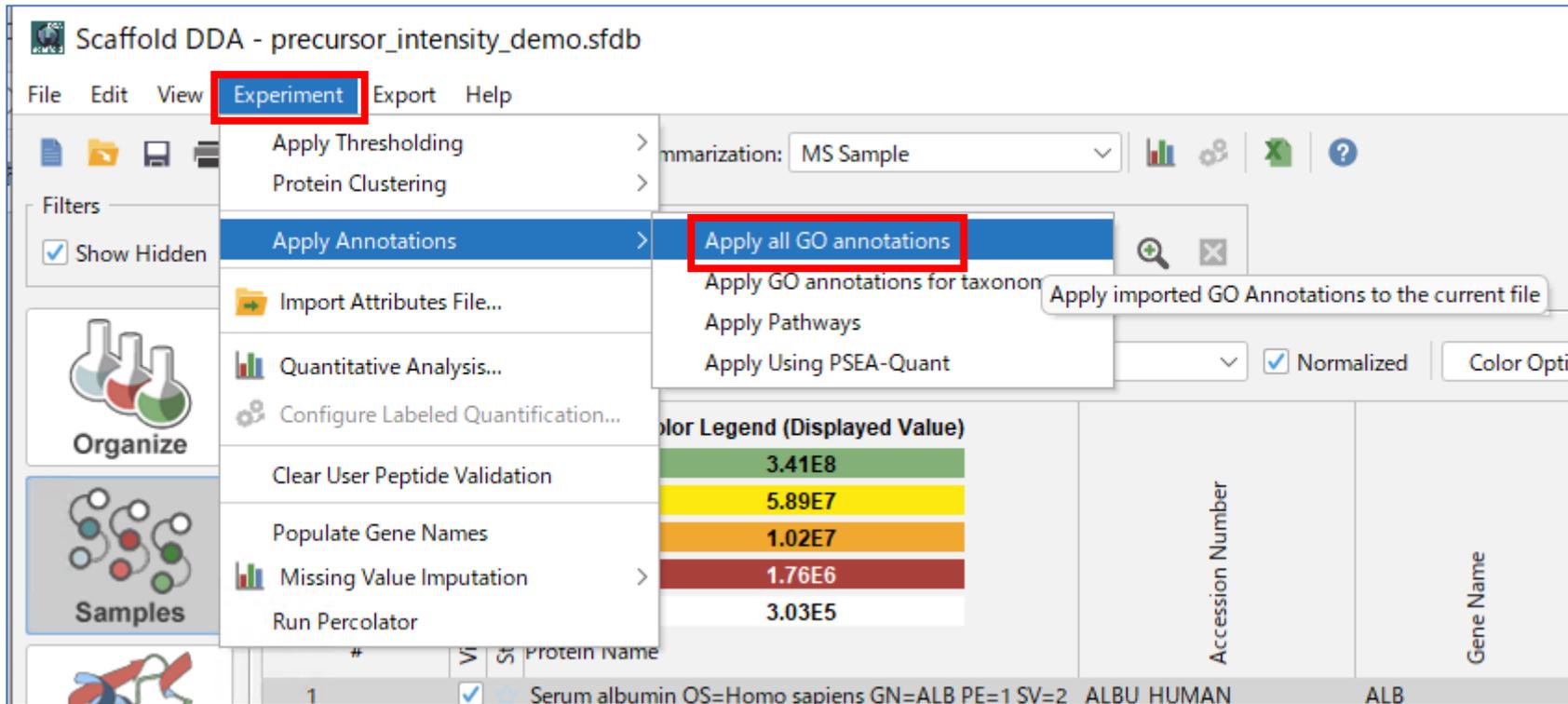


3. ダウンロード先または
gaf(gaf.gz)ファイルを選択、
「Add」ボタン



GO情報のannotation付与方法

- 1.(menu) Experiment
- > Apply Annotations
- > Apply all GO annotations



The screenshot shows the Scaffold DDA software interface. The 'Experiment' menu is open, and the 'Apply Annotations' option is selected. A sub-menu is displayed, with 'Apply all GO annotations' highlighted. A tooltip for this option reads: 'Apply imported GO Annotations to the current file'. The background shows a data table with columns for '#', 'Protein Name', 'Accession Number', and 'Gene Name'. The first row is highlighted in green and contains the following data:

#	Protein Name	Accession Number	Gene Name
1	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2	ALBU HUMAN	ALB

Scaffold DIA - Demo2_HeLa_insulin_6_files_30_proteins.sdia

File Edit View Experiment Export Tools Help

Summarization: Treatment Protein FDR 1.0% FDR Min # Peptides 2

Filters: Show Hidden Name/Accession p-value filter GO Term

Display Type: Exclusive Intensity Normalized Log Intensities Color Options...

Color Legend (Displayed Value): 6.36E8, 1.22E8, 2.33E7

Biological Process Cellular Component Molecular ...

Control Insulin

Proteins panel highlighted in red box.

タンパク質を選択した状態でダブルクリック、またはProteins パネルをクリックすると... →次頁

#	Visible	Star	Protein Name	Accession	MW	pI	Abundance	Score	Identified	Biological Process	Cellular Component	Molecular Function	Control	Insulin
1	<input checked="" type="checkbox"/>		sp P49792 RBP2_HUMAN Ribosomal protein L23 OS=Homo sapiens	sp P49792	24 kDa	4.5	100%	0.31	Homo sapiens	ribosome	ribosome	ribosome	2.45E8	2.24E8
2	<input checked="" type="checkbox"/>		sp P15924 DESP_HUMAN Desmin OS=Homo sapiens	sp P15924	48 kDa	5.0	100%	0.49	Homo sapiens	cytoskeleton	cytoskeleton	cytoskeleton	1.32E8	1.39E8
3	<input checked="" type="checkbox"/>		sp Q43707 ACTN4_HUMAN Alpha-actinin-4 OS=Homo sapiens	sp Q43707	103 kDa	5.0	63%	0.23	Homo sapiens	cytoskeleton	cytoskeleton	cytoskeleton	6.17E8	5.81E8
4	<input checked="" type="checkbox"/>		sp Q5T4S7 UBR4_HUMAN Ubiquitin-protein ligase E3 UBR4 OS=Homo sapiens	sp Q5T4S7	87 kDa	5.0	100%	0.090	Homo sapiens	ubiquitin-mediated proteolysis	cytoskeleton	ubiquitin-protein ligase	8.93E7	8.06E7
5	<input checked="" type="checkbox"/>		sp P11388 TOP2A_HUMAN DNA topoisomerase II alpha OS=Homo sapiens	sp P11388	174 kDa	8.5	100%	0.18	Homo sapiens	cell cycle	nucleus	topoisomerase activity	4.16E8	4.49E8
6	<input checked="" type="checkbox"/>		sp P18206 VINC_HUMAN Vinculin OS=Homo sapiens	sp P18206	124 kDa	4.5	100%	0.14	Homo sapiens	cell cycle	cytoskeleton	cell-cell junction	6.58E8	5.44E8
7	<input checked="" type="checkbox"/>		sp P12814 ACTN1_HUMAN Alpha-actinin-1 OS=Homo sapiens	sp P12814	103 kDa	5.0	63%	0.18	Homo sapiens	cytoskeleton	cytoskeleton	cytoskeleton	1.77E8	1.68E8
8	<input checked="" type="checkbox"/>		sp Q9NR30 DDX21_HUMAN Nucleolar RNA helicase 2 OS=Homo sapiens	sp Q9NR30	87 kDa	5.0	100%	0.14	Homo sapiens	ribosome	nucleolus	ribosome	5.71E8	5.91E8
9	<input checked="" type="checkbox"/>		sp Q9P8F0 PPP1R1_HUMAN Protein phosphatase 1 regulatory subunit 1B OS=Homo sapiens	sp Q9P8F0	152 kDa	5.0	100%	0.14	Homo sapiens	cell cycle	cytoskeleton	phosphatase activity	1.22E8	1.04E8
10	<input checked="" type="checkbox"/>		sp Q86UP2 KTN1_HUMAN Kinectin OS=Homo sapiens	sp Q86UP2	156 kDa	4.5	100%	0.18	Homo sapiens	cytoskeleton	cytoskeleton	cytoskeleton	1.24E8	1.09E8
11	<input checked="" type="checkbox"/>		sp P14631 KIF11_HUMAN Kinesin family class I member 1 OS=Homo sapiens	sp P14631	146 kDa	1.0	100%	0.14	Homo sapiens	cell cycle	cytoskeleton	cytoskeleton	2.66E8	2.42E8
12	<input checked="" type="checkbox"/>		sp P34932 HSP74_HUMAN Heat shock 70 kDa protein OS=Homo sapiens	sp P34932	94 kDa	5.0	100%	0.65	Homo sapiens	cell cycle	cytoskeleton	chaperone activity	5.61E8	5.55E8
13	<input checked="" type="checkbox"/>		sp P26639 SYTC_HUMAN Threonine--tRNA ligase, cytosolic OS=Homo sapiens	sp P26639	83 kDa	5.0	100%	0.51	Homo sapiens	translation	cytoskeleton	aminoacyl-tRNA synthetase activity	8.78E8	9.20E8
14	<input checked="" type="checkbox"/>		sp Q14152 EIF3A_HUMAN Eukaryotic translation initiation factor 3 subunit A OS=Homo sapiens	sp Q14152	167 kDa	5.0	100%	0.83	Homo sapiens	translation	cytoskeleton	translation	4.83E8	4.82E8
15	<input checked="" type="checkbox"/>		sp P22102 PUR2_HUMAN Trifunctional purine biosynthetic pathway regulator OS=Homo sapiens	sp P22102	108 kDa	5.0	100%	0.65	Homo sapiens	translation	cytoskeleton	translation	5.69E8	5.46E8
16	<input checked="" type="checkbox"/>		sp P50990 TCPO_HUMAN T-complex protein 1 subunit gamma OS=Homo sapiens	sp P50990	60 kDa	5.0	100%	0.25	Homo sapiens	cell cycle	cytoskeleton	cytoskeleton	1.20E9	1.27E9
17	<input checked="" type="checkbox"/>		sp P78371 TCPB_HUMAN T-complex protein 1 subunit beta OS=Homo sapiens	sp P78371	57 kDa	5.0	100%	0.39	Homo sapiens	cell cycle	cytoskeleton	cytoskeleton	7.59E8	8.82E8
18	<input checked="" type="checkbox"/>		sp Q04637 IF4G1_HUMAN Eukaryotic translation initiation factor 4 gamma OS=Homo sapiens	sp Q04637	175 kDa	5.0	100%	0.51	Homo sapiens	translation	cytoskeleton	translation	4.85E8	4.79E8
19	<input checked="" type="checkbox"/>		sp P14625 ENPL_HUMAN Endoplasmic reticulum protein OS=Homo sapiens	sp P14625	92 kDa	5.0	100%	0.89	Homo sapiens	cell cycle	cytoskeleton	cytoskeleton	1.19E9	1.17E9
20	<input checked="" type="checkbox"/>		sp P08729 K2C7_HUMAN Keratin, type II cytoskeletal class I OS=Homo sapiens	sp P08729	51 kDa	5.0	100%	0.61	Homo sapiens	cytoskeleton	cytoskeleton	cytoskeleton	2.17E9	2.18E9
21	<input checked="" type="checkbox"/>		sp P00558 PGK1_HUMAN Phosphoglycerate kinase 1 OS=Homo sapiens	sp P00558	45 kDa	5.0	100%	0.52	Homo sapiens	cell cycle	cytoskeleton	kinase activity	4.58E9	4.82E9
22	<input checked="" type="checkbox"/>		sp P06733 ENO4_HUMAN Alpha-enolase OS=Homo sapiens	sp P06733	47 kDa	5.0	100%	0.13	Homo sapiens	cell cycle	cytoskeleton	kinase activity	7.42E9	1.01E10
23	<input checked="" type="checkbox"/>		sp Q5UIP0 RIF1_HUMAN Telomere-associated protein OS=Homo sapiens	sp Q5UIP0	274 kDa	5.0	100%	0.71	Homo sapiens	cell cycle	cytoskeleton	telomerase activity	3.07E7	2.96E7

Proteins
0.0% FDR (attained)
29 Targets
0 Decoys

Peptides
0.0% FDR (attained)
986 Targets
0 Decoys

Proteins 画面

ペプチド/タンパク質の
定性/定量 結果を
より詳細に検証

Protein Similar Proteins

Showing All Peptides | ... | sp|Q5T4S7|UBR4_HUMAN

→ ①/②スライド 20

Quantifi...	Peptide Sequence	Quantified Mat...	Fixed Modifications	Variable Modificati...	Start	Stop	Protein Accessi...	Proba...	Mass
<input type="checkbox"/>	AAPPPPPPPPLESSPR	0 of 6			606	622	sp Q5T4S7 UBR4...	86%	1,702.904
<input checked="" type="checkbox"/>	AEHASSLLELASTTK	0 of 6			1065	1079	sp Q5T4S7 UBR4...	100%	1,556.805
<input type="checkbox"/>	ALGTLGTTNEK	0 of 6			4803	4814	sp Q5T4S7 UBR4...		
<input type="checkbox"/>	APSYIEIFGR	0 of 6			2364	2373	sp Q5T4S7 UBR4...		
<input checked="" type="checkbox"/>	AQQALSELHTEK	2 of 6			1828	1850	sp Q5T4S7 UBR4...	100%	1,452.757

① 同定ペプチド

Sequence	Modifications	Charge	Sample	Quant. Intensity	# Qua...	RT Start (min)	RT Center (min)	RT Stop (min)	Precur...
AEHASSLLELASTTK		2	Control_1	-	0	-	-	-	779.410
AEHASSLLELASTTK		2	Control_2	-	0	-	-	-	779.410
AEHASSLLELASTTK		2	Control_3	-	0	-	-	-	
AEHASSLLELASTTK		2	Insulin_1	-	0	-	-	-	

② 同定スペクトル

Protein Sequence | Protein Level Charts | Chromatograms | Fragment Intensities | Fragmentation Table

sp|Q5T4S7|UBR4_HUMAN E3 ubiquitin-protein ligase UBR4 OS=Homo sapiens GN=UBR4 PE=1 SV=1

→ ③ スライド 21

Q H N L L S P P F G W A S G S Q D S N S R R A T T P L Y

H F S S D A V P H P R F Y C V L S P E A S E D D L N R L D S V A C D V L F S K L 96

V K Y D E L Y A A L T A L L A A G S Q L D T V R R K E N K N V T A L E A C A

Y Y F L I L W R I L G I L P P S K **T Y I N Q L S M N S P E M S E C I L H T L R** 1040

756 out of 5183 (14.59%) amino acids identified with 23 modifications

③ 関連グラフ

Organize

Samples

Proteins

Visualize

Analysis

Publish

Proteins
0.0% FDR (attained)
29 Targets
0 Decoys

Peptides
0.0% FDR (attained)
1002 Targets
0 Decoys

①

同定ペプチド

Quantified	Peptide Sequence	Quantified Matches	Fixed Modif...	Variable Mo...	Start	Stop	Protein Accessi...	Proba...	Mass
<input type="checkbox"/>	AAPPPPPPPPLESSPR	0 of 6			606	622	sp Q5T4S7 UBR4...	86%	1,702.904
<input type="checkbox"/>	AEHASSLLELASTTK	0 of 6			1065	1079	sp Q5T4S7 UBR4...	100%	1,556.805
<input type="checkbox"/>	ALGTLGTTNEK	0 of 6			4803	4814	sp Q5T4S7 UBR4...	99%	1,234.623
<input type="checkbox"/>	APSYIEIFGR	0 of 6			2364	2373	sp Q5T4S7 UBR4...	98%	1,151.598
<input checked="" type="checkbox"/>	AQQALSELHVEK	0 of 6			1888	1898	sp Q5T4S7 UBR4...	100%	1,452.757
<input type="checkbox"/>									485.677
<input checked="" type="checkbox"/>									620.342
<input checked="" type="checkbox"/>									621.247

Quantified

定量計算に使用される条件(一定数以上のフラグメントの検出など)を満たしたタンパク質の定量計算に使用されているかどうか。
手動でチェックを外すと定量計算から外される。

Quantified Matches

全サンプルに対して定量計算に使用されたサンプルがいくつあるのか

②

同定スペクトル

Sequence	Modifications	Charge	Sample	Quant. Intensity	# Qua...	RT Start (min)	RT Center (min)	RT Stop (min)	Precursor MZ	Attribute
SALQYDTLISLMEHL...		3	Control_1	2.72E5	5	68.91	69.16	69.36	874.455	control
SALQYDTLISLMEHL...		3	Control_2	—	0	68.55	—	69.32	874.455	control
ASVVTASSGSALQYDTLISLMEHL...		3	Control_3	4.455E5	5	68.65	68.89	69.13	874.455	control
ASVVTASSGSALQYDTLISLMEHL...		3	Control_2	3.8E5	5	68.97	69.21	69.41	874.455	insulin

Quant. Intensity

定量計算に利用したフラグメントピーク強度の和

of Quant. fragment

定量計算に利用したフラグメントピーク数

RT Start, RT Center, RT Stop

該当ペプチドを検出し定量計算に利用したRTの開始から終了までの時間

全長とcoverage、修飾残基

③
関連グラフ

Protein Sequence | Protein Level Charts | Chromatograms | Fragment Intensities | Fragmentation Table

sp|Q5T4S7|UBR4_HUMAN E3 ubiquitin-protein ligase UBR4 OS=Homo sapiens GN=UBR4 PE=1 SV=1

Q H N L L S P P F G W A S G S Q D S N S R R A T T P L Y H G F K E V E E N W S K 920

H F S S D A V P H P R F Y C V L S P E A S E D D L N R L D S V A C D V L F S K L 960

V K Y D E L Y A A L T A L L A A G S Q L D T V R R K E N K N V T A L E A C A L Q 1000

Y Y F L I L W R I L G I L P P S K **T Y I** **N Q L S M N S P E M** **S E C D I L H T L R** 1040

756 out of 5183 (14.59%) amino acids identified with 23 modifications

ペプチド/フラグメントのクロマトグラム

Showing All Peptides | ... | sp|Q13423|INN

Peptide Sequence	Spectral Matches	Modifications	Start	Stop	Protein Accession	Proba...	Mass
I I L D A L Q A K	3	C3 Carbamidomethyl...	1071	1079	sp Q13423 NNTM...	100%	1,033.523
I K T T V L A M D Q V P R	3		171	182	sp Q13423 NNTM...	100%	1,357.739
Q G F N W W E S G A G E A S K	3		85	100	sp Q13423 NNTM...	100%	1,577.769
S L G A E P L E V D L K	3		268	279	sp Q13423 NNTM...	98%	1,269.682

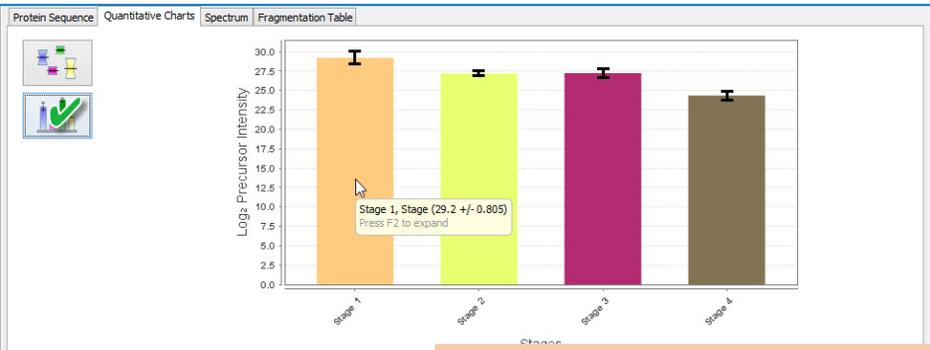
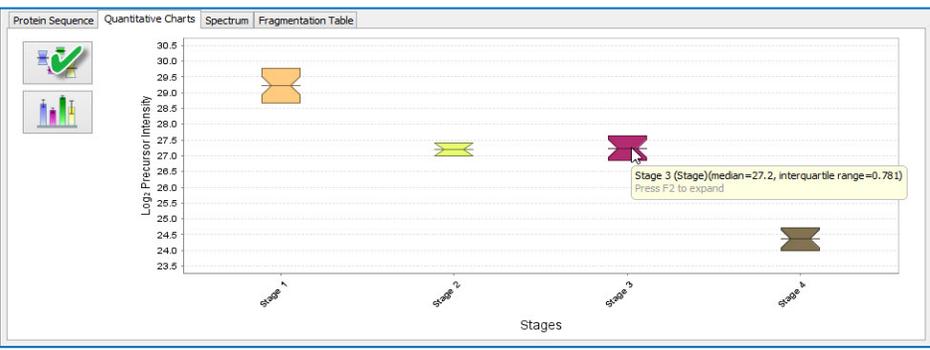
Valid	Sequence	Modifications	Charge	Sample	Quant. Intensity	# Qu...	RT Start (min)	RT Center (min)	RT Stop (min)	Precu...	Attrib...
<input checked="" type="checkbox"/>	Q G F N W W E S G A G E A S K	2	2	20170430_HeLa_DIA_control_2...	6.961E6	5	41.04	41.37	41.63	789.892	Control
<input checked="" type="checkbox"/>	Q G F N W W E S G A G E A S K	2	2	20170430_HeLa_DIA_insulin_1...	7.946E6	5	41.53	41.87	42.09	789.892	Insulin
<input checked="" type="checkbox"/>	Q G F N W W E S G A G E A S K	2	2	20170430_HeLa_DIA_insulin_ra...	8.576E6	5	41.55	41.88	42.14	789.892	Rapa...

Protein Sequence | Quantitative Charts | Chromatograms | Fragmentation Table

Precursor

Fragments

定量にはPrecursorのピーク強度情報を使っていない



タンパク質の定量値

Visualize 画面

有意な変動をしているタンパク質の解析 & データのバラつきをはじめとした確からしさを検証するグラフを提供

Scaffold DIA Viewer - Demo2_HeLa_insulin_6_files_30_proteins.sdia

File Edit View Experiment Export Help

Summarization: Condition Protein FDR 1.0% FDR Min # Peptides 2

Filters: Show Hidden Name/Accession p-value filter GO Term

Quantitation: Principal Component Analysis (Log₁₀ Intensity)

Volcano plot

Significance Marker: Significance Marker 0-Fold Change Marker

Multiselect Action: Zoom

Reference: control

Quantitative Scatterplot

contr... vs insulin

45° -line Regression Line

Multiselect Action: Zoom

Biological Process

Alphabetical 3D Values Legend

GO情報

Proteins: 0.0% FDR (attained), 29 Targets, 0 Decoys

Peptides: 0.0% FDR (attained), 1002 Targets, 0 Decoys

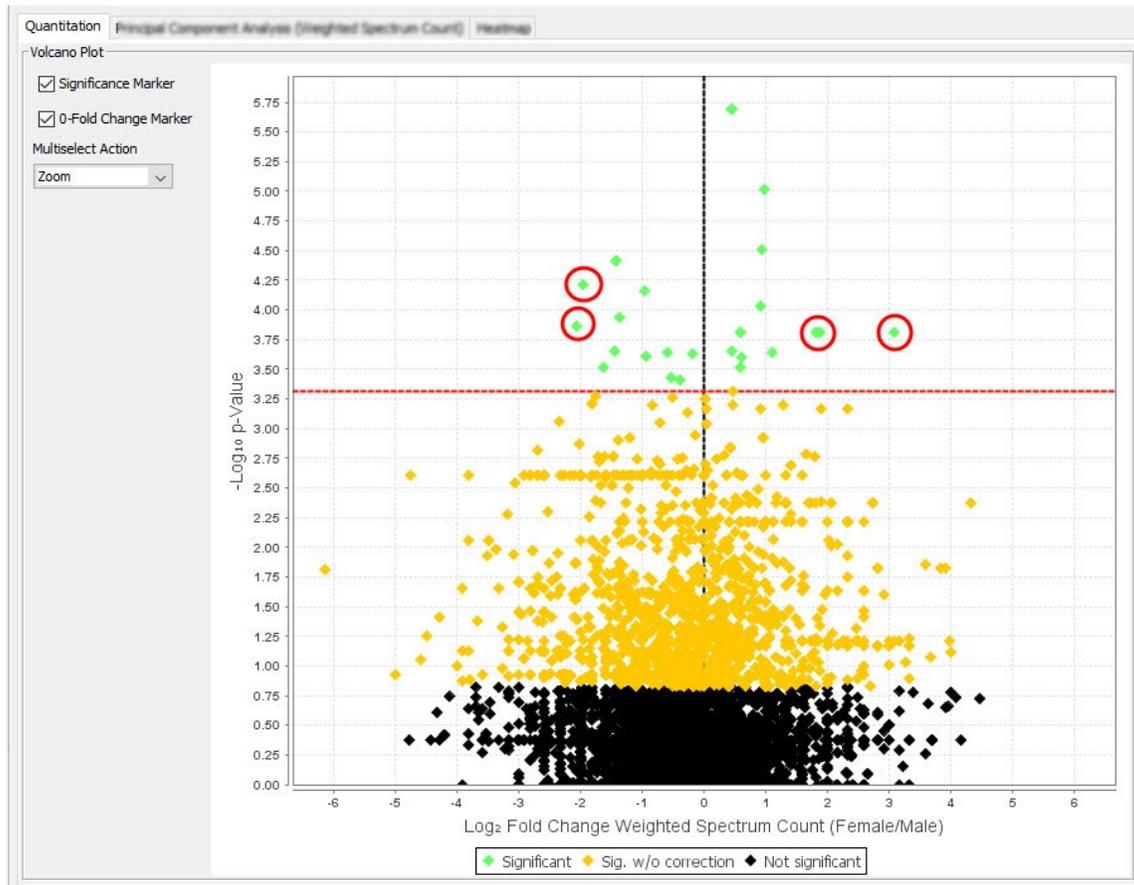
CV vs Mean Analyte Intensity



Volcano plot

Scatter plot

GO情報



Volcano plot

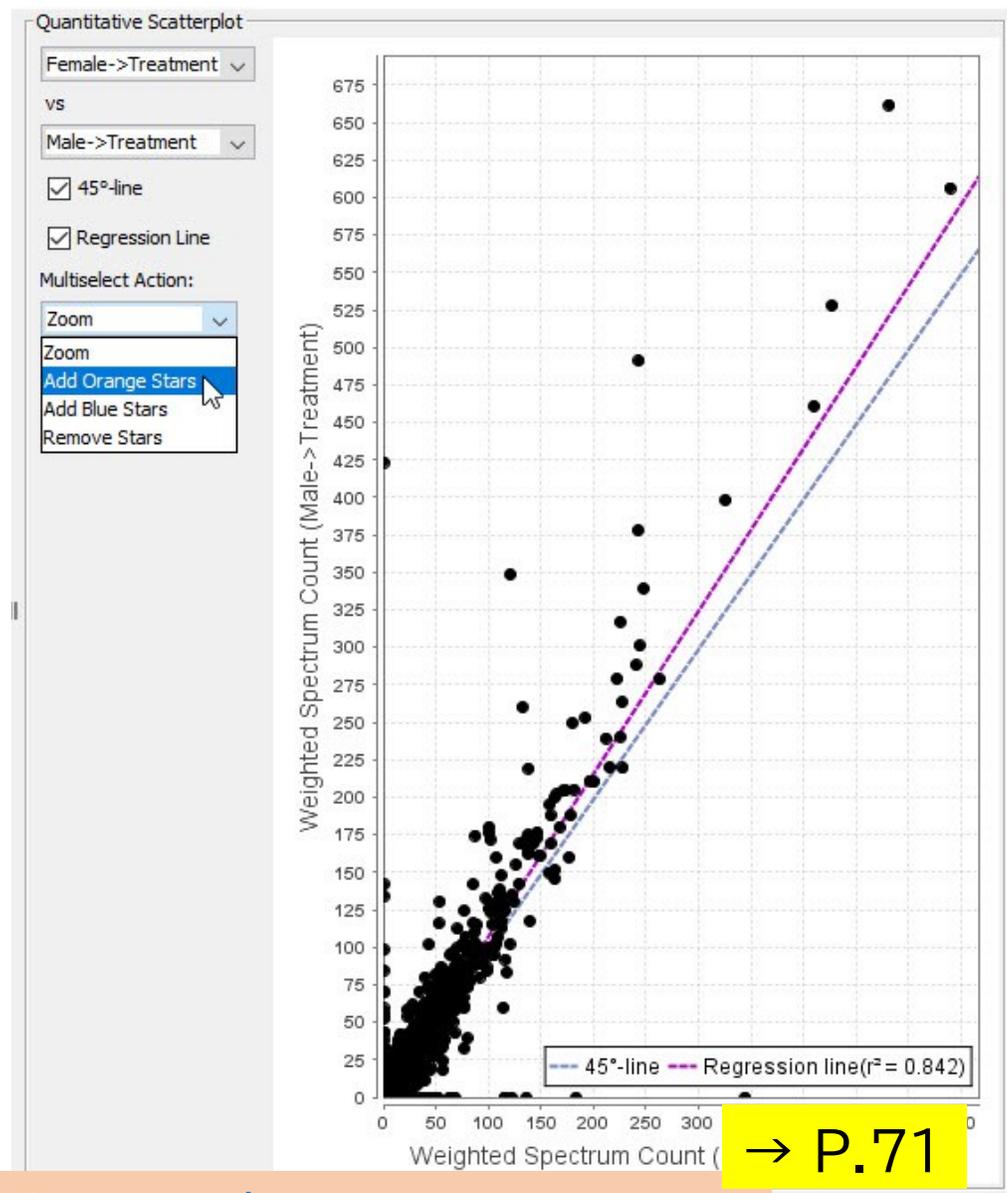
縦軸 $-\text{Log}_{10}(p)$

横軸 $\text{Log}_2(\text{Fold Change})$

赤点線 多重検定検証の $p(q, \alpha)$

変動タンパク質をratio, p-value両面から探す

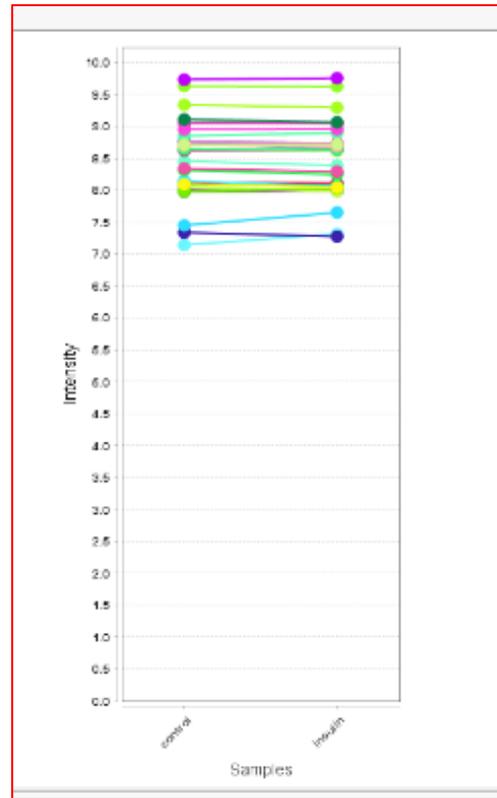
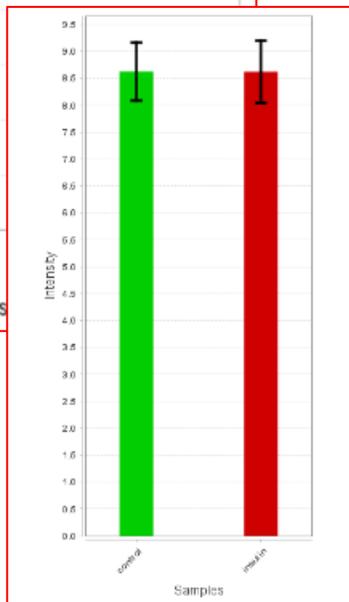
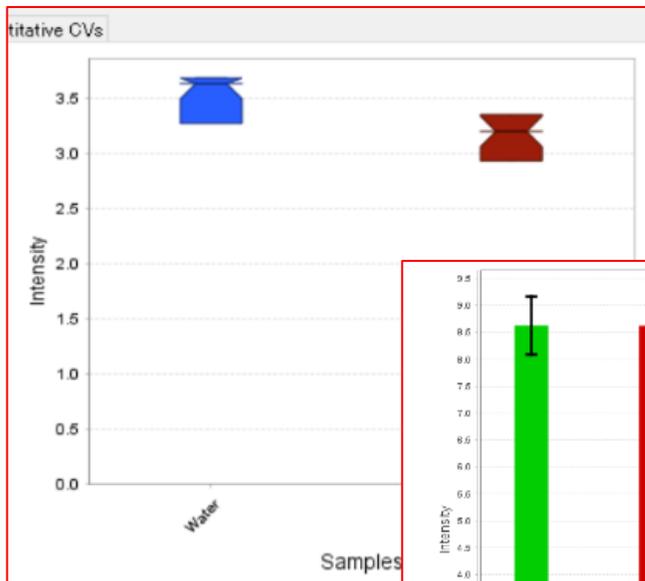
→ P.70



Scatterplot

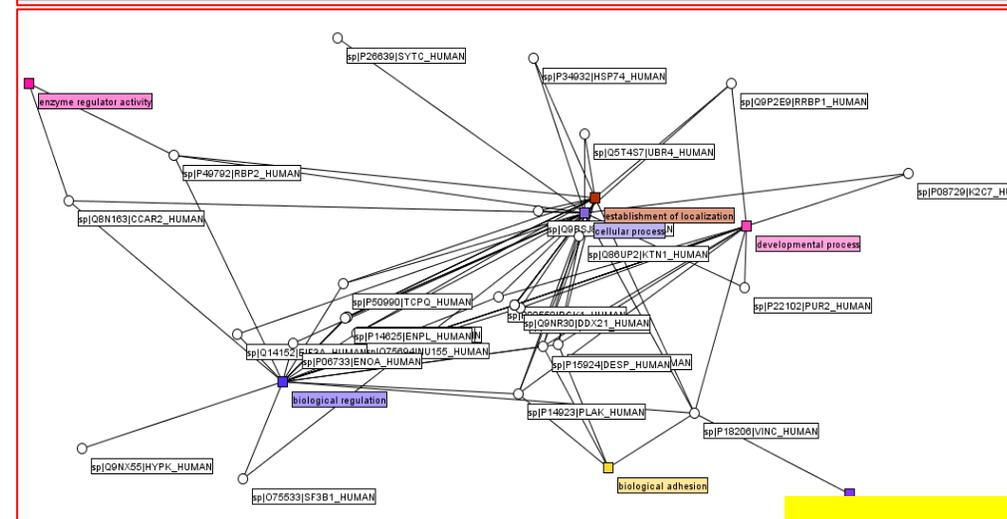
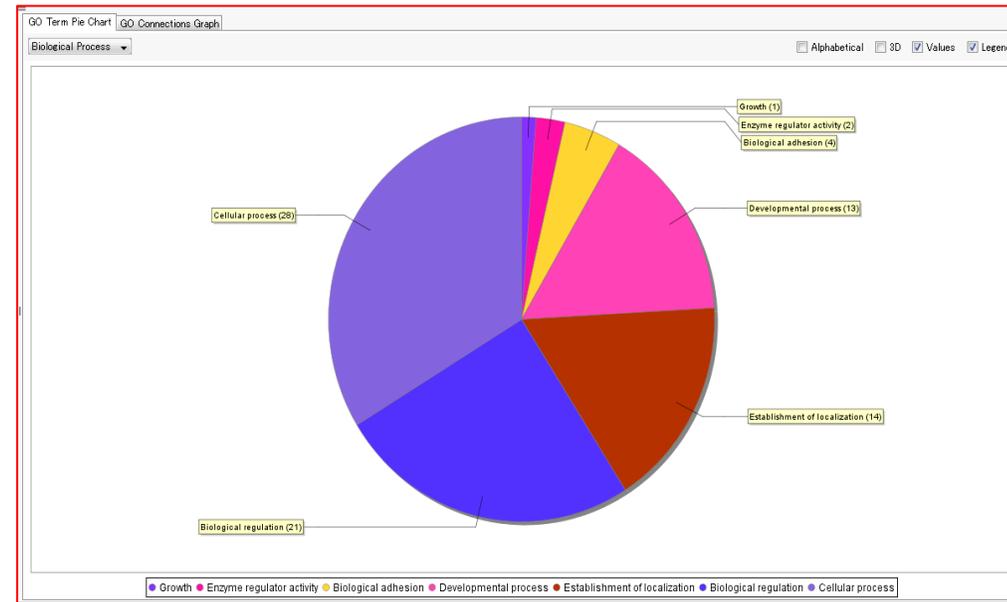
$y=x$ や回帰直線から大きく外れるタンパク質を探す

→ P.71



→ P.72

サンプル毎の定量値に関するグラフ
箱ひげ図(左)、棒グラフ(中)、Trend line(右)
サンプル毎の定量値ばらつきをチェックする

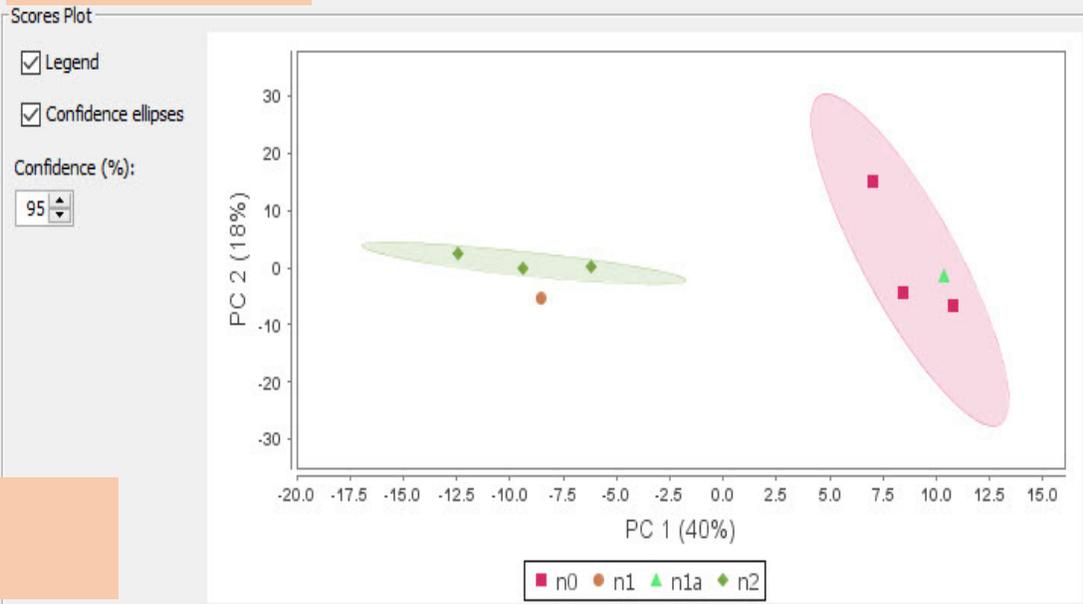


→ P.73

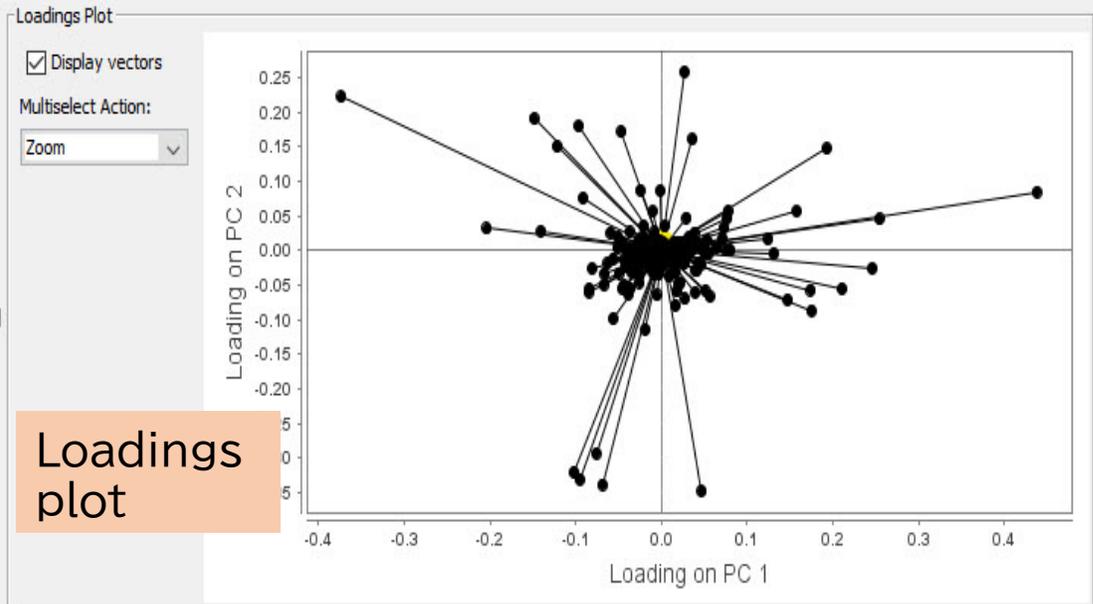
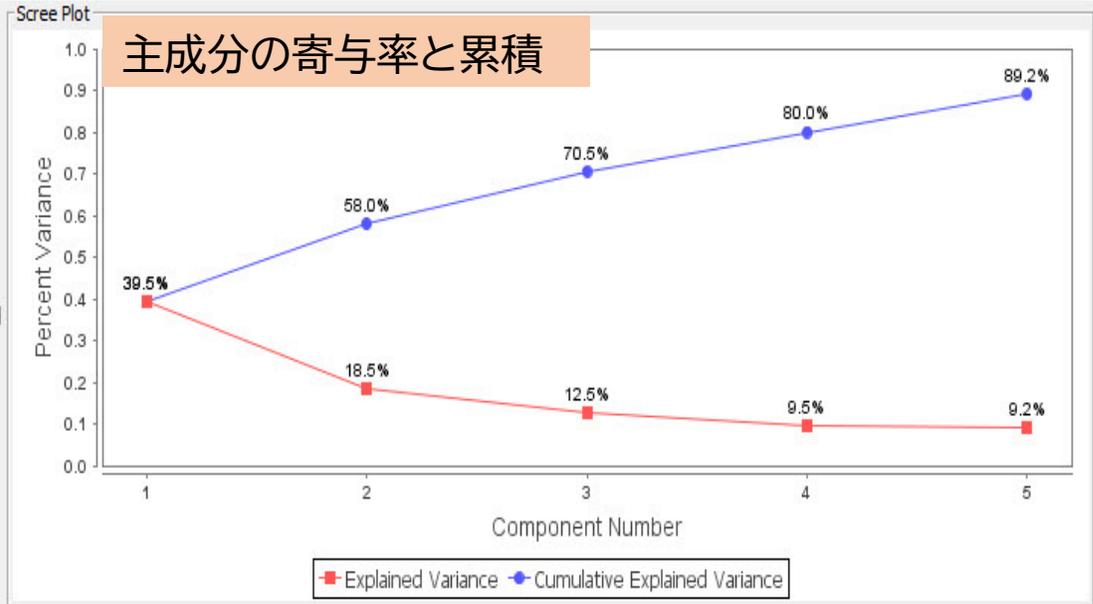
GO情報
円グラフ(上) ネットワーク図(下)



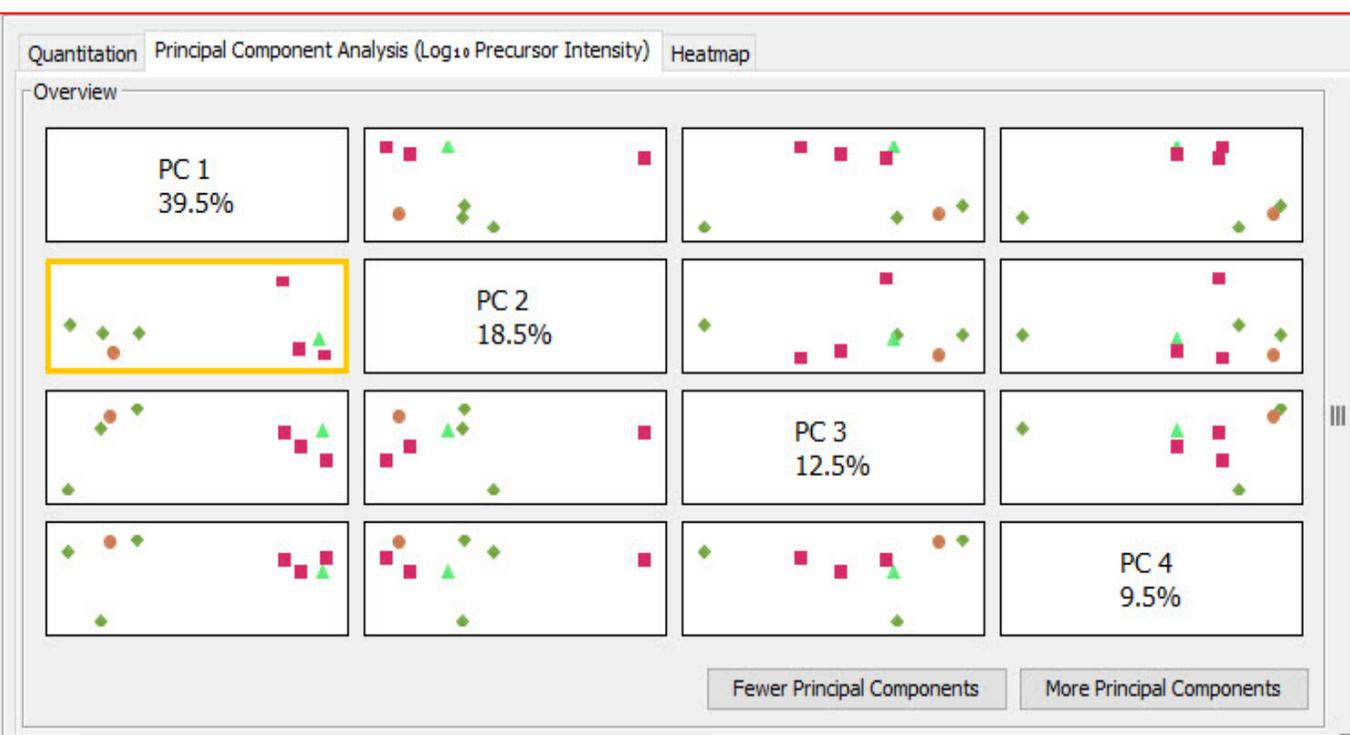
PCA Score Scatterplot



Score plot

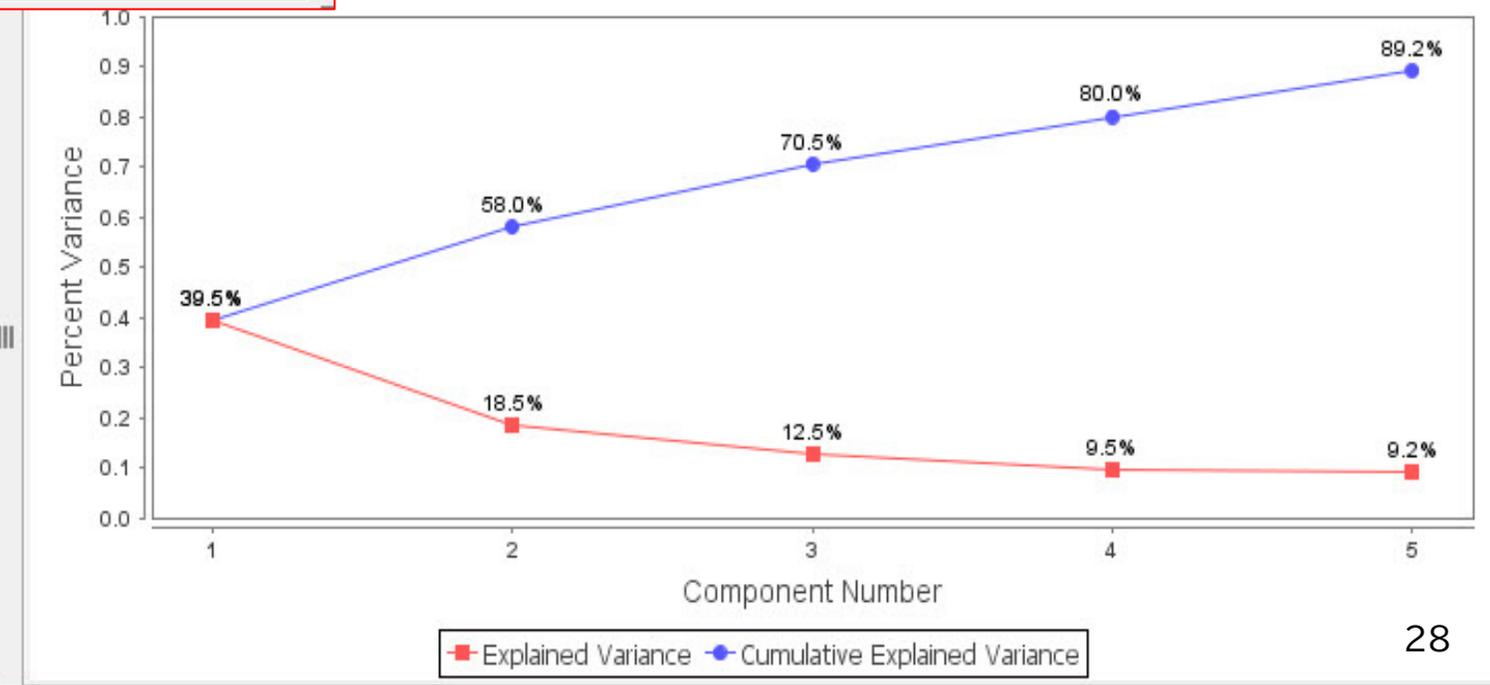


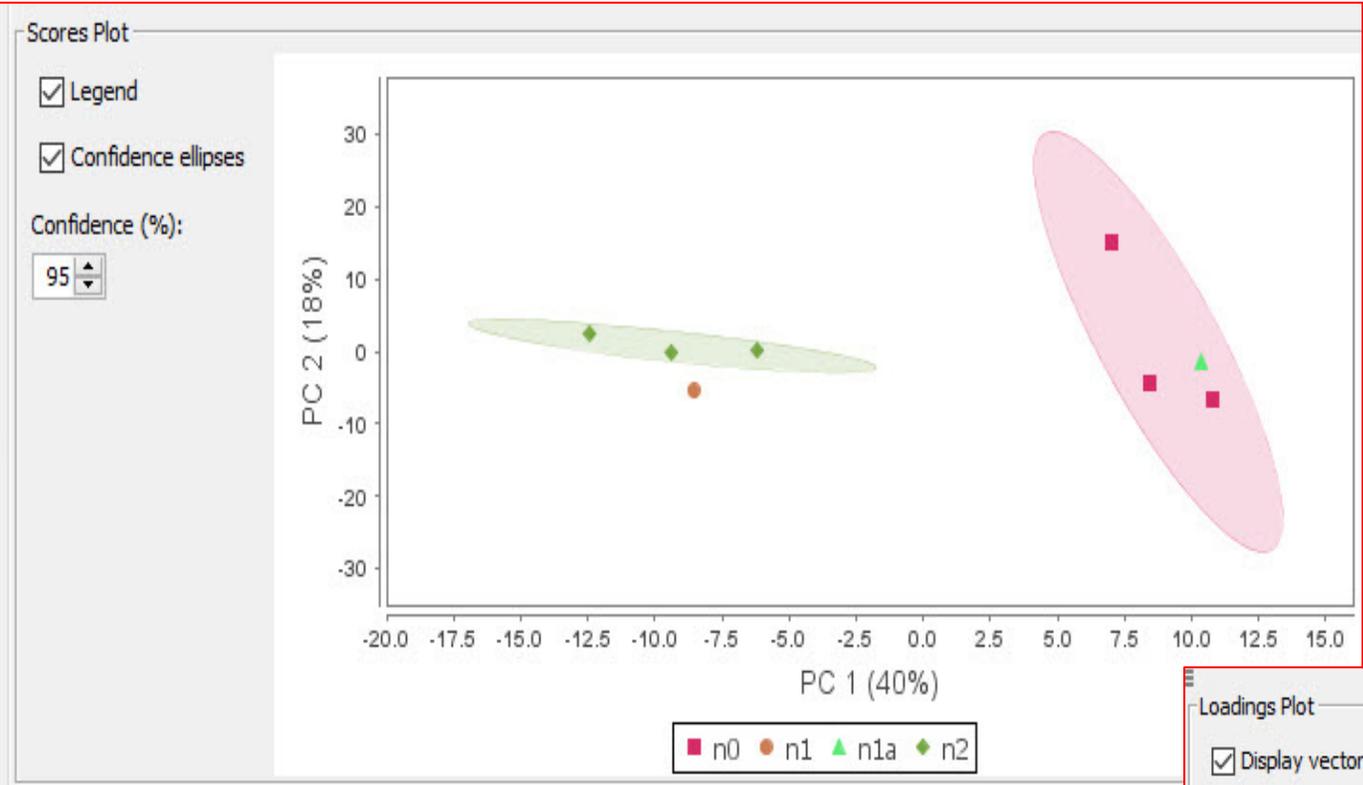
Loadings plot



各主成分の寄与率と累積

PCA score scatterplot
主成分別
プロットはsampleに該当

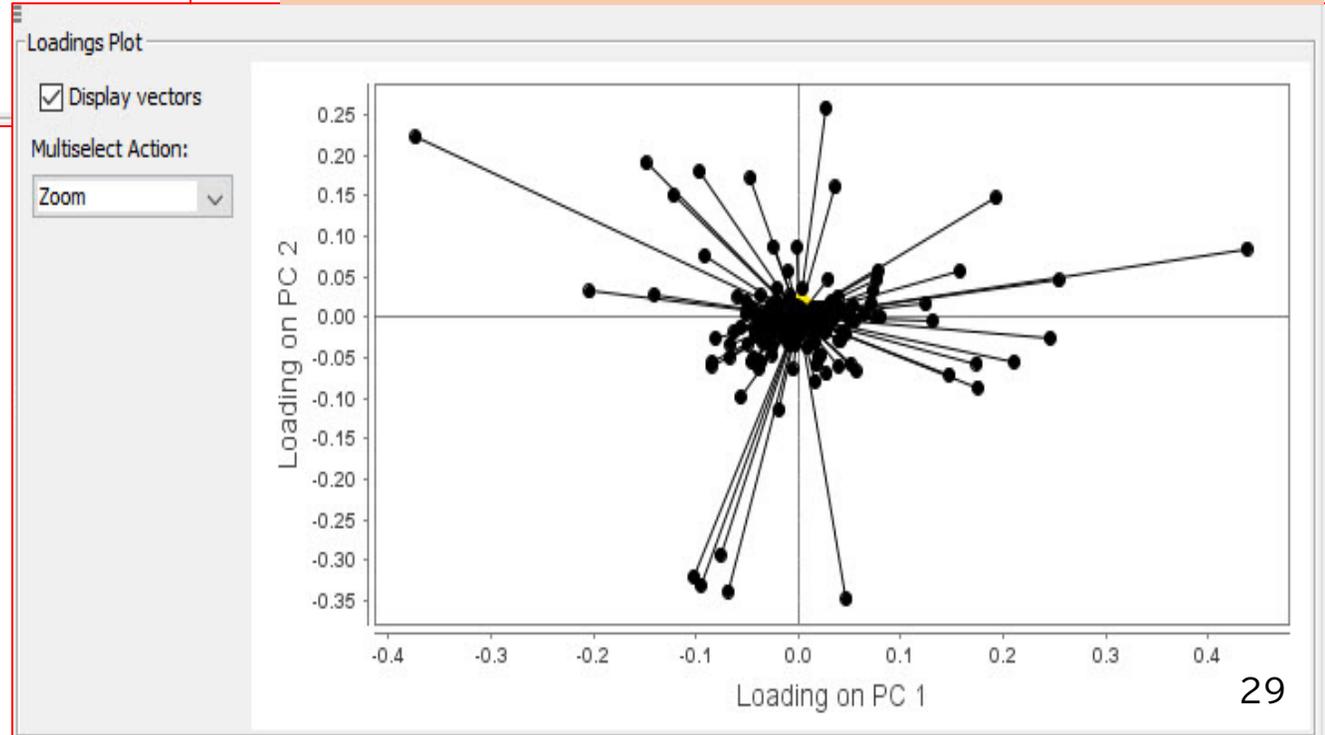




Loadings Plot

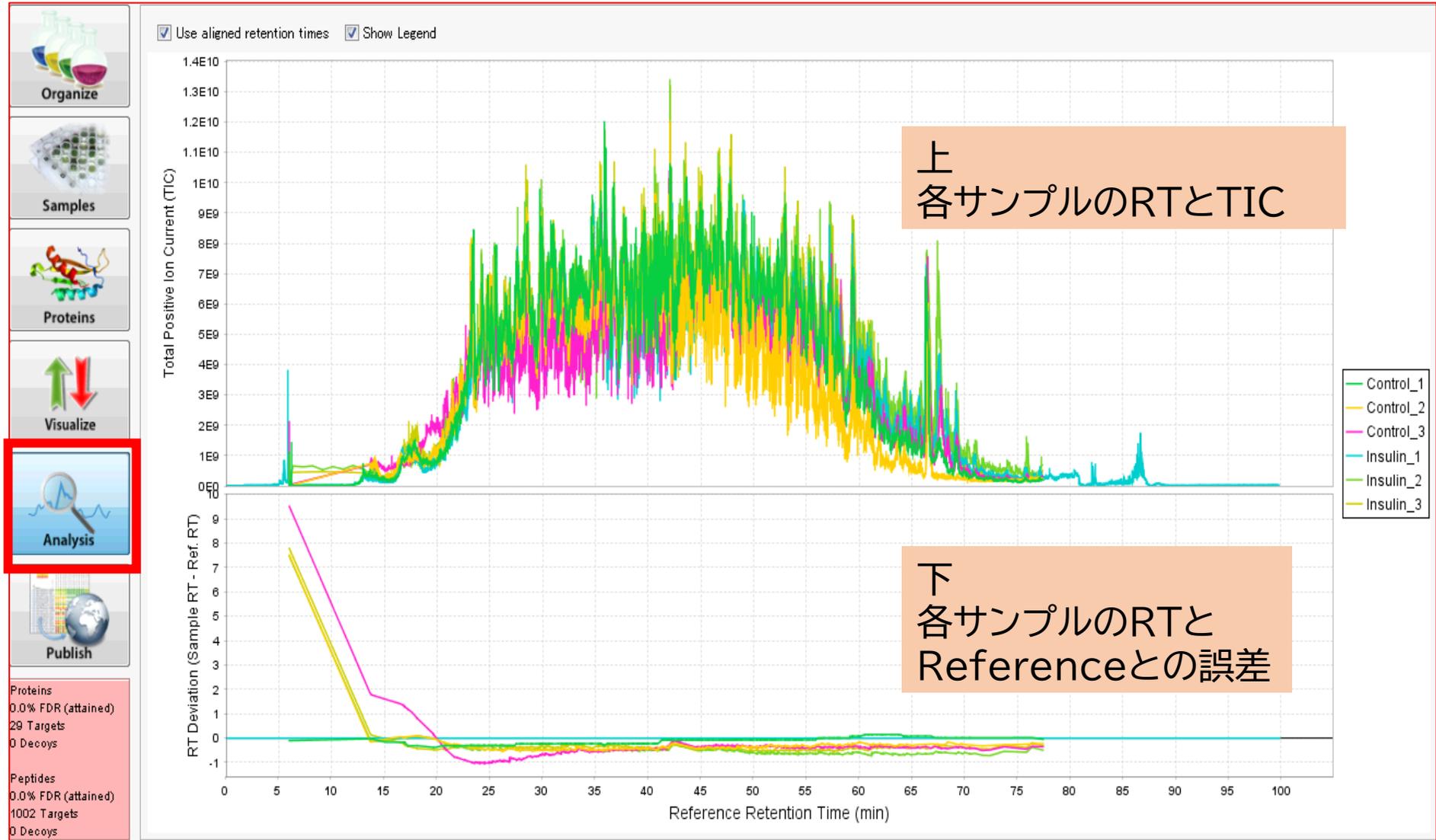
各タンパク質の主成分への寄与を表すプロット
左上で選択した主成分組み合わせと連動

Score plot
左上で選択した主成分組み合わせのスコア分布図
拡大版



Analysis 画面

TICとLC保持時間の
サンプル間誤差を確認
できる



Publish 画面

検索パラメータの確認、
論文のMethodのような
文章の作成

Experiment Methods SQL Report

Search

Search

- Search Library demo2elib
- Processing Directory #work
- Protein Sequence Database uniprot-swissprot-human.fasta
- Perform RT alignment between reference false
- Fragmentation CID
- Precursor Tolerance 10.0 ppm
- Fragment Tolerance 10.0 ppm
- Library Fragment Tolerance 10.0 ppm
- Peptide FDR Threshold 0.01
- Data Acquisition Type Overlapping DIA
- Digestion Enzyme Trypsin
- Peptide Length [6-30]
- Peptide Charge [2-8]
- Max Missed Cleavages 1
- Modifications Carbamidomethylation C 57.0214635 Non-termin...

Analysis

- Shared Evidence Clustering Perfect
- Target Protein FDR 0.01
- Minimum Number of Peptides 2
- Grouping Applied in Version 1.0.0
- Thresholding Applied in Version 1.0.0
- Clustering Applied in Version 1.0.0
- Quantify on Exclusive Peptides true

Advanced

- Precursor Window Size Deduced from file
- Minimum Number of Quant Ions 3
- Maximum Number of Quant Ions 5

Version

- Scaffold DIA 1.0.0
- Encyclopedia 0.6.12
- ProteoWizard 3.0.11748
- Percolator 3.01.nightly-13-655e4c7-dirty

パラメーター一覧

ANALYSIS OVERVIEW

DIA data were analyzed using Scaffold DIA (1.0.0).

RAW DATA PROCESSING

Raw data files were converted to mzML format using ProteoWizard (3.0.11748). Deconvolution of overlapping windows was performed.

SPECTRAL LIBRARY SEARCH

Analytic samples were aligned based on retention times and individually searched against *demo2.elib* with a peptide mass tolerance of 10.0 ppm and a fragment mass tolerance of 10.0 ppm. Fixed modifications considered were: Carbamidomethylation C. The digestion enzyme was assumed to be Trypsin with a maximum of 1 missed cleavage site(s) allowed. Only peptides with charges in the range [2-3] and length in the range [6-30] were considered. Peptides identified in each sample were filtered by Percolator (3.01.nightly-13-655e4c7-dirty) to achieve a maximum FDR of 0.01. Individual search results were combined and peptide identifications were assigned posterior error probabilities and filtered to an FDR threshold of 0.01 by Percolator (3.01.nightly-13-655e4c7-dirty).

QUANTIFICATION

Peptide quantification was performed by Encyclopedia (0.6.12). For each peptide, the 5 highest quality fragment ions were selected for quantitation. Only peptides exclusive to each protein or cluster were used for quantification.

CRITERIA FOR PROTEIN IDENTIFICATION

Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis were grouped to satisfy the principles of parsimony. Proteins with a minimum of 2 identified peptides were thresholded to achieve a protein FDR threshold of 1.0%.

GO ANNOTATION

Proteins were annotated with GO terms from: UniProt, InterPro, GO_Central, Reactome, GOC, HPA, Ensembl, IntAct, ParkinsonsUK-UCL, NTNU_SB, LIFEdb, FlyBase, BHF-UCL, HGNC, MGI, SYSCILIA_CCNET, CACAO, AgBase, PINC, ARUK-UCL, CAFA, MTBBASE, Alzheimers_University_of_Toronto, WormBase, GDB, SynGO-UCL, DFLAT, SGD, dictyBase and SynGO

CITATIONS

ProteoWizard

A cross-platform toolkit for mass spectrometry and proteomics.
Chambers, M.C., MacLean, B., Burke, R., Amode, D., Ruderman, D.L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T.A., Brusniak, M.-Y., Paulse, C., Creasy, D., Flashner, L., Kani, K., Moulding, C., Seymour, S.L., Nuwaysir, L.M., Lefebvre, B., Kuhlmann, F., Roark, J., Rainer, P., Detlev, S., Hemenway, T., Huhmer, A., Langridge, J., Connolly, B., Chadick, T., Holly, K., Eckels, J., Deutsch, E.W., Moritz, R.L., Katz, J.E., Agus, D.B., MacCoss, M., Tabb, D.L. & Mallick, P.
Nature Biotechnology 30, 918-920 (2012)
<http://www.nature.com/nbt/journal/v30/n10/full/nbt.2377.html>

Percolator

Semi-supervised learning for peptide identification from shotgun proteomics datasets
Lukac Koll, Jesse Costantini, Jason Weston, William Stafford Noble, and Michael J. MacCoss

Copy Text to Clipboard

Export Publish Report

Export Supplementary Data

Method文章

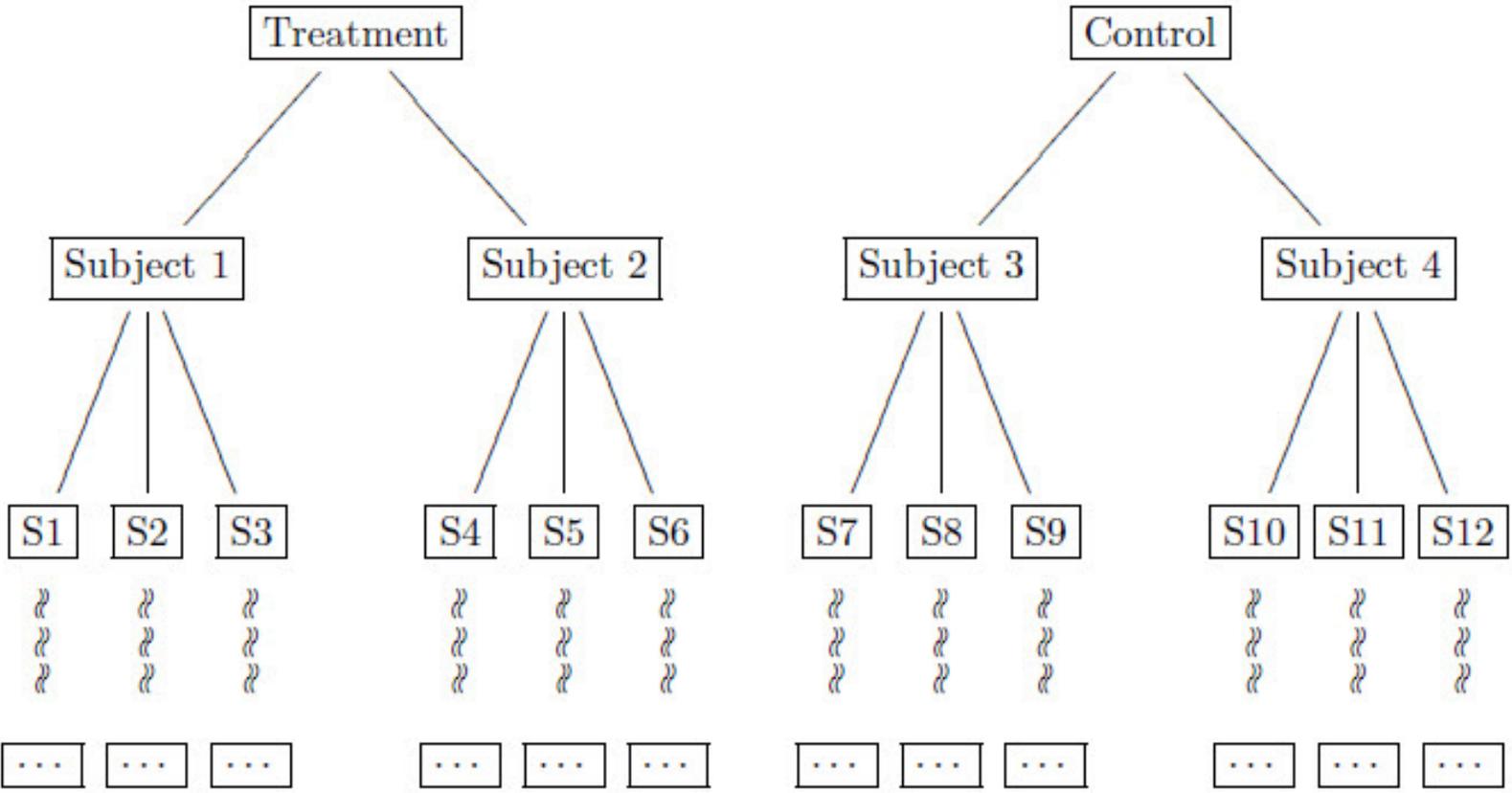
定量：階層構造化と属性付与の例

Biological Classes
↓ Experimental Variation ↓

Biological Replicates
(within biological classes)
↓ Biological Variation ↓

Technical Replicates
(within biological replicates)
Technical Variation

Fraction(s)
(within technical replicates)



eriment Export Help

Summarization: **MS Sample** Protein FDR: 1.0% FDR Min # Peptides: 2

Name/Accession p-value filter GO Term (712 of 826 filtered out)

Table

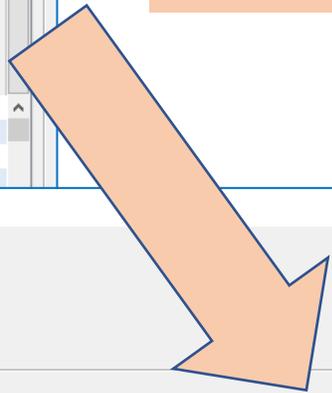
Display Type: Log₁₀ Intensity Normalized Color Options...

Color Legend (Displayed Value)

4.41
3.75
3.09
2.43
1.77

#	Visible	Star	Protein Name	Accession Number	Taxonomy	S1			S2		
						Rep1	Rep2	Rep3	Rep1	Rep2	Rep3
1	<input checked="" type="checkbox"/>		sp P00490 PHSM_ECOLI Maltodextrin phosphorylase OS=Escherichia coli (strain K12) GN=malP PE=1 SV=7	sp P00490 PHSM_ECOLI		2.21	2.31	2.33	2.90	2.75	2.74
2	<input checked="" type="checkbox"/>		sp P41250 SYG_HUMAN Glycine--tRNA ligase OS=Homo sapiens GN=GARS PE=1 SV=3	sp P41250 SYG_HUMAN	Homo sapiens	3.20	3.12	3.11	3.10	3.08	3.20
3	<input checked="" type="checkbox"/>		sp P0AFG6 ODO2_ECOLI Dihydrodipolysine-residue succinyltransferase component of 2-oxoglutarate dehydro...	sp P0AFG6 ODO2_ECOLI		2.17	2.35	2.39	2.56	2.66	2.65
4	<input checked="" type="checkbox"/>		sp P24752 THIL_HUMAN Acetyl-CoA acetyltransferase, mitochondrial OS=Homo sapiens GN=ACAT1 PE=1 SV=1	sp P24752 THIL_HUMAN	Homo sapiens	2.60	2.78	2.77	2.63	2.68	2.66

データをまとめる
単位を変更可能



eriment Export Help

Summarization: **Sample** Protein FDR: 1.0% FDR Min # Peptides: 2

Name/Accession p-value filter GO Term

Table

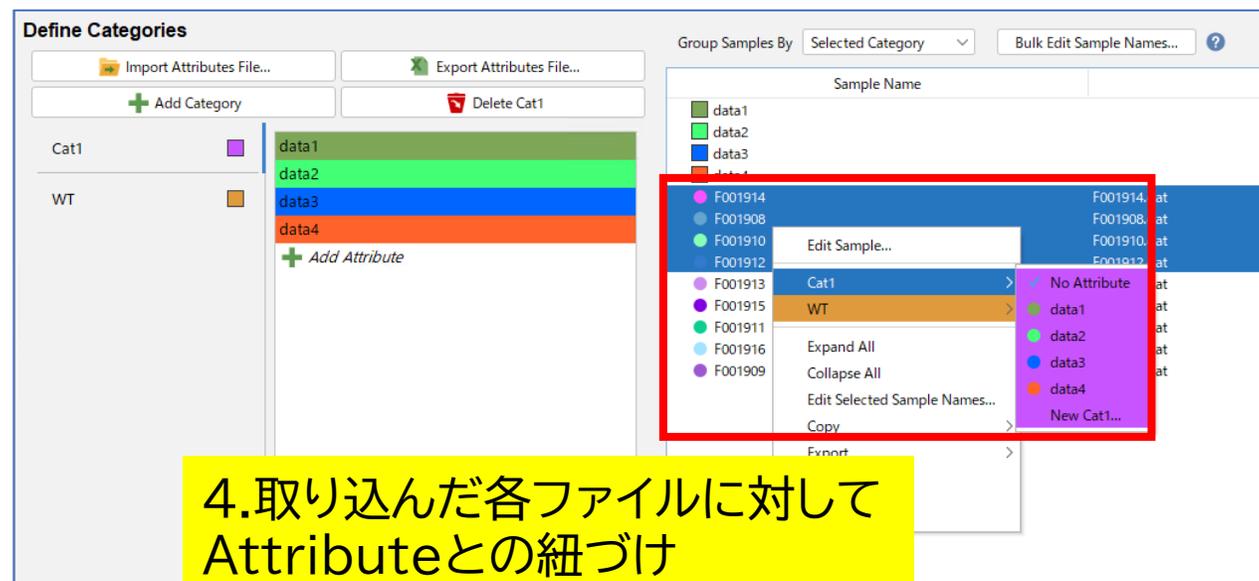
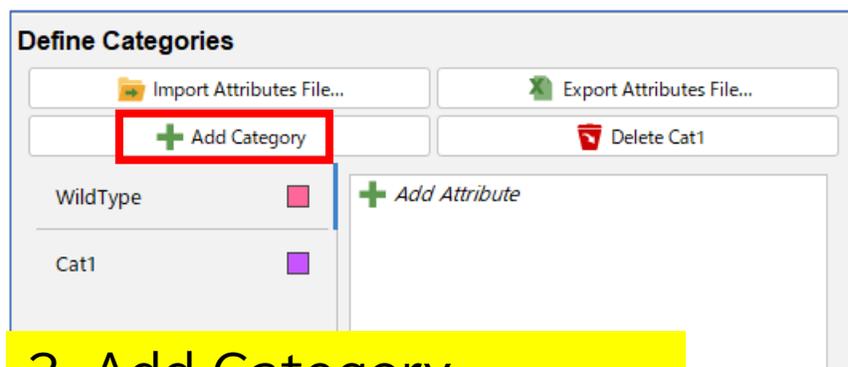
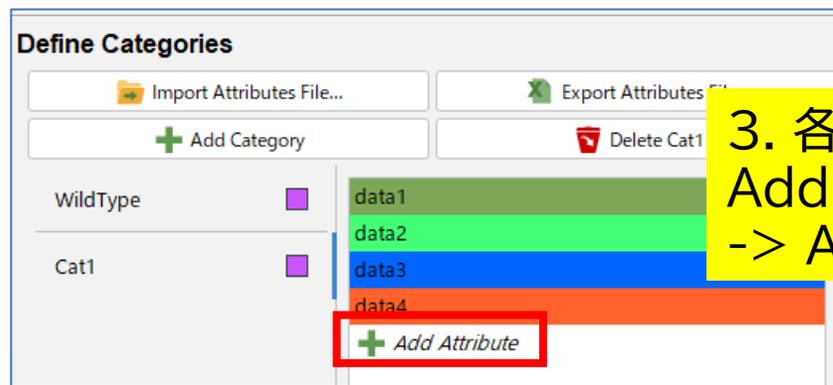
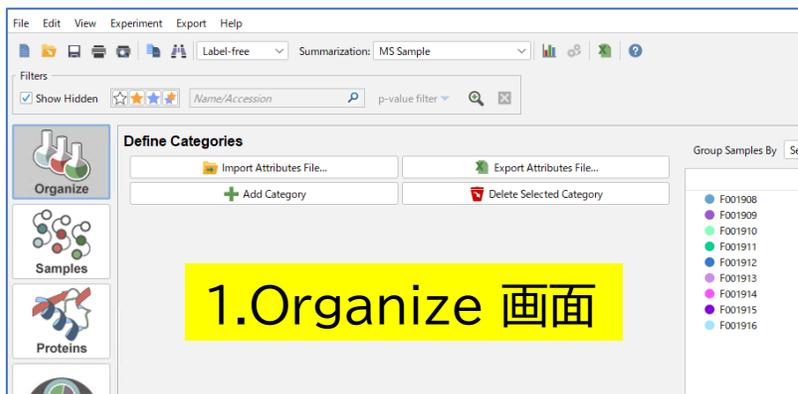
Display Type: Log₁₀ Intensity Normalized Color Options...

Color Legend (Displayed Value)

4.41
3.75
3.09
2.43
1.77

#	Visible	Star	Protein Name	Accession Number	Taxonomy	S1		S2	
						Rep1	Rep2	Rep1	Rep2
1	<input checked="" type="checkbox"/>		sp P00490 PHSM_ECOLI Maltodextrin phosphorylase OS=Escherichia coli (strain K12) GN=malP PE=1 SV=7	sp P00490 PHSM_ECOLI		2.31	2.75	2.31	2.75
2	<input checked="" type="checkbox"/>		sp P41250 SYG_HUMAN Glycine--tRNA ligase OS=Homo sapiens GN=GARS PE=1 SV=3	sp P41250 SYG_HUMAN	Homo sapiens	3.12	3.10	3.12	3.10
3	<input checked="" type="checkbox"/>		sp P0AFG6 ODO2_ECOLI Dihydrodipolysine-residue succinyltransferase component of 2-oxoglutarate dehydro...	sp P0AFG6 ODO2_ECOLI		2.35	2.65	2.35	2.65
4	<input checked="" type="checkbox"/>		sp P24752 THIL_HUMAN Acetyl-CoA acetyltransferase, mitochondrial OS=Homo sapiens GN=ACAT1 PE=1 SV=1	sp P24752 THIL_HUMAN	Homo sapiens	2.77	2.66	2.77	2.66
5	<input checked="" type="checkbox"/>		sp P23381 SYWC_HUMAN Tryptophan--tRNA ligase, cytoplasmic OS=Homo sapiens GN=WARS PE=1 SV=2	sp P23381 SYWC_HUMAN	Homo sapiens	2.82	2.84	2.82	2.84
6	<input checked="" type="checkbox"/>		sp Q9UII2 ATIF1_HUMAN ATPase inhibitor, mitochondrial OS=Homo sapiens GN=ATPIF1 PE=1 SV=1	sp Q9UII2 ATIF1_HUMAN	Homo sapiens	3.09	3.20	3.09	3.20
7	<input checked="" type="checkbox"/>		tr C8Z7Y6 C8Z7Y6_YEAS8 Ret2p OS=Saccharomyces cerevisiae (strain Lalvin EC1118 / Prise de mousse) GN=E...	tr C8Z7Y6 C8Z7Y6_YEAS8		2.90	2.55	2.90	2.55
8	<input checked="" type="checkbox"/>		tr C8ZBZ3 C8ZBZ3_YEAS8 Fas1p OS=Saccharomyces cerevisiae (strain Lalvin EC1118 / Prise de mousse) GN=E...	tr C8ZBZ3 C8ZBZ3_YEAS8		3.66	3.35	3.66	3.35
9	<input checked="" type="checkbox"/>		sp Q8NC51 PAIRB_HUMAN Plasminogen activator inhibitor 1 RNA-binding protein OS=Homo sapiens GN=SERBP...	sp Q8NC51 PAIRB_HUMAN	Homo sapiens	3.90	3.93	3.90	3.93

階層設定の操作



検定

Statistical Test

<input type="radio"/> ANOVA / t-test	P	≥2 Background×Concentrations
<input type="radio"/> Permutation Test	NP	≥2 Background×Concentrations
<input type="radio"/> Mann-Whitney U Test	NP	Exactly 2 Background×Concentrations
<input type="radio"/> Kruskal-Wallis Test	NP	≥2 Background×Concentrations
<input checked="" type="radio"/> None		

Permutation Test

- ノンパラメトリック
- 2群 ○ 3群以上 ○
- ベースはF検定、群間のランダムなデータを入れ替えF値を計算し続ける
- 10000回の 入れ替え計算を行い、入れ替え前のF値より有意差以上に差があった回数を10000 (データ入れ替えの試行回数)で割った値をp-value とする。

t-test / ANOVA

- パラメトリック
- 2群 ○ 3群以上 ○
- ANOVA両側検定を行う。(2群しかない場合 t検定と同じ。)

Mann Whitney U test / Kruskal-Wallis test

- ノンパラメトリック
- 2群→Mann Whitney U, 3群以上→Kruskal-Wallis
- 同じ分布の形、スケールである事を前提とする

多重比較補正

Multiple Test Correction

- Control FWER with Hochberg's step-up and Holm's step-down
- Control FDR with standard Benjamini-Hochberg procedure
- Control FWER with Hochberg's step-up and Holm's step-down
- No correction

多重検定時の第一種過誤に対応する補正

FDR: (False Discovery Rate)
BH法

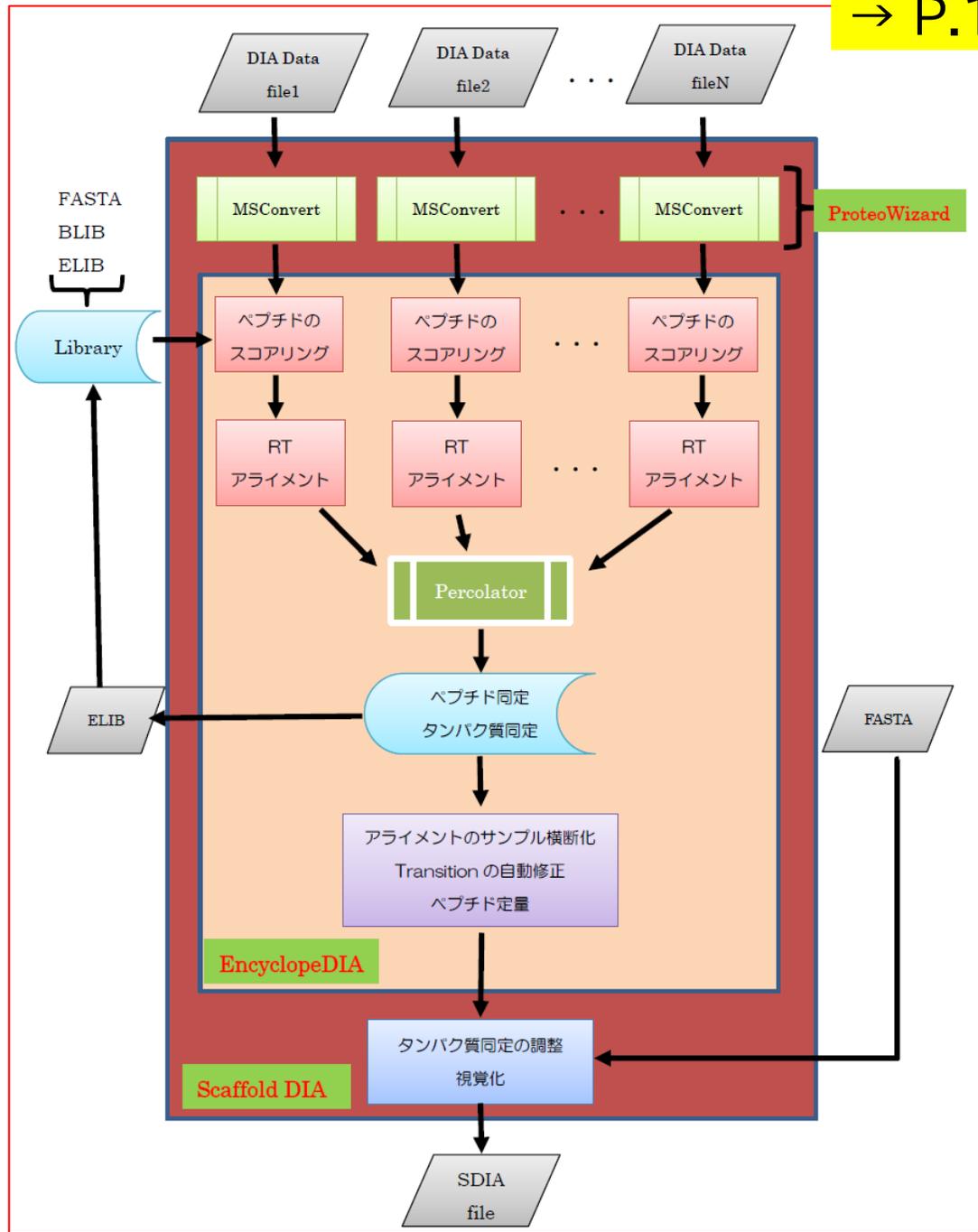
FWER: Family Wise Error Rate
ホッホベルクのステップアップ手順&ホルムの
ステップダウン

解析操作の流れ

- ソフトウェア起動
- 新規作成、Ctrl + N
(menuの「File」-> New やアイコンクリック)
- 検索パラメータ指定、検索開始
- データ取り込み完了
- 属性付与、階層構造化
- データ解析(定性解析、定量解析、検定)
- レポート

内部プログラム

- MSConvert
 - データ変換
- EncyclopeDIA
 - ペプチドピックアップ
 - RTアライメント
 - ペプチド同定
 - タンパク質同定
 - ペプチド定量



その他のトピックス(+日本語マニュアル対応ページ)

・タンパク質のグループ化、クラスター化 (→P.89)

シェアペプチドの存在とタンパク質のグループ化ルール

・Report (→P.91)

Scaffold DIAで出力可能なファイル

・Prositとの連携によるライブラリフリーサーチ (→P.93)

[旧バージョン,3.4.1以下] Prositでライブラリを作成し
Scaffold で使用可能にするための 操作方法

英語マニュアルのAppendix

- **Appendix A.** Structure of Scaffold DIA files (*.sdia)
- **Appendix B.** Computation of FDR in Scaffold DIA
- **Appendix C.** Summarization: Rolling up Values
- **Appendix D.** Missing Values
- **Appendix E.** Shared Evidence Clustering Algorithm
- **Appendix F.** Heat map clustering
- **Appendix G.** Techniques to Control the Family-wise Error Rate
- **Appendix H.** Using Principal Component Analysis
- **Appendix I.** How PCA is Performed in Scaffold DIA
- **Appendix J.** Description of Mouse Right Click Context Menu Commands

インストール環境、その他

■インストール環境

・対応OS

Windows 10,11 (64bit)。MacやLinux利用の場合、mzMLを自身で準備

・メモリ

最低 4GB以上

しかし64GB以上を強く推奨(大規模解析なら128GB以上)

・ストレージ

最低 数百 MB 以上。高速SSD の使用を推奨

・CPU

使用可能なコア数の上限が 64コア