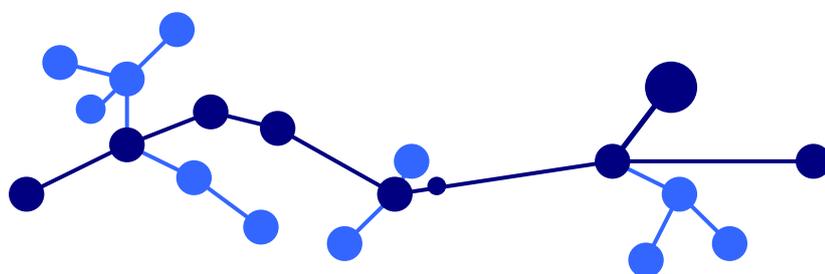


IPA Winter 2020  
Analysis Match  
ユーザーマニュアル



**トミーデジタルバイオロジー株式会社**

## 目次

本マニュアルについて .....	3
Analysis Match について .....	3
Analysis Match のデータソース .....	3
Analysis Match の使用方法 .....	4
解析結果のフィルタリング .....	5
メタデータの評価 .....	6
データセットの検索 .....	7
ヒートマップの表示 .....	8
ヒートマップのフィルタリング .....	10
Activity Plot について .....	11
Activity Plot の起動方法 .....	12
Activity Plot の使用方法 .....	14
解析データセットのレポジトリ .....	18
データセットの解析方法 .....	19
シグネチャの作成 .....	20
シグネチャのスコアリング .....	21
Customize Table でカラムに追加または削除できるメタデータ .....	22
製品に関するお問い合わせについて .....	23

## 本マニュアルについて

本マニュアルは IPA の追加有料オプション機能 Analysis Match に関して解説したものです。

Analysis Match は IPA 本体とは別途契約を結ぶことで使用が可能となります。

IPA の動作環境やログイン方法など、Analysis Match 以外の機能に関しては IPA Winter 2020 ユーザーガイドをご参照ください。

## Analysis Match について

Analysis Match は IPA に機能を追加する有料オプションです。

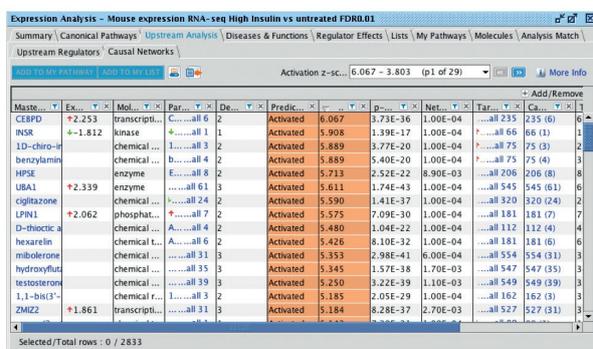
自身の IPA Core Analysis 解析結果と類似、または反対する生物学的結果を持つ、他の IPA Core Analysis を、自身の過去の解析結果や、公共データベース由来のヒトおよびマウスの疾患/がんなどに関する発現解析結果から自動的に探索し、比較解析ができるようになります。

この比較解析は、Canonical Pathways, Upstream Regulators, Causal Networks, Diseases and Functions 解析結果の共通パターンに基づいてマッチングされます。

## Analysis Match のデータソース

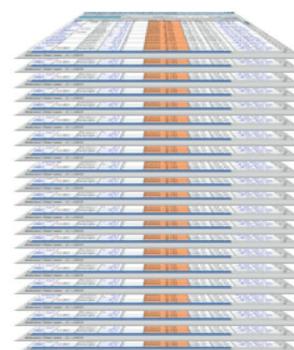
Analysis Match に使用される解析は、SRA、GEO、Array Express、TCGA、LINCS などにある 310,000 以上のヒトおよびマウスの疾患/がんなどに関する発現解析結果を QIAGEN の OmicSoft チームが精査し再構築したものです。

2020 年 12 月現在、80,000 以上の解析結果が利用可能です。



Maste...	Ex...	Mol...	Par...	De...	Predic...	p...	Net...	Tar...	Ca...
CEBPD	+2.253	transcripto...	C.....all 6	2	Activated	5.067	3.73E-36	1.00E-04	.....all 235 (6)
WDR	+1.812	kinase	+.....all 1	1	Activated	5.908	1.39E-17	1.00E-04	+.....all 66 (8)
1D-chiro-		chemical ...	1.....all 3	2	Activated	5.889	3.77E-20	1.00E-04	+.....all 75 (5)
benzylamin		chemical ...	b.....all 4	2	Activated	5.889	5.40E-20	1.00E-04	+.....all 75 (5)
HPSE		enzyme	E.....all 8	2	Activated	5.713	2.52E-22	8.90E-03	.....all 206 (8)
UBA1	+2.339	enzyme	.....all 61	3	Activated	5.611	1.74E-43	1.00E-04	.....all 545 (61)
ciglitazone		chemical ...	.....all 24	2	Activated	5.590	1.41E-37	1.00E-04	.....all 320 (24)
LPN1	+2.062	phosphat...	+.....all 7	2	Activated	5.575	7.09E-30	1.00E-04	.....all 181 (7)
D-thucic a		chemical ...	A.....all 4	2	Activated	5.480	3.04E-22	1.00E-04	.....all 132 (4)
hexarelin		chemical ...	A.....all 6	2	Activated	5.426	8.10E-32	1.00E-04	.....all 181 (6)
mibolerone		chemical ...	.....all 31	3	Activated	5.353	2.98E-41	6.00E-04	.....all 554 (31)
hydroxyflut		chemical ...	.....all 35	3	Activated	5.345	1.57E-38	1.70E-03	.....all 547 (35)
testosterone		chemical ...	.....all 39	3	Activated	5.250	3.22E-39	1.10E-03	.....all 549 (39)
1,1-bis(3'-		chemical r...	1.....all 3	2	Activated	5.185	2.05E-29	1.00E-04	.....all 162 (3)
ZMIZ2	+1.861	transcripto...	.....all 31	3	Activated	5.184	8.28E-37	2.70E-03	.....all 527 (31)

Match against  
→  
>80,000 analyses



## Analysis Match の使用方法

Analysis Match は、IPA Core Analysis 解析結果ウィンドウの上にあるタブの一番右側に表示されます。Analysis Match のライセンスを取得されている場合は、自身の解析結果と類似、または反対する生物学的結果を持つ、他の IPA Core Analysis が自動的にランク付けされて表示されています。

デフォルトでは、解析結果は全体的な類似性スコアに基づいて、自身の解析結果と最も類似しているものから、最も類似性の低いものへとランク付けされます。解析結果は、Canonical Pathways, Upstream Regulators, Causal Networks, Diseases and Functions の各解析結果から、比較解析用に生成されたシグネチャに基づいてマッチングされます。シグネチャの生成とマッチングの基準の詳細については、[こちら](#)をご参照ください。

上図は、全体的な z スコアのパーセンテージに基づいて、最も類似した解析結果と、相違した解析結果のみを表示するようにフィルタリングされています。

z スコアのカラムの左側にある 4 つの色付きのカラムは、自身の解析結果に対する各解析結果のシグネチャの類似度の割合を表しています。紫色は類似性が高いことを示し、水色は類似性が低いことを示しています。

左から、Canonical Pathways (CP), Upstream Regulators (UR), Causal Networks (CN), Downstream Effects (Diseases & Functions) (DE) の各解析結果のシグネチャの類似度の割合を表すカラムです。

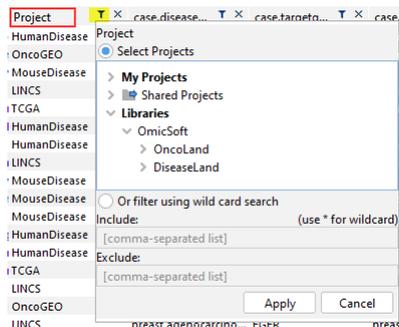
z スコアのカラムは、これら 4 つのシグネチャの類似度の平均です。z スコアのカラムの右側にある白と紫のカラムは、シグネチャごとにフィッシャーの正確確率検定で算出された p 値を表示しています。

## 解析結果のフィルタリング

カラムの上部にある漏斗アイコンをクリックして、数値またはテキストを入力することで、任意のカラムで結果をフィルタリングすることができます。z スコアカラムの場合、入力したカットオフ値は絶対値として扱われます。たとえば、50 を入力した場合、結果はスコア>50 または<-50 のものにフィルタリングされます。

結果を特定の(またはすべての)データソースや、自身のプロジェクトのみに制限することができます。

プロジェクトのフィルターファンネルアイコン(前ページの図で左側オレンジ色の漏斗アイコン)をクリックしてから、1つ以上のデータソースをクリックして選択します。ラジオボタンを使用してワイルドカード検索に切り替えることもできます。



それぞれの解析結果にあるメタデータでフィルタリングすることもできます。それぞれの解析結果には、生物種、比較の種類(疾患対正常、処理対対照など)、その他多くの注釈が付けられています。

Customize Table メニューを使用して、カラムに追加または削除できるメタデータのリストは[こちら](#)をご参照ください。

たとえば、sampledatamode フィールドをカラムに追加し、RNAseq の解析結果のみにフィルタリングすることなどが可能です。

Analysis No.	Project	Compari...	sample...	weblink	CP (z-sc...	UR (z-sc...	CN (z-sc...	DE (z-sc...	z-sc...	CP (p-va...	UR (p-va...	CN (p-va...	DE (p-va...	p-value
114	head and ne	TCGA	Other Compari...	NFE2L2_Somatic...	RnaSeq_Transcr...	https://cancergen	28.40	50.00	19.60	1.46E-07	2.34E-31	2.89E-02	19.51	
56	central squam	TCGA	Other Compari...	ATR_Somatic...	RnaSeq_Transcr...	https://cancergen	31.11	46.77	19.47	2.25E-11	2E-32	2.89E-02	21.94	
160	lung squamo	TCGA	Other Compari...	KEAP1_Somatic...	RnaSeq_Transcr...	https://cancergen	31.11	46.77	19.47	1.41E-07	3.38E-27		16.66	
6	emphysema [ju	MouseDisease	Treatment vs. C...	SubjectTreatme...	RnaSeq_Transcr...	https://www.ncbi	28.40	43.30	17.92	4.65E-05	3.47E-21		12.40	
1	normal control	MouseDisease	Treatment vs. C...	TreatmentTreat...	RnaSeq_Transcr...	https://www.ncbi		35.36	34.30	3.46E-01	1.1E-12	3.04E-08	9.97	
122	bladder caro	TCGA	Other Compari...	NFE2L2_Somatic...	RnaSeq_Transcr...	https://cancergen	28.40	38.19	16.65	8.46E-07	1.74E-16		10.92	
140	head and ne	TCGA	Other Compari...	KEAP1_Somatic...	RnaSeq_Transcr...	https://cancergen	28.40	36.80	16.30	3.81E-07	1.24E-13	3.21E-02	10.41	
37	papillary renal	TCGA	Other Compari...	CEB1_Somatic...	RnaSeq_Transcr...	https://cancergen	28.40	33.85	15.56	7.44E-10	1.22E-12	3.26E-03	11.76	
69	small intestine	OncoGEO	Treatment vs. C...	TreatmentTreat...	RnaSeq_Transcr...	https://www.ncbi		25.00	34.30	3.11E-02	2.08E-05	3.04E-08	6.85	
46	small intestine	OncoGEO	Treatment vs. C...	TreatmentTreat...	RnaSeq_Transcr...	https://www.ncbi	31.11	27.00	14.53	1.75E-06	5.11E-06		5.52	
42	head and neck	TCGA	Other Compari...	HGF_Somatic...	RnaSeq_Transcr...	https://cancergen	25.40	32.27	14.42	3.32E-04	1.35E-09		6.17	
44	normal control	MouseDisease	Tissue1 vs. Tissu...	Tissue => small...	RnaSeq_Transcr...	https://trace.ncbi	25.40	22.82	34.30	14.28	2.09E-04	6.56E-08	5.43	
162	small intestine	OncoGEO	Treatment vs. C...	TreatmentTreat...	RnaSeq_Transcr...	https://www.ncbi	28.40	27.00	14.02	3.77E-05	6.79E-11		7.30	
15	neurodegener	MouseDisease	Other Compari...	Genotype => TB...	RnaSeq_Transcr...	https://www.ncbi	28.40	25.00	13.35	3.81E-07	5.98E-05	9.04E-02	5.84	
29	bile duct cano	TCGA	Other Compari...	DOT1L_Somatic...	RnaSeq_Transcr...	https://cancergen	25.40	27.00	13.10	1.56E-07	1.67E-08		7.29	
24	systemic lupu	HumanDisease	CellType1 vs. Cel...	DiseaseState:Cell...	RnaSeq_Transcr...	https://www.ncbi	10.37		42.01	13.09	3.31E-06		8.46	
3	normal control	OncoGEO	Other Compari...	Transfection =>...	RnaSeq_Transcr...	https://www.ncbi	28.40	22.82	12.80	5.15E-05	5.89E-04		3.76	
21	hepatocellular	HumanDisease	Treatment1 vs. T...	ExperimentGrou...	RnaSeq_Transcr...	https://www.ncbi		51.03	12.76	1.12E-05	3.5E-33	2.89E-02	19.47	
3	normal control	MouseDisease	Treatment vs. C...	PreTreatmentTr...	RnaSeq_Transcr...	https://www.ncbi	25.40	25.00	12.60	7.85E-04	5.98E-05		3.66	
21	normal control	MouseDisease	Treatment1 vs. T...	CellDescription...	RnaSeq_Transcr...	https://www.ncbi	25.40	25.00	12.60	7.85E-04	5.98E-05		3.66	
2	normal control	MouseDisease	Treatment vs. C...	DiseaseState:Fre...	RnaSeq_Transcr...	https://www.ncbi	25.40	25.00	12.60	7.85E-04	5.98E-05	1.1E-02	4.54	
133	lung squamo	TCGA	Other Compari...	NFE2L2_Somatic...	RnaSeq_Transcr...	https://cancergen	25.40	24.06	12.36	1.28E-05	3.48E-23		13.68	
3	normal control	MouseDisease	Other Compari...	CellDescription...	RnaSeq_Transcr...	https://www.ncbi			48.51	12.13	8.41E-03	3.89E-14	7.74	
10	colon cancer	OncoGEO	Treatment vs. C...	Transfection =>...	RnaSeq_Transcr...	https://www.ncbi	25.40	22.82	12.06	5.24E-04	5.89E-04		3.26	
17	Alzheimer's d	MouseDisease	Other Compari...	Tissue:Genotyp...	RnaSeq_Transcr...	https://www.ncbi	25.40	22.82	12.06	1.48E-03	8.01E-05		5.66	
16	normal control	MouseDisease	Tissue1 vs. Tissu...	Tissue => jejun...	RnaSeq_Transcr...	http://www.ncbi		13.69	34.30	12.00	4.01E-04	1.16E-05	4.17	
98	nonalcoholic	HumanDisease	Treatment1 vs. T...	GenotypeSampl...	RnaSeq_Transcr...	https://www.ncbi	31.11	16.67	11.94	2.12E-01	3.31E-06	5.98E-05	5.19	
126	small intestine	OncoGEO	Treatment vs. C...	Treatment => b...	RnaSeq_Transcr...	https://www.ncbi	28.40	19.29	11.92	5.68E-05	5.11E-06	7.54E-02	5.33	
21	normal control	MouseDisease	Other Compari...	Age(month):Ge...	RnaSeq_Transcr...	https://www.ncbi	25.40	21.82	11.81	6.84E-02	1.21E-04	4.57E-15	9.71	
1	acute myeloid	HumanDisease	CellType1 vs. Cel...	CellType => mo...	RnaSeq_Transcr...	https://www.ncbi	12.70		34.30	11.75	7.85E-04	9.51E-07	4.56	
11	normal control	MouseDisease	Treatment1 vs. T...	GenotypeSampl...	RnaSeq_Transcr...	https://www.ncbi	26.94	19.29	11.56	4.61E-02	6.45E-09	5.11E-06	8.20	
1	familial dyspl	HumanDisease	Other Compari...	DiseaseStage =>...	RnaSeq_Transcr...	https://www.ncbi	25.40	20.41	11.45	2.44E-04	4.76E-03		2.97	
20	hepatocellular	HumanDisease	Treatment1 vs. T...	ExperimentGrou...	RnaSeq_Transcr...	https://www.ncbi		45.64	11.41	2.08E-04	1.86E-24		13.71	
13	normal control	MouseDisease	Other Compari...	SampleMaterial...	RnaSeq_Transcr...	https://www.ncbi		45.37	11.34	6.6E-02	4.95E-01	7.07E-11	5.82	

## メタデータの評価

目的のデータセットを選択し、**Evaluate Metadata** ボタンを使用することで、90 を超えるキュレートされたメタデータフィールドからメタデータを自動的に分析し、選択した解析に共通性があるかどうかを検出することができます。

Metadata field	Significant term	p-value	Selected analyses with t...	Total analyses with t...	Selected analyses wit...	Total analyses with a...
platformname	NGS.Illumina.HiSeq2500	3.81E-08	11	2951	29	62813
control.diseasestate	experimental autoimmune e...	2.79E-05	2	17	29	62813
case.diseasestate	experimental autoimmune e...	9.49E-05	2	31	29	62813
platformname	NGS.Illumina.HiSeq2000	1.92E-04	12	8441	29	62813
case.dosage	CTLA4-Ig 7.5 ug/ml;anti-CD3 ...	2.88E-04	1	1	11	38206
control.dosage	anti-CD3 2 ug/ml;anti-CD28 ...	2.88E-04	1	1	11	38206
case.cellline	A2780	2.93E-04	1	8	1	27286
case.celltype	anergic CD4+ T cell	3.25E-04	1	1	13	40007
case.treatmentstatus	low carbohydrate diet	4.14E-04	1	1	1	2418
case.diseasestate	atypical deletion Williams sy...	4.62E-04	1	1	29	62813
comparisoncontrast	Treatment:TreatTime[hours] ...	4.62E-04	1	1	29	62813

クエリーデータは、NRF2 アクチベータ CDDO-me で処理されたマウスの腎臓からのものですが、それに一致する解析として、喫煙に暴露されたマウスの肺に由来する解析がエンリッチしています。この例では、最も類似していた30の解析が選択されましたが、そのうちの8つは「3R4F 喫煙暴露」からのものです。データベース中の80,000の解析のうちで同じ処理を行っているものは59のみであるため、上位30の結果の中にそのうちの8つが表示されることは非常に重要です(p 値 = 2.17E-13)。

### カラムの定義

**Metadata field:**重要な用語(キーワードまたはフレーズ)が見つかった OmicSoft メタデータフィールドの名前。

**Significant term:**選択した一連の解析で大幅にエンリッチしていることがわかった実際の用語。

**p-value:**選択した解析の特定の用語に対して計算されたフィッシャーの正確検定の p 値。

**Selected analyses with term:**選択された解析のうち、重要な用語を持つ解析数。

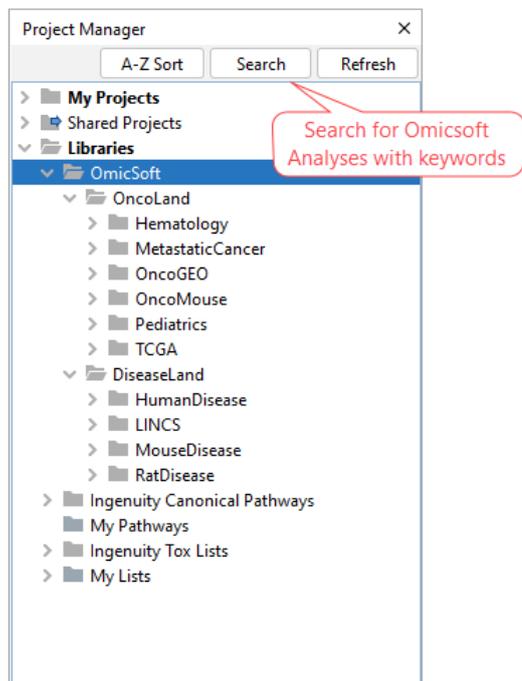
**Total analyses with term:**重要な用語を含む選択された解析の、全体における解析数。

**Selected analyses with any value in field:**選択した解析に関連する用語が含まれている可能性のある解析数。

**Total analyses with any value in field:**そのメタデータフィールドに任意の用語がある解析数。

## データセットの検索

OmicSoft のデータセットは IPA のライブラリに保存されているため、検索機能を使用して目的のデータセットを探ることができます。ただしこれらは読み取り専用で、IPA のライブラリからダウンロードすることはできません。



以下の例は、ヒト(human)の喘息(asthma)解析結果で、albuterol を含まないものを検索しています。このような検索結果から、ダブルクリックして解析を開くか、最大 20 個を選択して比較解析することができます。

Dataset and Analysis Search

Search: human AND asthma NOT albuterol

Found 486 results in 91ms for query [human AND asthma NOT albuterol]

Folder Types: dataset (247), analysis (239)

Projects: HumanDisease (484)

Folder Name	Folder Type	Creation Date	Folder Id
2- paucigranulocytic asthma [sputum] NA 6186	analysis	2020/10/01 01:40:29	9125456
1- neutrophilic asthma [sputum] NA 6185	analysis	2020/10/01 19:06:52	9136692
1- asthma [sputum] NA 9627	analysis	2020/10/10 01:32:57	9196174
4- asthma [peripheral blood] LpA 5825	analysis	2020/09/30 15:42:41	9117256
1- asthma [bronchial epithelium] Transfection_HRV-infected 2473	analysis	2020/10/10 01:58:39	9196361
5- asthma [nasal epithelium] NA 15079	analysis	2020/09/30 11:42:30	9113984
1- asthma [nasal epithelium] NA 15075	analysis	2020/10/10 01:40:42	9196223
1- asthma [airway smooth muscle] FBS 8445	analysis	2020/10/10 02:00:25	9196375
1- asthma [tracheal epithelium] Infection_rhinovirus 14079	analysis	2020/10/10 01:50:57	9196301
1- asthma [bronchial epithelium] Transfection_HRV-infected 2471	analysis	2020/10/10 01:43:46	9196245
8- asthma [nasal epithelium] NA 15082	analysis	2020/10/07 15:21:00	9169417
11- asthma [peripheral blood] NA 1737	analysis	2020/10/01 12:57:29	9133210
12- asthma [peripheral blood] NA 1738	analysis	2020/10/01 11:24:35	9132200
1- allergic asthma [bronchial epithelium] NA 5879	analysis	2020/10/10 02:08:26	9196427
1- asthma [sputum] NA 6004	analysis	2020/10/10 01:51:40	9196307
1- asthma [airway epithelium] NA 2957	analysis	2020/10/10 01:48:13	9196275
1- asthma [bronchial epithelium] NA 6387	analysis	2020/10/10 01:44:25	9196251
1- asthma [lung] NA 2658	analysis	2020/10/10 01:42:13	9196232
1- asthma [peripheral blood] NA 5286	analysis	2020/10/10 01:39:23	9196213
2- asthma [peripheral blood] NA 5287	analysis	2020/10/08 23:44:00	9185304

OmicSoft > DiseaseLand > HumanDisease > Analyses

2- paucigranulocytic asthma [sputum] NA 6186

Case/Control Differences

Key	Case	Control
diseasestate	paucigranulocytic asthma	eosinophilic asthma

Comparison Context

comparisoncategory Disease1 vs. Disease2  
 comparisoncontrast DiseaseState => paucigranulocytic asthma vs eosinophilic asthma  
 organism human  
 platformname illumina.HumanRef-8\_V2\_0\_R1\_11223162\_A  
 tissue sputum  
 treatment NA

All Experiment Metadata

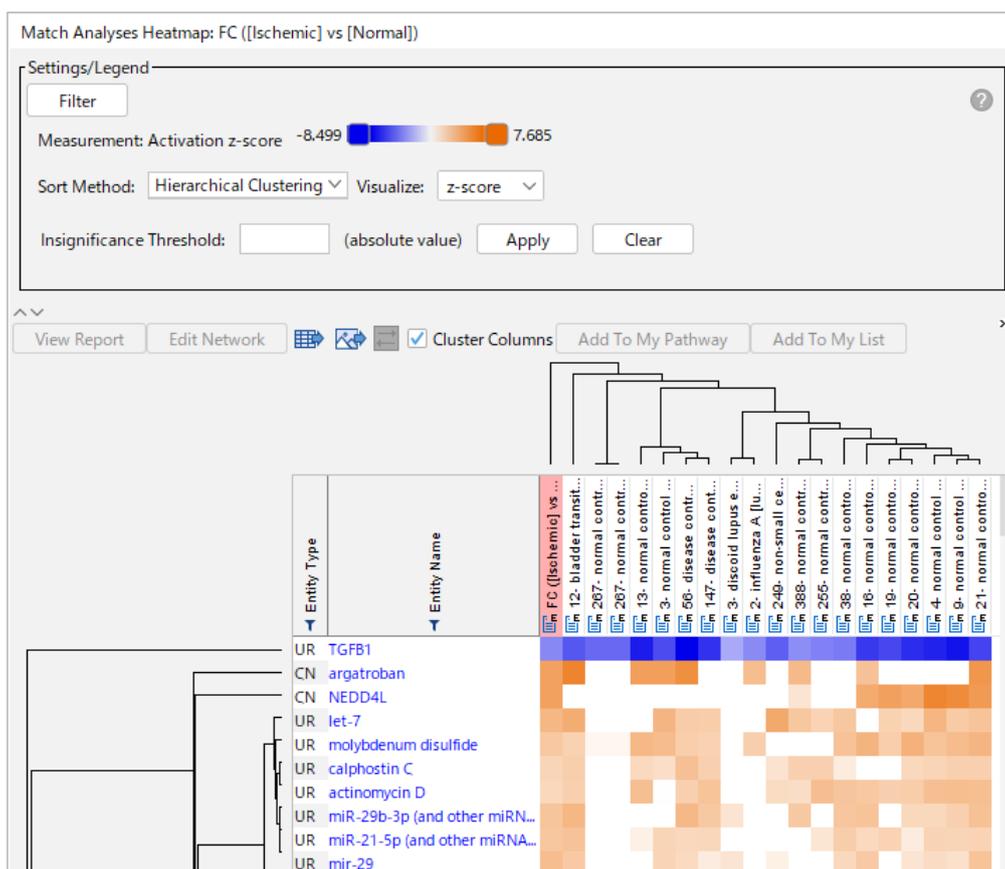
case.diseasestate paucigranulocytic asthma  
 case.dosage[gy] .  
 case.samples GSM1096958;GSM1096959;GSM1096960;GSM1096961;GSM1096962;GSM1096963;C  
 case.samplesource sputum  
 case.samplingtime[dp] .  
 case.samplingtime[hours] .  
 case.samplingtime[hp] .  
 case.tissue sputum  
 case.treatment NA  
 case.treattime[days] .  
 case.treattime[hours] .  
 case.treattime[minutes] .

collection ImmunoHuman  
 comparisoncategory Disease1 vs. Disease2  
 comparisoncontrast DiseaseState => paucigranulocytic asthma vs eosinophilic asthma

## ヒートマップの表示

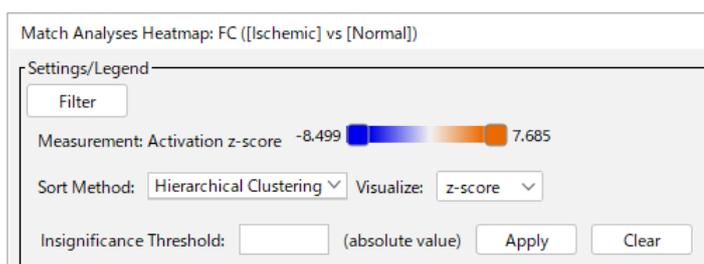
ソートやフィルタリングなどで解析結果をリスト表示した後、目的の解析結果を選択して View as Heatmap ボタンで、ヒートマップを表示することができます。

エンティティ (row) ごとに各解析結果のシグネチャの類似度の割合がクラスター化されて表示されます。ヒートマップの上にある Cluster Columns にチェックを入れると各解析 (Columns) をクラスター化することができます。

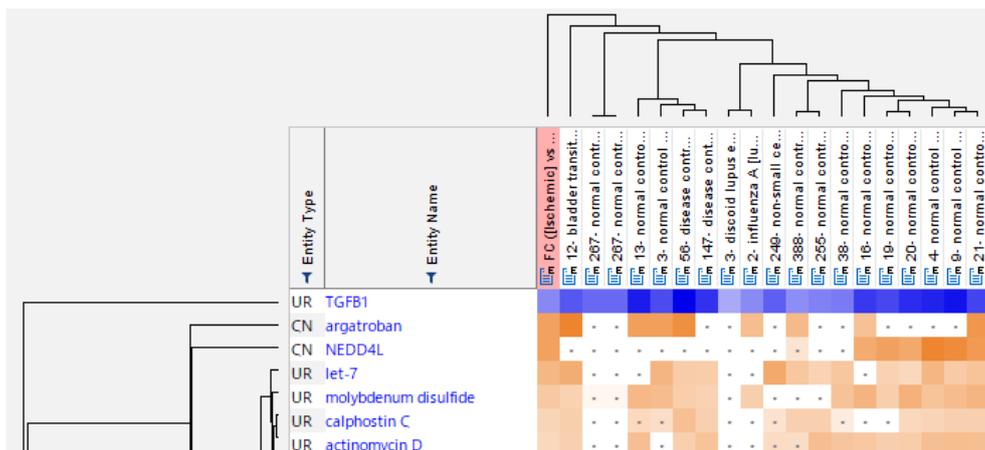


ヒートマップ内のそれぞれのセルは、解析内のエンティティの z スコアで色付けされています。オレンジ色は正の z スコアを持ち、青い正方形は負の z スコアを持ちます。また自身の解析 (ピンク色で強調表示) は、左側の列にリストされているすべてのエンティティに対して有意な z スコア ( $|z| > 2$ ) を持っていますが、その他の選択された解析は有意な z スコアを有していない場合があります。

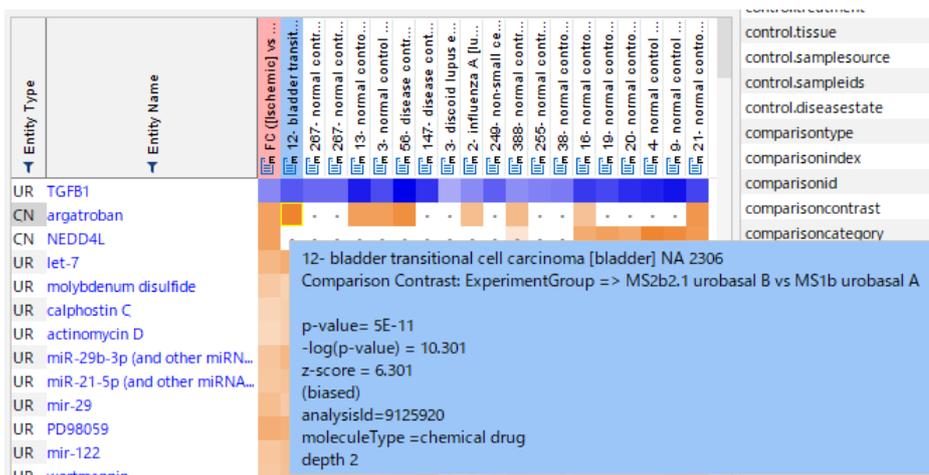
Insignificance Threshold でカットオフを設定することが可能です。



例えばカットオフを 2 に設定すると、z スコアが 2 未満のセルにはドットが表示されます。



Analysis Match ヒートマップの正方形または列の上にカーソルを置くと、OmicSoft 解析の比較対照メタデータフィールドが表示されます。このフィールドは、各発現差解析データセットの実験サンプルと対照サンプルの間で比較されている内容を簡潔に記述したものです。



## ヒートマップのフィルタリング

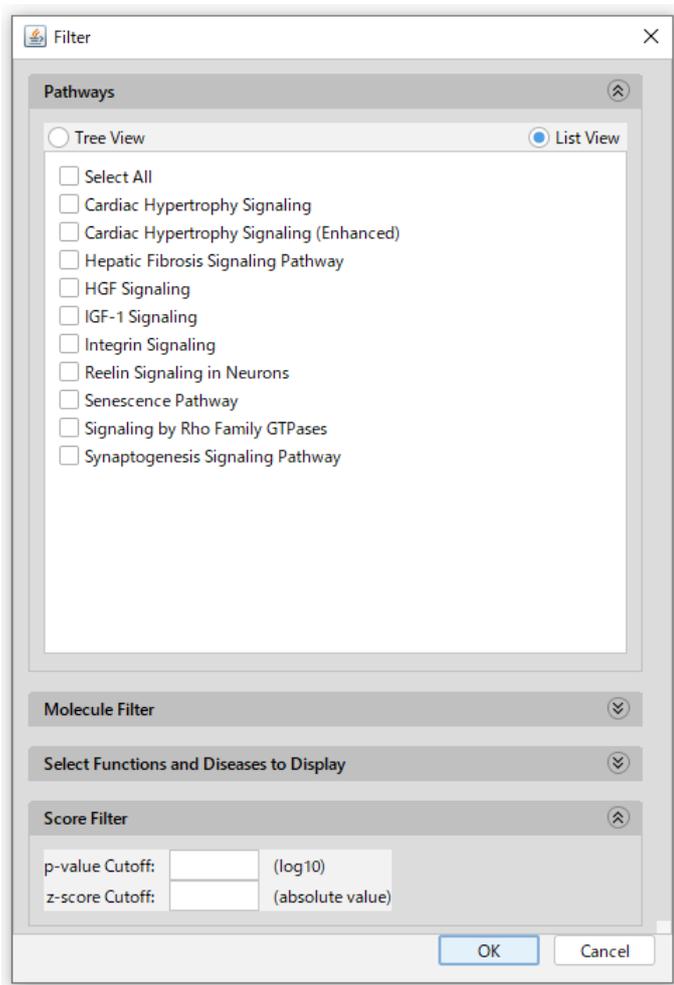
ヒートマップは、いくつかの方法でフィルタリングができます。

Entity Type にあるフィルターファンネルを使用して、Upstream Regulators は UR、Causal Networks は CN、Canonical Pathways は CP、Downstream Effects(Diseases&Functions)は DE でフィルタリングできます。

”UR, CN”のような組み合わせでフィルタリングすることで、両方を見ることができます。

Entity Name のフィルターファンネルを使用すると、特定のエンティティ名をフィルタリングできます。例えば PPARA、PPARG などを見つけるには、PPARA を検索するか、PPAR\* などのワイルドカード(アスタリスク)を使用します。

ウィンドウの左上にある FILTER ボタンで 4 種類のシグネチャのすべてに対してフィルタリングをすることができます。例えば、Molecule Type フィルターを使用して、Upstream Regulators と Causal Networks のレギュレータ分子を、「転写レギュレータ」のみに制限することなどができます。

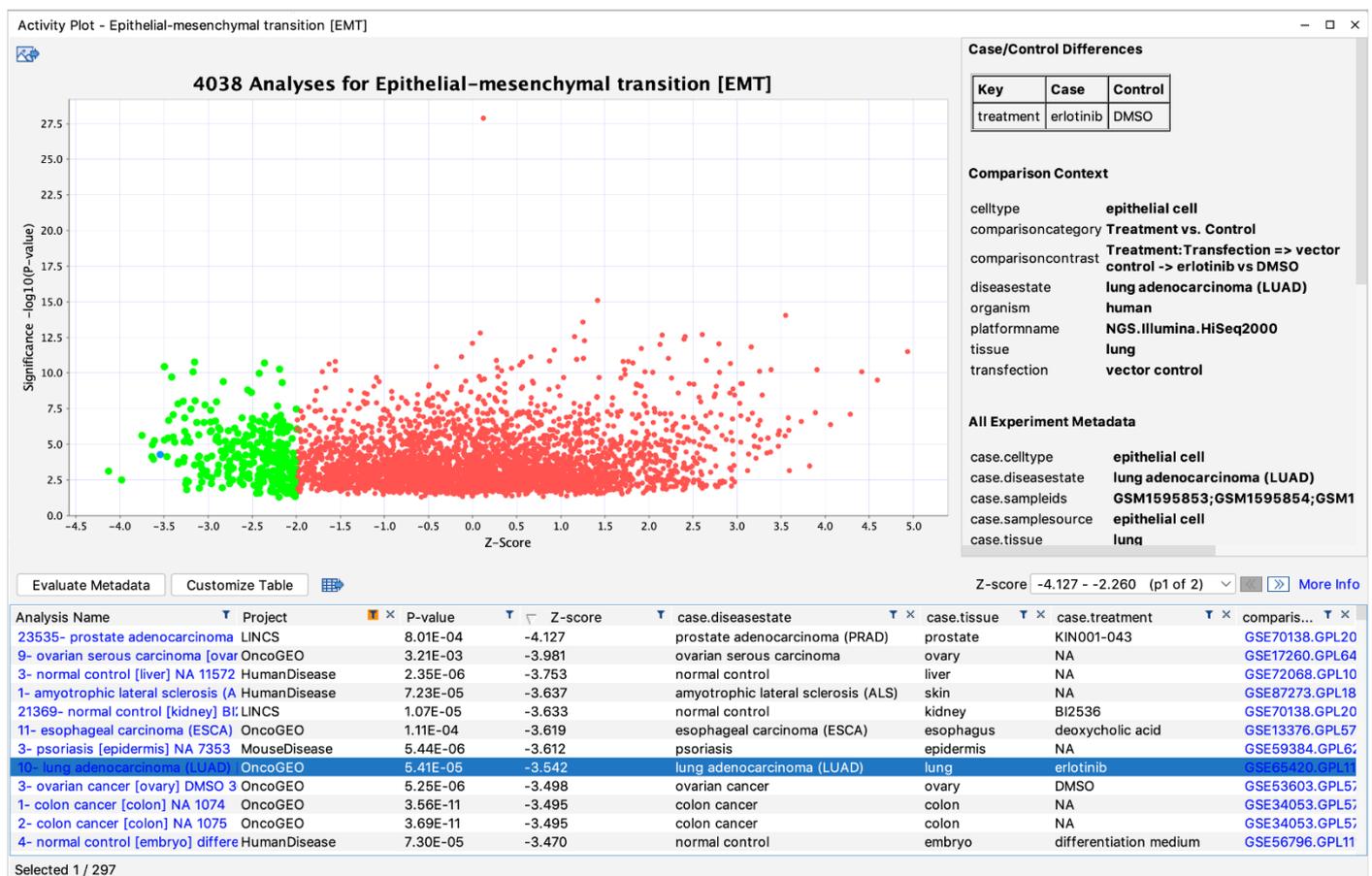


## Activity Plot について

Activity Plot 機能では、80,000 以上の OmicSoft Land 分析における Canonical Pathways、Upstream Regulators / Causal Network master regulators、Diseases and Functions などの単一の IPA エンティティのアクティビティを可視化し、探索することができます。疾患の状態、薬剤や他の治療法、ネットワークアウトなどを表す何千ものデータセットにわたって予測される生物学的活性を Analysis Match データベースで探索することにより、IPA エンティティに関する見識更に高めることができます。Activity Plot 機能は、Analysis Match ライセンスに含まれています。

この機能を使用すると、「どの治療法が上皮から間葉系への移行(EMT)を阻害すると予測されるか?」「どのような疾患状態で ILK シグナル伝達経路が活性化されるのか?」「どのタイプのがんにおいて STAT3 が上流制御因子として活性化されているか?」などの質問に答えることができます。

下の図は、Analysis Match データベースの 80,000 以上の解析における上皮から間葉系への移行機能の予測される活性を示しています。EMT は 3545 の解析で有意に変動しており(プロット内のドット、プロットタイトル)、さらに EMT が強く阻害されている(z スコア < -2)解析のサブセットを選択し、comparison type 「Treatment vs. Control」についてフィルタリングしたものです(緑色の点)。



## Activity Plot の起動方法

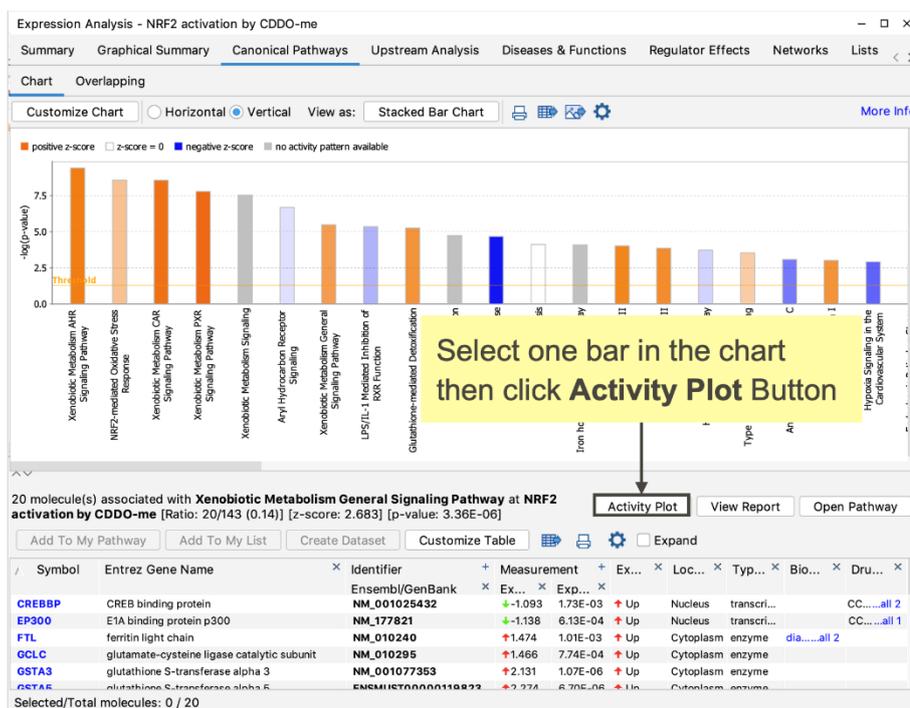
特定のエンティティの Activity Plot を起動するには、主に 2 つの方法があります。

検索結果からの場合は、適切な検索タブに興味のある用語を入力して検索し、検索結果の中から 1 つのチェックボックスを選択して、ツールバーの Activity Plot ボタンをクリックします。

上流調節因子として表示するか、マスターレギュレータ(depth 1~3)として表示するかを選択します。

Core Analysis 結果からの場合は、Canonical Pathways、Upstream Analysis、Diseases & Functions のいずれかの結果から起動することができます。

Canonical Pathways タブでは、棒グラフで興味のある Pathway をクリックし、Activity Plot ボタンをクリックします。



Upstream Analysis タブでは、興味のある Upstream regulator または Causal network を選択し、Activity Plot ボタンをクリックします。

Upstream ...	Expr Fold ...	Molecule Type	Predict...	Activati...	p-v...	Target ...	Mecha...
NRF2		transcription regulator	Activated	4.568	4.29E-20	↑ABCC1, ...all 69	239 (1)
GSR	↑1.566	enzyme	Inhibited	-4.562	1.01E-19	↑ABCC2, ...all 21	
TXNRD1	↑1.958	enzyme	Inhibited	-4.543	2.60E-17	↑ABCC2, ...all 21	
Breast cancer		disease	Inhibited	-0.642	5.65E-17	↑ABCC1, ...all 119	
tert-butyl-hydroquino		chemical rea	Inhibited	2.667	6.76E-14	↑ABCC1, ...all 25	272 (8)
1,2-dithiol-3-thione		chemical reagent	Activated	2.830	4.96E-13	↑ABCC3, ...all 37	104 (4)
MAFG	↑1.155	transcription regulator	Activated	2.611	7.06E-11	↑ABCC3, ...all 11	

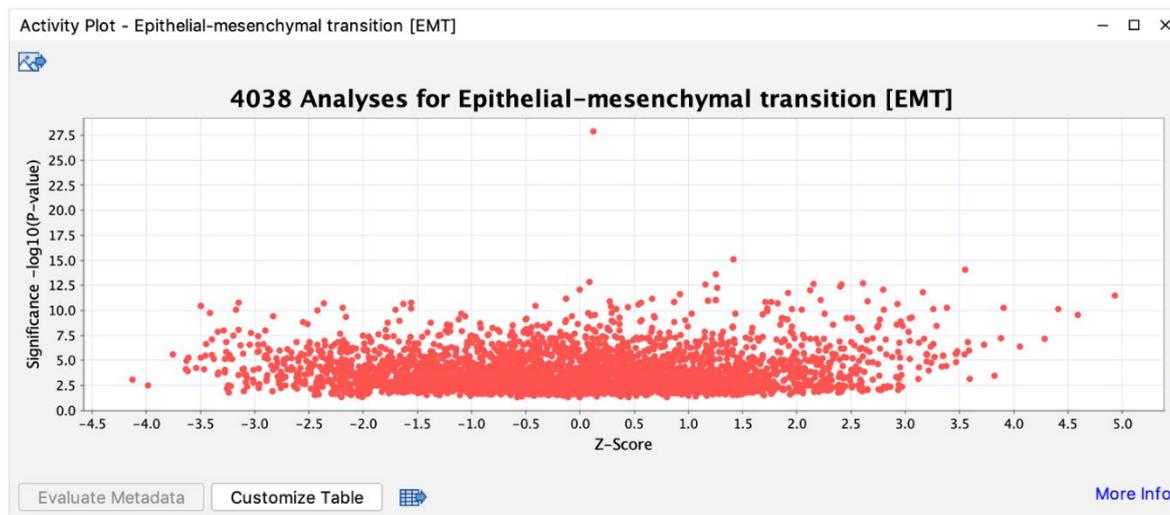
Diseases and Bio Function タブまたは Tox Functions tab タブでは、興味のある Disease または Function を選択し、Activity Plot ボタンをクリックします。

Categories	Diseases or Functions A...	p-value	Predicted ...	Activati...	Molecules	# Molecules
Developmental Disorder	Midline defect	2.40E-04	Increased	4.236	↑ALDH1A1, ...all 31	31
Drug Metabolism	Metabolism of xenobiotic	3.50E-12	Increased	2.621	↑ABCC2, ...all 26	26
Drug Metabolism	Conjugation of glutathione	1.42E-04	Increased	2.439	↑GSTA3, ↑G...all 6	6
Lipid Metabolism, Smal...	Conjugation of lipid	7.31E-05	Increased	2.425	↑GSTP1, ↑S...all 6	6
Molecular Transport	Release of heavy metal	5.05E-05	Increased	2.224	↑ABCC1, ↑G...all 6	6
Amino Acid Metabolis...	Transport of methotrexate	1.57E-05	Increased	2.202	↑ABCC1, ↑A...all 5	5
Drug Metabolism	Joining of glutathione	1.79E-05	Increased	2.147	↑GLRX, ↑GS...all 7	7
Drug Metabolism, Prote...	Synthesis of glutathione	1.42E-04	Increased	2.000	↑GCLC, ↑G...all 6	6

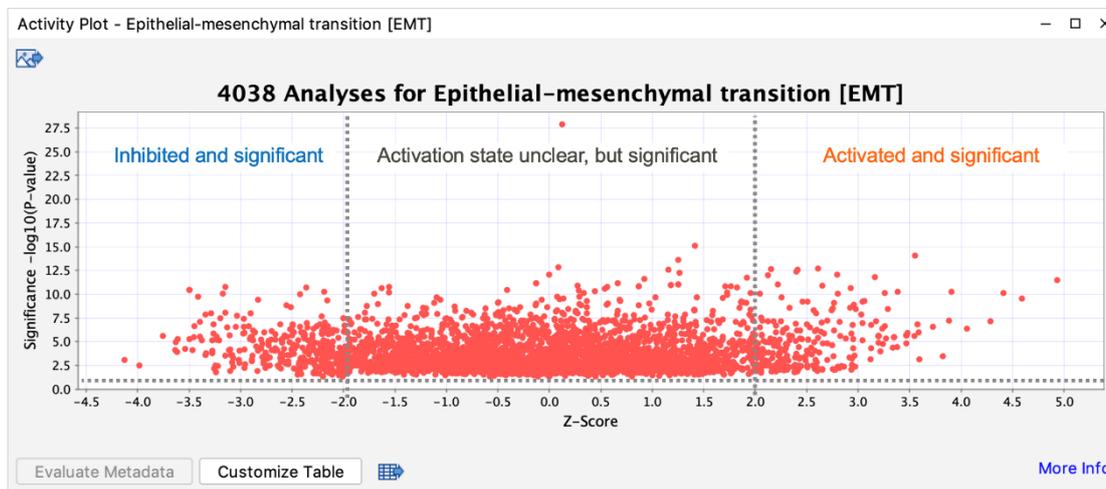
解析中のエンティティからアクティビティプロットを起動すると、その解析の "ドット" が自動的に選択され(明るい緑色で表示)、プロットの下の方に表示されます。

## Activity Plot の使用方法

Activity Plot を起動すると、この例(Biological function "epithelial-mesenchymal transition" (EMT))のようなチャートが表示されます。

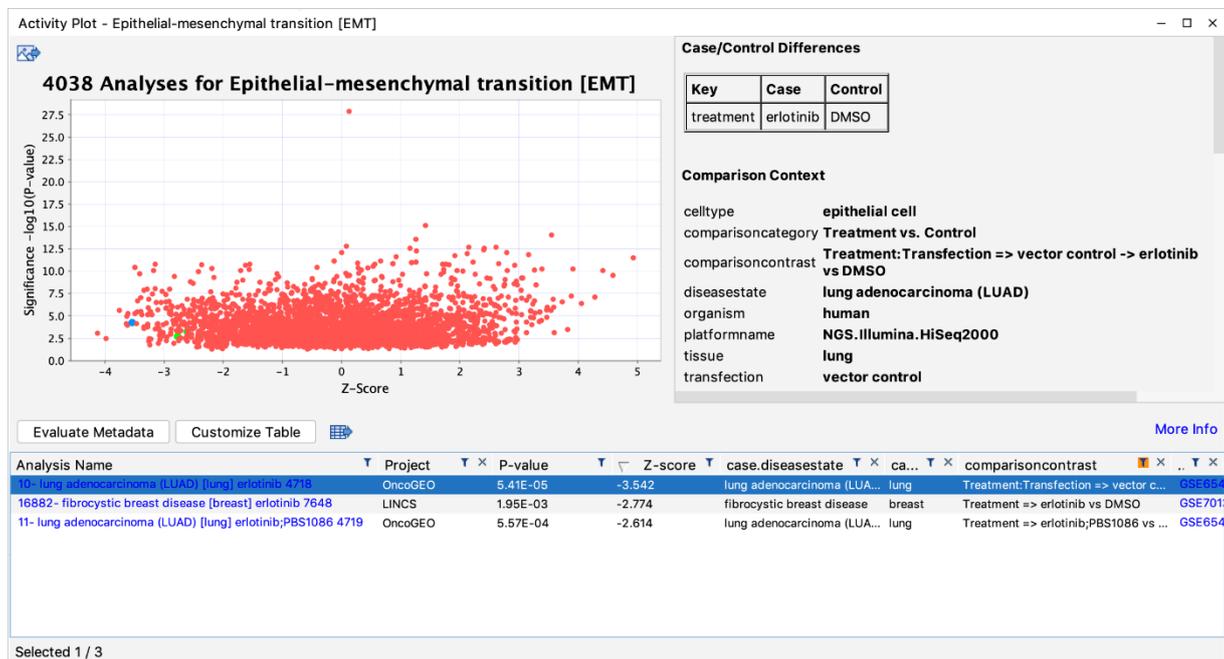


これは、Analysis Match のリポジトリから OmicSoft のすべての分析にわたって、1つのエンティティ(この例では EMT)の予測された活性の散布図です。各赤い点は1つの分析を表しています。分析のエンティティの z スコアが x 軸にプロットされ、p 値の  $-\log$  が y 軸にプロットされています。エンティティの z-スコアが絶対値 1.75 未満、または p 値が 0.05 を超える場合はプロットに表示されないことに注意してください。

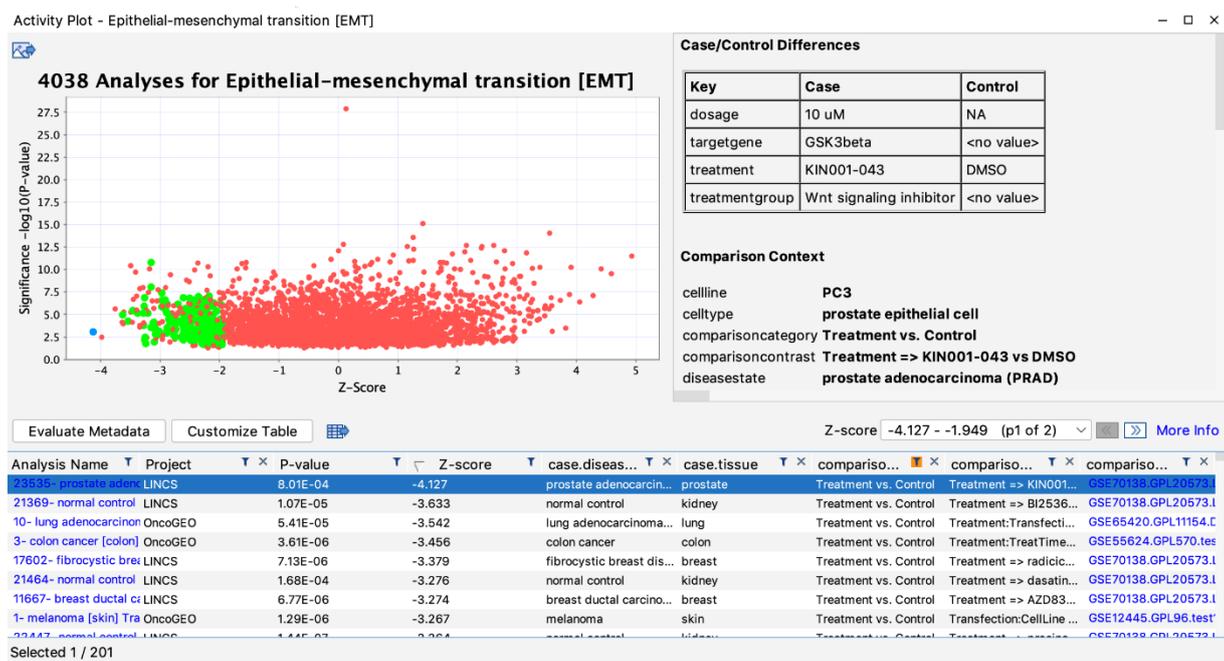


チャートの中で一つの分析を選択すると、チャートの下の表に分析の詳細が表示され、右上に分析の詳細が表示されたメタデータパネルが表示されます。下の例では、EMT の z スコアが強くマイナスにある分析が選択されています(緑色の点)。

**Customize Table** ボタンを押すと、表にメタデータの列を追加することができます。

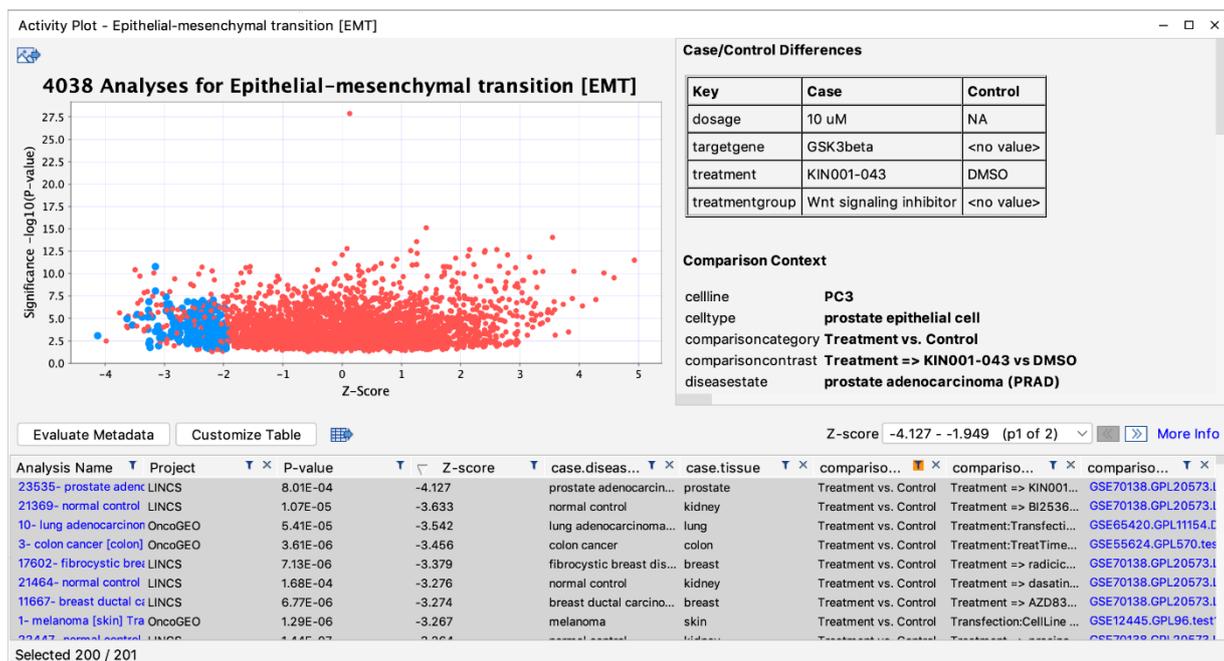


一度に複数の分析を選択したり、テーブル内の漏斗ボタンを使って結果をフィルタリングしたりすることができます。下のスクリーンショットでは、負の z スコア(EMT が阻害されることが予測されている)を持つ分析を選択し、”comparisoncategory” のカラムで”Treatment vs. Control” のフィルタリングを行い、何らかの治療を表す分析に焦点を当てています。表に存在する分析は、チャート内では緑色のドットで示されています。



選択した分析のメタデータを統計的に評価することで、分析データセットに関する関連情報を明らかにすることができます。

テーブル内で分析データセットを選択し(選択した分析のドットが青くなります)、**Evaluate Metadata** ボタンをクリックします。



メタデータのテーブルが表示されます。漏斗ボタンでフィルタリングすることができます。(下図では”case.treatment” フィールドの結果のみを表示するようにフィルタリングされています)

Significant metadata in 200 selected analyses (Epithelial-mesenchymal transition [EMT])

Customize Table 200 repository analyses selected

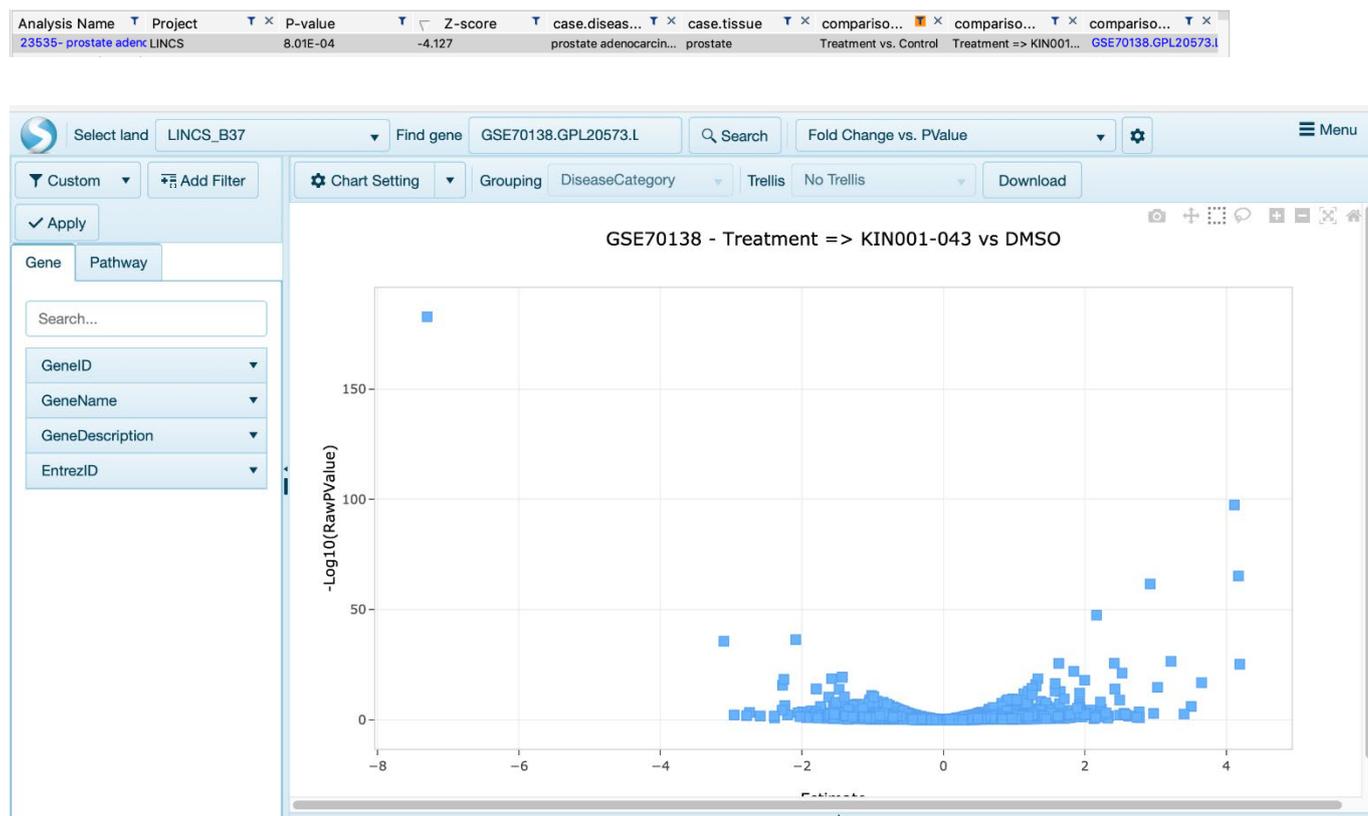
Metadata fi...	Significant ...	p-value	Selected an...	Total analy...	Selected an...	Total analy...
case.treatment	dasatinib	2.38E-11	10	135	200	62813
case.treatment	luminespib	1.04E-07	7	116	200	62813
case.treatment	pracinostat	3.43E-07	6	85	200	62813
case.treatment	NVP-TNKS656;selum...	1.01E-05	2	2	200	62813

Selected 0 / 48

アクティビティプロット左上のエクスポートボタンで、アクティビティプロットを画像(JPEG、PDF、PNG、TIFF、SVG)としてエクスポートすることができます。

テーブル上側のエクスポートボタンで、選択した解析をテキストまたは.xls としてエクスポートすることができます。

Land Explorer のライセンスをお持ちの方は、下図の“comparison ID” 列(右から 2 番め)にある「GSE65 …」のリンクをクリックすると、データセットをボルケーノプロットで表示することができます。



Land Explorer に関する詳細は下記をご参照ください。

<https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/content-exploration-and-databases/qiagen-omicsoft-land-explorer/>

## 解析データセットのレポジトリ

解析データセットはいくつかのプロジェクトフォルダにまとめられています。

それぞれは2020年12月現在では以下のような内訳になっています。

### DiseaseLand

#### HumanDisease (16,290)

- 582 diseases
- 294 tissues
- 68 expression platforms
- 4,394 RNA-seq datasets

#### MouseDisease (14,638)

- 360 diseases
- 248 tissues
- 56 expression platforms
- 7,242 RNA-seq datasets

#### RatDisease (5,480)

- 48 diseases
- 78 tissues
- 1,680 RNA-seq datasets

#### LINCS (28,234)

- 23 cell lines
- 374 chemical treatments or gene overexpression
- 226 different targets (or groups of target genes)

### OncoLand

#### OncoGeo (7,379)

- 169 cancers
- 95 tissues
- 51 expression platforms
- 1,835 RNA-seq datasets

#### TCGA (4,789)

- 33 cancers
- 27 tissues
- 385 different mutation

#### Pediatrics (444)

- 47 cancers
- 23 tissues

#### Metastatic Cancer (81)

- 46 cancers
- 73 tissues
- 196 RNA-seq datasets

#### Hematology (2,511)

- 54 cancers
- 14 tissue types
- 80 cell types
- 495 RNA-seq datasets

#### OncoMouse (933)

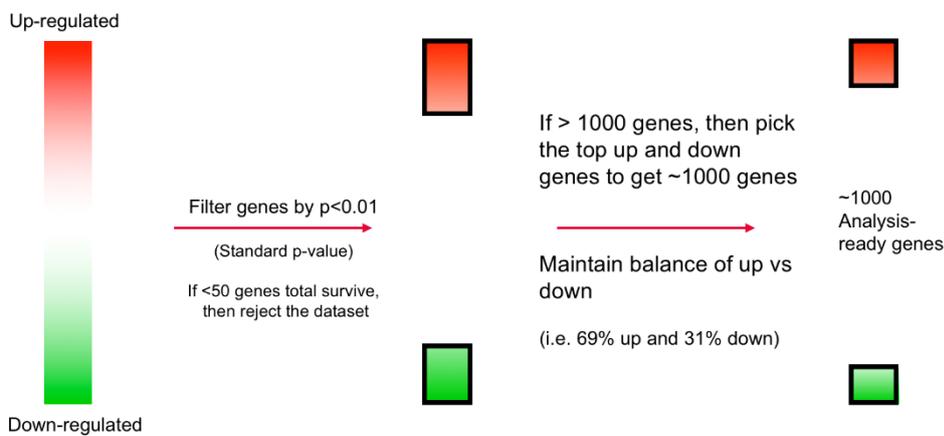
- 46 cancers
- 36 tissues
- 19 expression platforms
- 259 RNA-seq datasets

## データセットの解析方法

公共リポジトリからのデータは完全に再処理、正規化、QA そしてアノテーション付けをされますが、得られたデータセットは、多数の異なる実験デザイン、細胞型、組織、アレイプラットフォーム、および RNA-seq テクノロジーから得られたものです。

IPA でそれらを分析するためには、この多様なレポジトリにまたがって 1 つまたは少数の標準的なカットオフを用いることは不可能であるため、次のような手法を用いて、各データセットに対してなるべく一様な解析となるように遺伝子を取得します。

### Example of Omicsoft Dataset



リポジトリには 56,000 以上のデータセットが含まれており、その大部分には 1000 以上の解析可能な遺伝子が含まれています。データセットは、アップダウンの比率を保ったまま、 $p$  値  $< 0.01$  のカットオフを通過し、1000 未満の遺伝子になるようにフィルタリングされます。

※LINCS では、 $p$  値のしきい値は 0.01 ではなく 0.05 に設定されています。

## シグネチャの作成

データセットが作成された後、IPA は解析用に最大で以下の 4 つのシグネチャの構成セットを作成します。

Canonical Pathways(最大 20 のパスウェイ)

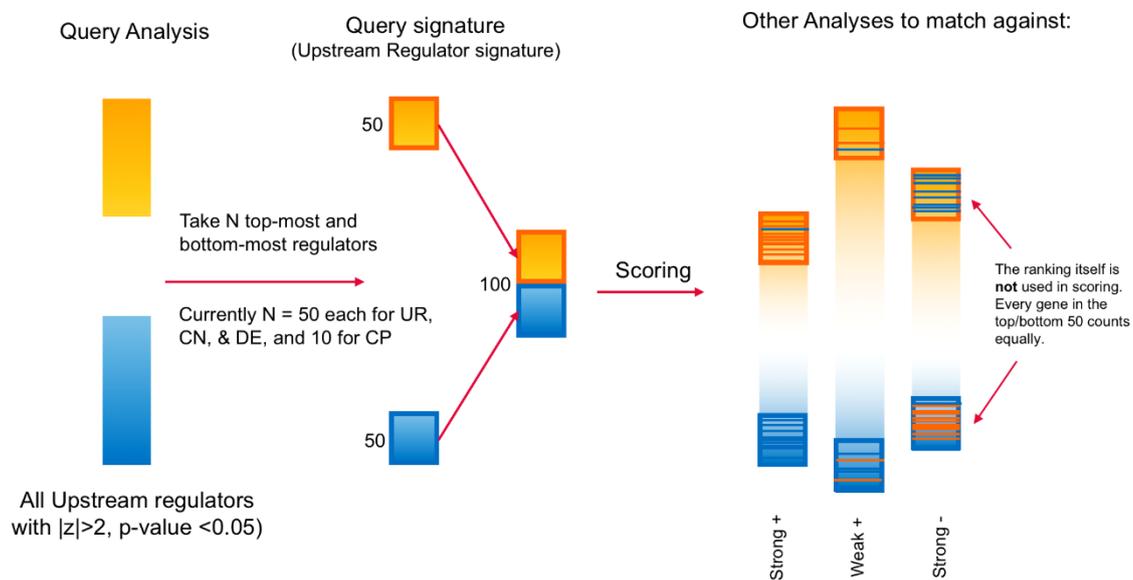
Upstream Regulators(最大 100 のレギュレータ)

Causal Networks(最大 100 のマスターレギュレータ)

Downstream Effects(Diseases&Functions)(最大 100 の疾患または機能)

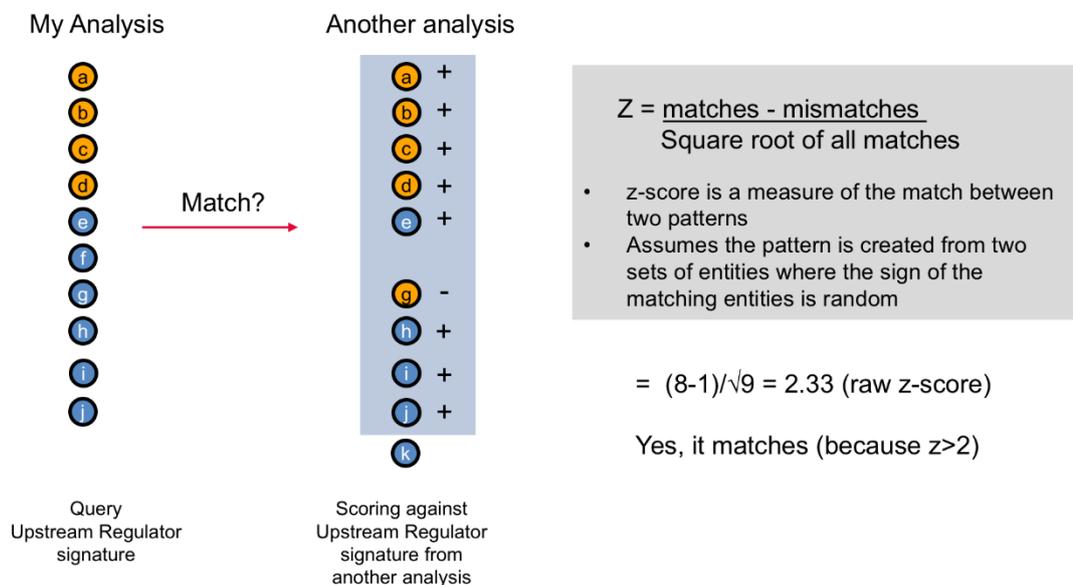
各シグネチャは、下図のように作成されます。

### Example

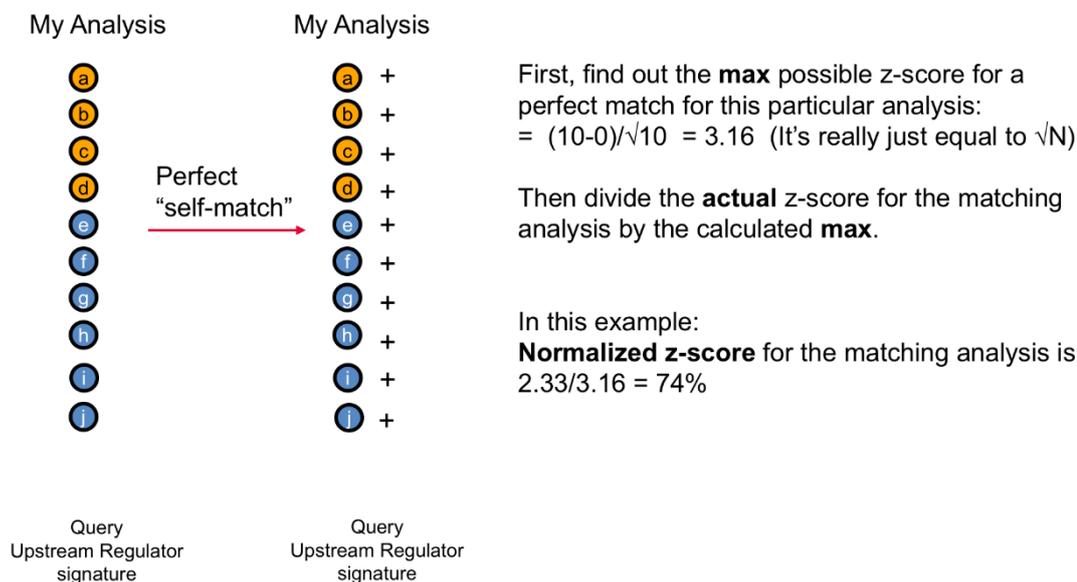


## シグネチャのスコアリング

クエリーのシグニチャと他の解析のシグニチャとの一致について、以下に示すように z スコアを計算します。



ここで計算された Raw z スコアは、Analysis Match タブのデフォルトでは非表示になっています。スコアをより有用にするために、以下のようにスコアは正規化されています。



その結果、実際の一致 (Raw z スコア) が最大値に対するパーセンテージとして計算されます。非常に強く一致するシグネチャでは最大で 80% になりますが、弱く一致するシグネチャは最小で 20% の Normalized z スコアとなります。

## Customize Table でカラムに追加または削除できるメタデータ

case.agecategory	case.tissue	control.gender
case.animalstrain	case.tissuedescription	control.genotype
case.cancergrading	case.tissueregion	control.infection
case.celldescription	case.transfection	control.pretreatment
case.cellline	case.treatment	control.response
case.cellmarkers	case.treatmentstatus	control.sampleids
case.cellsubgroup	case.treattime	control.samplematerial
case.celltype	case.treattime[days]	control.samplepathology
case.clinicalphenotype	case.treattime[hours]	control.samplesource
case.diseaseonsettype	case.treattime[minutes]	control.samplingtime
case.diseasestage	case.treattime[weeks]	control.smokingstatus
case.diseasestate	case.vaccine	control.subjectinfection
case.dosage	case.virusstrain	control.subjecttreatment
case.dosage[gy]	collection	control.tissue
case.ethnicity	comparisoncategory	control.tissuedescription
case.gender	comparisoncontrast	control.transfection
case.genotype	comparisonid	control.treatment
case.infection	comparisonindex	control.treatmentstatus
case.plasmasamplingtime[months]	comparisontype	control.treattime
case.pretreatment	control.agecategory	control.treattime[days]
case.response	control.animalstrain	control.treattime[hours]
case.sampleids	control.cancergrading	control.treattime[minutes]
case.samplematerial	control.celldescription	control.treattime[weeks]
case.samplepathology	control.cellline	control.vaccine
case.samplesource	control.cellmarkers	control.virusstrain
case.samplingtime	control.cellsubgroup	genemodelid
case.samplingtime[dpi]	control.celltype	length
case.samplingtime[hours]	control.clinicalphenotype	observation_name
case.samplingtime[hpi]	control.diseaseonsettype	organism
case.smokingstatus	control.diseasestage	projectname
case.subjectinfection	control.diseasestate	platformname
case.subjecttreatment	control.dosage	sampledatamode
case.symptom	control.dosage[gy]	
	control.ethnicity	

## 製品に関するお問い合わせについて

本マニュアルに記載がされていない詳細な機能に関しては IPA Winter 2020 ユーザーガイドもご参照ください。  
その他、IPA をご利用になられてご不明な点や問題が生じた場合などは弊社サポート窓口までお問合せください。

トミーデジタルバイオロジー株式会社

カスタマーサポート

TEL: 03-6240-0451

e-mail: [support@digital-biology.co.jp](mailto:support@digital-biology.co.jp)

電話サポート対応時間 平日 9:00～17:30