

遺伝子発現DBの使い方 at AJACS安芸

情報・システム研究機構 データサイエンス共同利用基盤施設 ライフサイエンス統合データベースセンター

坊農 秀雅 <http://bonohu.jp/> bono@dbcls.rois.ac.jp <mailto:bono@dbcls.rois.ac.jp>

2016年7月5日

これは統合データベース講習会AJACS安芸「遺伝子発現DBの使い方」の資料です。

概要

本講習は、だれでも自由に使うことができる公共DBやウェブツールを活用して、研究のさまざまな場面で調べることの多い個々の遺伝子発現DBを簡単に調べるための方法と基礎知識について学びます。とくに、需要の増している公共遺伝子発現DBからデータを検索し取得してくる方法について詳しく説明、実習します。また、自ら行なった大規模発現解析の(あるいは公共DBから取得・解析した)結果として得られた数百～数千におよぶ遺伝子セットについて、生物学的な解釈をする方法とその結果の考察を実践します。

講習の流れ

今回の講習では、コンピュータを使って以下の内容について説明します。

- 研究現場で頻繁に使われるDBやツールを知る: 統合TV
- 遺伝子発現DBとは
- 公共の遺伝子発現DBの使い方
 - EBI ArrayExpress
 - 【実習1】 ArrayExpressを使って、興味ある実験データセットを検索する
 - NCBI Gene Expression Omnibus (GEO)
 - AOE (All of gene expression)
 - 【実習2】 AOEを使って、興味ある実験データセットを絞り込む
- 個々の遺伝子の発現プロファイルを調べる
 - RefEx

- 【実習3】RefExを使って、組織特異的遺伝子を検索する
 - 数十～数千の遺伝子群の生物学的解釈
 - DAVID
 - 【実習4】DAVIDを用いて、発現データの結果を生物学的に解釈する
-

講習に際しての注意とお願い

- みんなで同時にアクセスするとサイトにつながりにくくなることが予想されます。
 - 資料を見ながら自力で進められそうな方はどんどん先に、そうでない方は講師と一緒にすすめていきましょう。
 - サイトの反応が悪い時はタイミングをずらして実行してみてください。
 - 反応が無いからと言って何度もクリックするとますます繋がらなくなってしまいます。おおらかな気持ちで臨みましょう。
 - わからないことがあったら挙手にてスタッフにお知らせください。
 - 遠慮は無用です(そのための講習会です!)。おいてけぼりは楽しくありません。
-

いざ講習、その前に

- 次世代シーケンス解析について
 - 「次世代シーケンサーデータの取り扱いと疾患ゲノム解析の基礎@」(2015年9月 AJACS津軽) <http://doi.org/10.7875/togotv.2015.110>
 - さらにもっと基礎から次世代シーケンス解析について知りたい方向け
 - 「バイオインフォマティクス人材育成カリキュラム（次世代シーケンサ）速習コース」の動画 https://www.youtube.com/playlist?list=PL0uaKHgcG00abmj1Nzs1SUhqKLjf_PFBB
-

研究現場で頻繁に使われるデータベースやツールを知る

統合TV <http://togotv.dbcls.jp/ja/>

- 生命科学分野の有用なデータベースやツールの使い方を動画で紹介するウェブサイト
 - <http://togotv.dbcls.jp/> <http://togotv.dbcls.jp/>
-

『統合TV』は、生命科学分野の有用なデータベースやツールの使い方を動画で紹介するウェブサイトです。

目的別に検索

- AJACS講習会資料
 - ゲノム・核酸配列解析
 - タンパク質配列・構造解析
 - 発現制御解析・可視化
 - 文献・辞書・プログラミング
 - 著名データベース
 - 学会講演・講習会
- 関連するタグから検索
- ゲノム (124) 遺伝子 (217)
 - タンパク質 (75)
 - 配列解析 (116)
 - 発現解析 (176) NGS (110)
 - 文献検索 (72) 情報収集 (47)

Q 全番組のリストから、調べたいDBやウェブツールに関するキーワードで検索! (全1080件)

番組のタイトルや画像をクリックすると番組の再生ページへ移動します。

表示件数を選ぶ▼ 検索窓にキーワードを入れると、入力の度ごとに即座に候補の番組が絞り込まれます

GeneTrail2を使って、エンリッチメント解析を行う

マイクロアレイ実験や次世代シーケンシング(NGS)などで得られた大量な遺伝子を扱う際に、どういう機能を持つ遺伝子が変化する傾向にあるかを知るエンリッチメント解析は非常に有用な解析手法です。GeneTrail2は、ウェブ上で多様なエンリッチメント解析ができるツールです。トランскriプトームだけでなく、タンパク質やmiRNA, SNPのエンリッチメント解析をすることもできます。本動画では、GeneTrail2を使ってNCBI GEOに登録されているマイクロアレイデータのエンリッチメント解析を行う方法について紹介します。



無料統計ソフトEZR(Easy R)を使ってマウス操作だけで多彩な統計解析をする～導入・基本編～

Rは、フリーで使える統計解析向けのプログラミング言語です。非常に多彩な解析を行うことができますが、コマンド操作が中心であるため、初心者には敷居が高い面もあります。



<https://gyazo.com/edbabee661757e2a50f6f8ee77c3e778>

- YouTube版もあります <http://youtube.com/togotv/> <http://youtube.com/togotv/>
- ウェブサイトへのアクセスから結果の見方まで、操作の一挙手一投足がわかります。
 - 講義・講習などの参考資料や後輩指導の教材として利用できます。
 - 本講習中、本家サイトが繋がらない時は、統合TVのYouTube版を見ればおよその内容がわかるようになっています。
 - 今回の講習に関連する内容の多くは、「発現解析」タグのついた動画があります。
 - 過去の講習会の内容はそのほとんどが統合TVに収録されており、いつでもどこでも繰り返し復習できるようになっています。
- お探しの動画が見つからない or 統合TV未掲載の場合は、統合TV番組リクエストフォーム <http://togotv.dbcls.jp/ja/contact.html>へどうぞ!
- 統合TVを作ってくれる方、募集中!!

<https://twitter.com/bonohu/status/747954940157399040>

遺伝子発現DBとは

mind the gap

遺伝子発現解析→DB(というか、公共アーカイブ)に登録

- 日本では論文投稿前が現在一般的
- 欧米では研究費の条件で多くの場合データを出したらわりとすぐ

その歴史

- マイクロアレイの発明→網羅的遺伝子発現定量が可能に→遺伝子発現DB
 - アレイのイメージデータ(CELファイルなど)
 - 定量データ(Series Matrix File)
 - DDBJ CIBEX <http://cibex.nig.ac.jp/> (更新停止)
 - NCBI Gene Expression Omnibus (GEO) <http://www.ncbi.nlm.nih.gov/geo/>
 - EBI ArrayExpress <http://www.ebi.ac.uk/arrayexpress/>
 - ArrayExpressはGEOのデータを定期的に取り込み続けている→こちらの使用を推奨
- 次世代シークエンサー(NGS: Next Generation Sequencer)の発明→個々のサンプルでの transcript sequencingが現実的に
 - データはNGS配列DB(SRA: Sequence Read Archive) and/or 遺伝子発現DBへ?
 - NCBI,EBIでは遺伝子発現DB(それぞれGEO,ArrayExpress)に登録すれば、配列データ(FASTQファイル)がSRAにも登録される状況
 - DDBJ(日本)もその仕組みを現在構築中
 - 現状、どこかの公共DBに登録されていればOK

発現定量のステップ

マイクロアレイ

マイクロアレイ解析



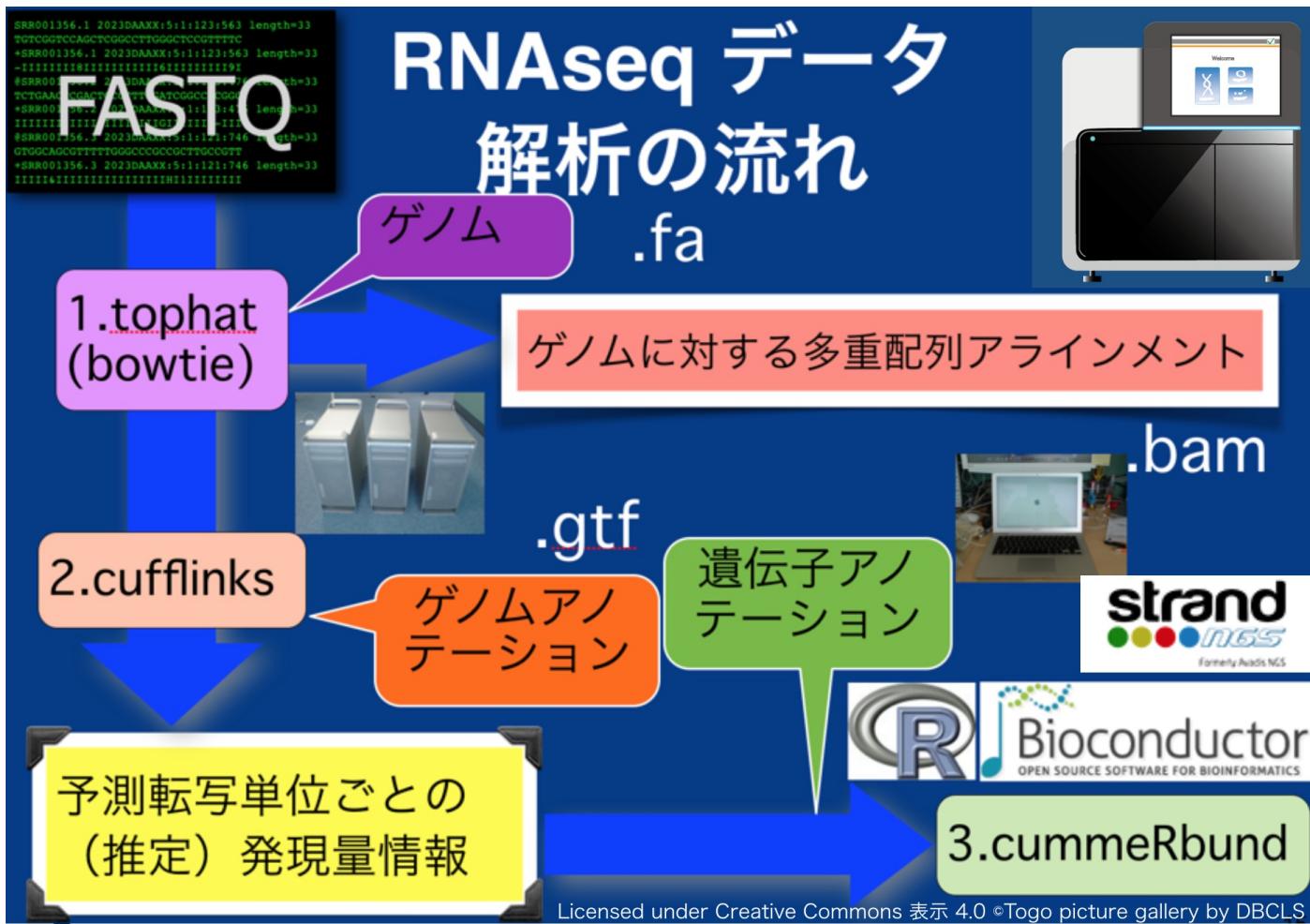
遺伝子アノ
テーション

oligoprobeに対応する
遺伝子ごとの発現量

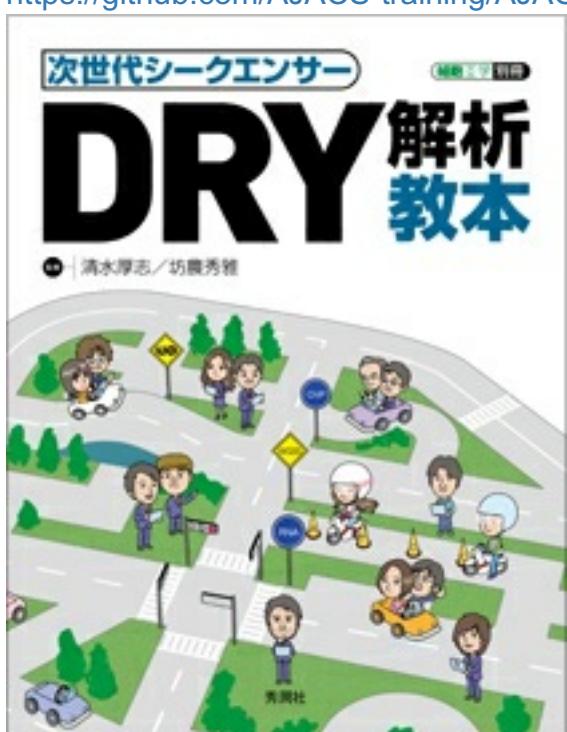


<https://gyazo.com/a5ea2e97d8fb84e7f21ecc1bb4b878cb>

NGS(RNA-seq)



詳しい手順は、次世代シークエンサーDRY解析教本「Level2② 発現解析」
<https://github.com/AJACS-training/AJACSA2p97-141>などを参考に



<http://gakken-mesh.jp/book/detail/9784780909203.html>

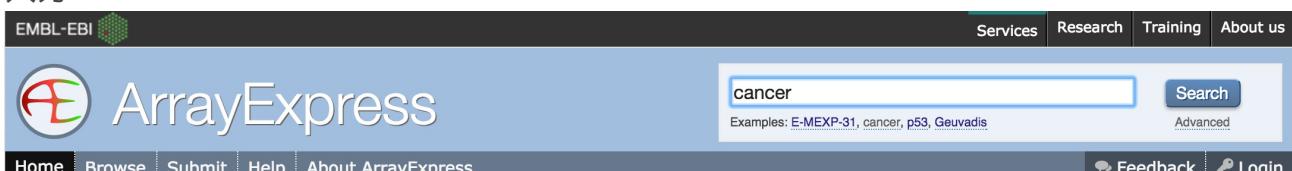
公共の遺伝子発現DBの使い方

EBI ArrayExpress

- EBIが提供・維持管理している遺伝子発現情報のデータベース
 - <http://www.ebi.ac.uk/arrayexpress/>
 - NCBI GEOのデータも取り込んでいて、こちらのほうがデータ数が多い
 - インターフェースが使いやすい

【実習1】 ArrayExpressを使って、興味あるマイクロアレイの実験データセットを検索する

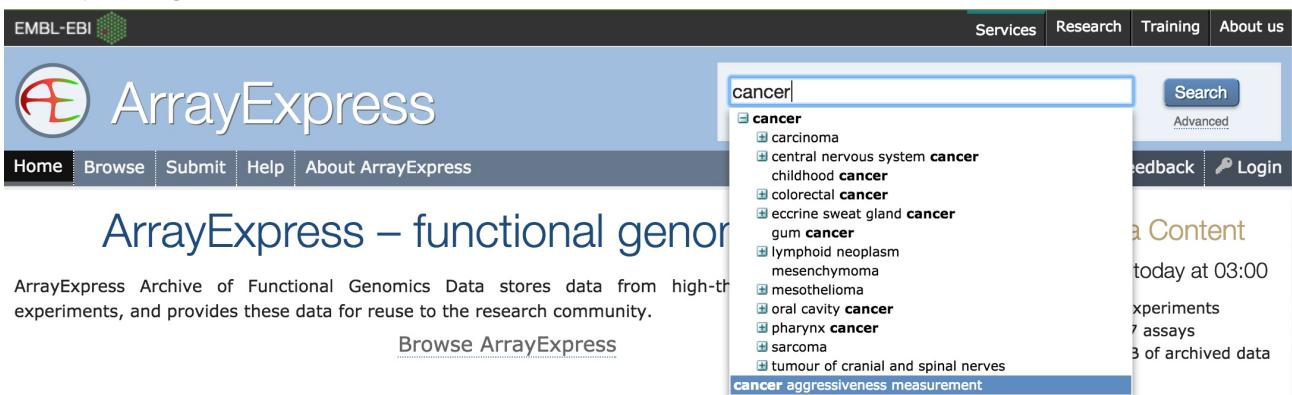
1. ArrayExpressのサイト <http://www.ebi.ac.uk/arrayexpress/>にアクセス、検索窓に'cancer'と入力



The screenshot shows the ArrayExpress homepage. At the top, there's a navigation bar with links for Services, Research, Training, and About us. Below that is a main header with the EMBL-EBI logo, the ArrayExpress logo, and a search bar containing the word 'cancer'. Below the search bar, there's a link to 'Advanced' search options. The main content area has a title 'ArrayExpress – functional genomics data' and a brief description of what the archive contains. There's also a 'Browse ArrayExpress' link. On the right side, there's a sidebar titled 'Data Content' with statistics: 'Updated today at 03:00', '66349 experiments', '2019317 assays', and '43.35 TB of archived data'.

<https://gyazo.com/c0cf9c24aff417fa51d24fc645f76f31>

2. 候補となるキーワードがいろいろ出てきました。cancerの左にある'+'をクリックして聞いてみましょう



This screenshot shows the same ArrayExpress homepage as above, but with the search results expanded. The search bar still contains 'cancer', and a dropdown menu on the left lists various cancer types: carcinoma, central nervous system cancer, childhood cancer, colorectal cancer, eccrine sweat gland cancer, gum cancer, lymphoid neoplasm, mesenchymoma, mesothelioma, oral cavity cancer, pharynx cancer, sarcoma, and tumour of cranial and spinal nerves. Below this list, there's a link to 'cancer aggressiveness measurement'. The sidebar on the right remains the same, showing updated statistics.

<https://gyazo.com/2e716af19911031e5ba8eba07d1dad37>

3. 今回はそのまま'Search'ボタンを押して、cancerで検索してみます
4. 黄色でハイライトされた'cancer'以外に、オレンジ色でcancer関係のキーワードもハイライトされているのを確認しましょう

Page 1 2 3 4 5 6 .. 608			Showing 1 - 25 of 15189 experiments				Page size 25 50 100 250 500				
Accession	Title	Type	Organism	Assays	Released	Processed	Raw	Views	Atlas		
E-GEOD-71472	Effect of DNMT3A R882H mutation or WT expression on epigenetic landscapes of hematopoietic stem/progenitor cells with NRAS G12D co-transduction (ChIP-seq)	ChIP-seq	Mus musculus	13	23/06/2016		-	-	-		
E-GEOD-71439	Expression profiling of murine leukemia stem cell (LSC) lines established ex vivo by coexpression of R882H-mutated DNMT3A and NRAS-G12D post treatment with Dot1l inhibitor (Microarray)	transcription profiling by array	Mus musculus	10	23/06/2016			-	-		
E-GEOD-71437	Effect of DNMT3A R882H mutation or WT on gene expression in hematopoietic stem/progenitor cells with NRAS G12D co-transduction (Microarray)	transcription profiling by array	Mus musculus	12	23/06/2016			-	-		
E-GEOD-39292	Oncogenic EGFR Represses the TET1 DNA Demethylase to Induce Silencing of Tumor Suppressors in Cancer Cells	transcription profiling by array	Homo sapiens	12	23/06/2016		-	-	-		
E-MTAB-2672	Functional and genetic heterogeneity of distinctive leukaemic stem cell populations in CD34+ human Acute Myeloid Leukaemia	RNA-seq of coding RNA	Homo sapiens	124	22/06/2016	-		23	-		
E-GEOD-83607	Loss of the DNA methyltransferase MET1 induces H3K9 hypermethylation at P ^c G target genes in Arabidopsis, and relocation of H3K27 trimethylation to transposons [RNA-seq]	RNA-seq of coding RNA	Arabidopsis thaliana	2	22/06/2016		-	-	-		
E-GEOD-83606	Loss of the DNA methyltransferase MET1 induces H3K9 hypermethylation at P ^c G target genes in Arabidopsis, and relocation of H3K27 trimethylation to	ChIP-chip by tiling array	Arabidopsis thaliana	6	22/06/2016			-	-		

<https://gyazo.com/ae8fe04aad31e08c299ede4e25d131de>

5. 単にcancerではヒット数が多いので、'oral cancer'で検索してみましょう

EMBL-EBI

Services | Research | Training | About us

ArrayExpress

Search results for **oral cancer**

Examples: E-MEXP-31, cancer, p53, Geuvadis

Feedback | Login

Filter search results

Show more data from EMBL-EBI

Page 1 2 3 4 5 6 .. 10			Showing 1 - 25 of 231 experiments				Page size 25 50 100 250 500				
Accession	Title	Type	Organism	Assays	Released	Processed	Raw	Views	Atlas		
E-GEOD-73521	NUPR1 a new target in liver cancer: implication in controlling cell growth, migration, invasion and sorafenib resistance	transcription profiling by array	Homo sapiens	2	15/06/2016			2	-		
E-MTAB-4863	Copy number Analysis of Head and Neck cancer cases from two high incidence regions	genotyping by array	Homo sapiens	225	14/06/2016	-		-	-		
E-GEOD-75540	A minimal DNA 5-methylcytosine signature of oral tongue squamous cell carcinoma	transcription profiling by array, methylation profiling by array	Homo sapiens	151	01/06/2016		-	29	-		
E-GEOD-75539	A minimal DNA 5-methylcytosine signature of oral tongue squamous cell carcinoma [HumanHT-12 V4.0]	transcription profiling by array	Homo sapiens	15	01/06/2016		-	10	-		
E-GEOD-75538	A minimal DNA 5-methylcytosine signature of oral tongue squamous cell carcinoma [HumanHT-12 WG-DASL V4.0]	transcription profiling by array	Homo sapiens	28	01/06/2016		-	9	-		
E-GEOD-75537	A minimal DNA 5-methylcytosine signature of oral tongue squamous cell carcinoma [methylation]	methylation profiling by array	Homo sapiens	108	01/06/2016		-	11	-		

<https://gyazo.com/8192a0e6a879f679e79203bff1f65a15>

6. デフォルトでは'Released'(データの公開日)が新しいものから古いもの順にならんでいます。'Views'(閲覧数)をクリックして並べ替えてみましょう



Filter search results

Search results for oral cancer

Show more data from EMBL-EBI

Page 1 2 3 4 5 6 .. 10

Showing 1 - 25 of 231 experiments

Page size 25 50 100 250 500

Accession	Title	Type	Organism	Assays	Released	Processed	Raw	Views	Atlas
E-MTAB-62	Human gene expression atlas of 5372 samples representing 369 different cell and tissue types, disease states and cell lines	transcription profiling by array	Homo sapiens, Mus musculus	5372	04/08/2010			41982	-
E-MTAB-2706	RNA-seq of 675 commonly used human cancer cell lines	RNA-seq of coding RNA	Homo sapiens	675	25/11/2014	-		12209	
E-TABM-185	Transcription profiling by array of integrated human experiments involving the hgu133a platform to investigate a global map of human gene expression	transcription profiling by array	Homo sapiens	5896	20/11/2007			9474	-
E-MTAB-3732	A comprehensive human expression map	transcription profiling by array	Homo sapiens	27871	23/07/2015		-	9205	-
E-GEOD-7307	Transcription profiling by array of human samples with various different diseases	transcription profiling by array	Homo sapiens	677	20/06/2008			7249	-
E-TABM-302	Transcription profiling of human head and neck squamous cell carcinomas to identify metastasis predictors and descriptors, and differentiation classifiers	transcription profiling by array	Homo sapiens	81	23/08/2007			1777	-

<https://gyazo.com/42740cf9772175e373ac4c81dafcfc31>

7. 詳細を見るには'Accession'のリンクをクリックします。番号の意味はこちらを参照
http://www.ebi.ac.uk/arrayexpress/help/accession_codes.html

Home Browse Submit Help About ArrayExpress Feedback Login

ArrayExpress > Search results for "oral cancer" > E-MTAB-62

E-MTAB-62 - Human gene expression atlas of 5372 samples representing 369 different cell and tissue types, disease states and cell lines

Status	Released on 4 August 2010, last updated on 2 May 2014
Organism	Homo sapiens, Mus musculus
Samples (5372)	Click for detailed sample information and links to data ↗ found inside: 600MPE breast cancer, a498 renal cell carcinoma, A549 lung adenocarcinoma, A673 Ewing tumor, ABC-1 lung adenocarcinoma, acute lymphoblastic leukemia, acute megakaryoblastic leukaemia, acute monoblastic and monocytic leukemia, acute myeloid leukemia, acute myelomonocytic leukemia, acute promyelocytic leukemia, alveolar rhabdomyosarcoma, AMO-1 myeloma, AU565 breast carcinoma, B-cell lymphoma, breast cancer, leukemia, nervous system neoplasm, non breast carcinoma, sarcoma, ...
Array (1)	A-AFFY-33 - Affymetrix GeneChip Human Genome HG-U133A [HG-U133A]
Protocols (1)	Click for detailed protocol information
Description	The experiment contains systematically annotated and consistently normalized human gene expression data matrix of 5372 samples integrated from 206 public experiments of a HG-U133A array platform. The dataset is a subset of a larger pool of 9004 samples gathered from ArrayExpress and GEO websites and checked for quality assessment and suitability for data co-analysis as described by Bolstad et al 2005. The sample annotations have been subject to semiautomatic curation and manual generalization of 369 biological groups which have been additionally organized into blood/non-blood, 14 and 15 meta-classes.
Experiment types	transcription profiling by array, co-expression, data integration
Contact	Margus Lukk <margus.lukk@cruk.cam.ac.uk>
Citations	Multiple receptor tyrosine kinases promote the <i>in vitro</i> phenotype of metastatic human osteosarcoma cell lines. Rettew AN, Young ED, Lev DC, Kleinerman ES, Abdul-Karim FW, Getty PJ, Greenfield EM. <i>Oncogenesis</i> 1:e34 (2012), Europe PMC 23552467 Coevolution reveals a network of human proteins originating with multicellularity. Bezginov A, Clark GW, Charlebois RL, Dar VU, Tillier ER. <i>Mol Biol Evol</i> 30(2):332-346 (2013), Europe PMC 22977115

<https://gyazo.com/2151dac0cce22ae9db7eb8814f72ea45>

8. ブラウザのバックボタンで戻ります。'Title'には研究内容のタイトル、'Type'には実験の種類、'Organism'には生物種が書かれています



Filter search results

Search results for oral cancer

Show more data from EMBL-EBI

Page 1 2 3 4 5 6 .. 10

Showing 1 - 25 of 231 experiments

Page size 25 50 100 250 500

Accession	Title	Type	Organism	Assays	Released	Processed	Raw	Views	Atlas
E-MTAB-62	Human gene expression atlas of 5372 samples representing 369 different cell and tissue types, disease states and cell lines	transcription profiling by array	Homo sapiens, Mus musculus	5372	04/08/2010			41982	-
E-MTAB-2706	RNA-seq of 675 commonly used human cancer cell lines	RNA-seq of coding RNA	Homo sapiens	675	25/11/2014	-		12209	
E-TABM-185	Transcription profiling by array of integrated human experiments involving the hgu133a platform to investigate a global map of human gene expression	transcription profiling by array	Homo sapiens	5896	20/11/2007			9474	-
E-MTAB-3732	A comprehensive human expression map	transcription profiling by array	Homo sapiens	27871	23/07/2015		-	9205	-
E-GEOD-7307	Transcription profiling by array of human samples with various different diseases	transcription profiling by array	Homo sapiens	677	20/06/2008			7249	-
E-TABM-302	Transcription profiling of human head and neck squamous cell carcinomas to identify metastasis predictors and descriptors, and differentiation classifiers	transcription profiling by array	Homo sapiens	81	23/08/2007			1777	-

<https://gyazo.com/42740cf9772175e373ac4c81dafcfc31>

- 'Processed'にアイコンがあるものは解析済みデータがあることを、'Raw'にアイコンがあるものは生データがダウンロード可能であることを示しています
- アイコンが下向き矢印のものは直接ダウンロードが始まり、クリップのものはダウンロードのできるサイトにジャンプします



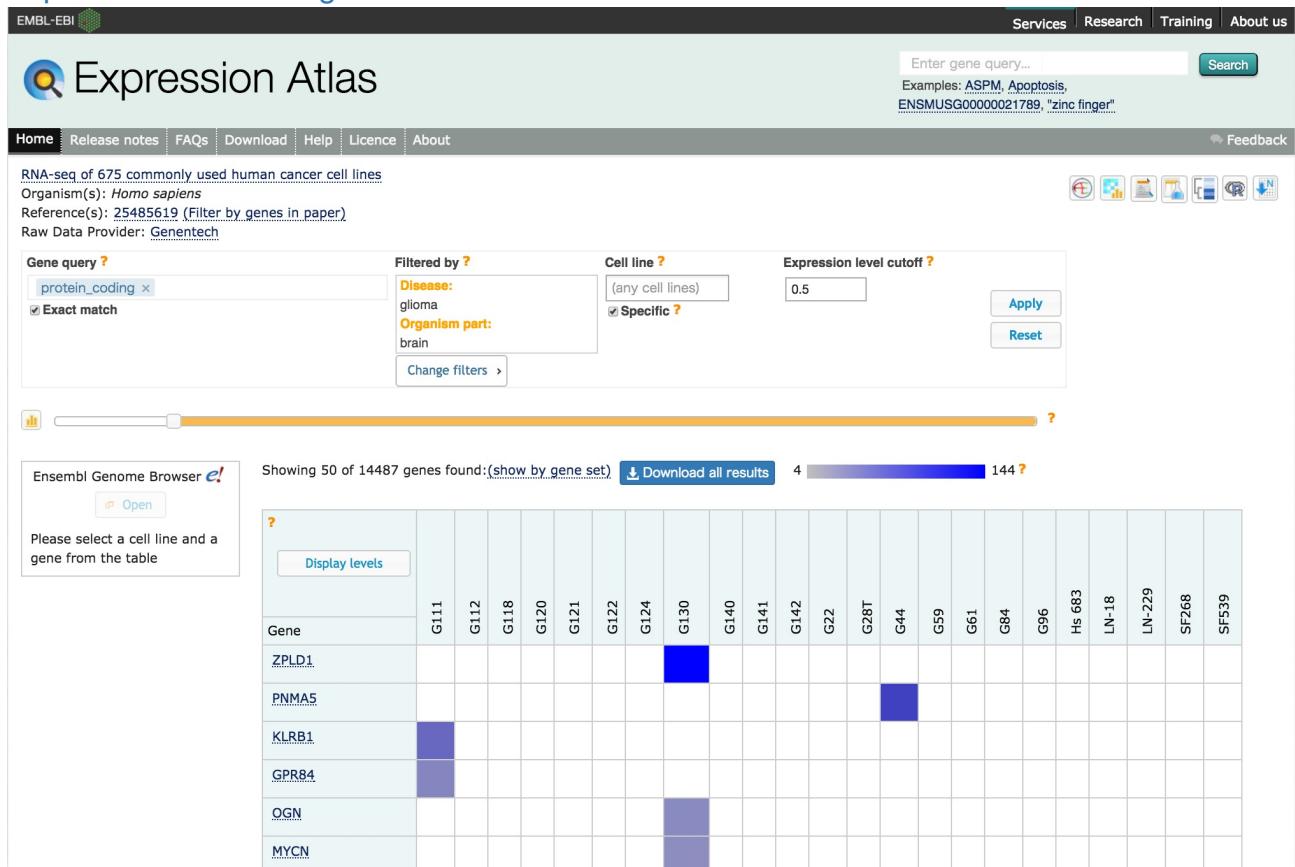
ArrayExpress > Browse > E-MTAB-62 > Files > Raw data

E-MTAB-62 - Human gene expression atlas of 5372 samples representing 369 different cell and tissue types, disease states and cell lines

E-MTAB-62.raw.1.zip	154.7 MB	31 January 2012, 14:35
E-MTAB-62.raw.2.zip	148.6 MB	31 January 2012, 14:36
E-MTAB-62.raw.3.zip	149.7 MB	31 January 2012, 14:38
E-MTAB-62.raw.4.zip	149.6 MB	31 January 2012, 14:39
E-MTAB-62.raw.5.zip	151.0 MB	31 January 2012, 14:40
E-MTAB-62.raw.6.zip	150.5 MB	31 January 2012, 14:42
E-MTAB-62.raw.7.zip	150.9 MB	31 January 2012, 14:43
E-MTAB-62.raw.8.zip	153.3 MB	31 January 2012, 14:45
E-MTAB-62.raw.9.zip	150.5 MB	31 January 2012, 14:46
E-MTAB-62.raw.10.zip	150.3 MB	31 January 2012, 14:31
E-MTAB-62.raw.11.zip	150.1 MB	31 January 2012, 14:33
E-MTAB-62.raw.12.zip	181.5 MB	31 January 2012, 14:33
E-MTAB-62.raw.13.zip	177.4 MB	31 January 2012, 14:34
E-MTAB-62.raw.14.zip	154.3 MB	31 January 2012, 14:34
E-MTAB-62.raw.15.zip	152.3 MB	31 January 2012, 14:34
E-MTAB-62.raw.16.zip	152.4 MB	31 January 2012, 14:34
E-MTAB-62.raw.17.zip	152.0 MB	31 January 2012, 14:34

<https://gyazo.com/1ba9b3841322ccfb0836ae4baf01ee40>

- ブラウザのバックボタンで戻ります。'Atlas'にアイコンがあるものはExpression Atlas



<https://gyazo.com/29b979ffee1e1605a6a8cd023af0639a>

12. 自分の研究テーマに近い、また興味のあるマイクロアレイデータが利用可能か検索してみましょう。

NCBI Gene Expression Omnibus (GEO)

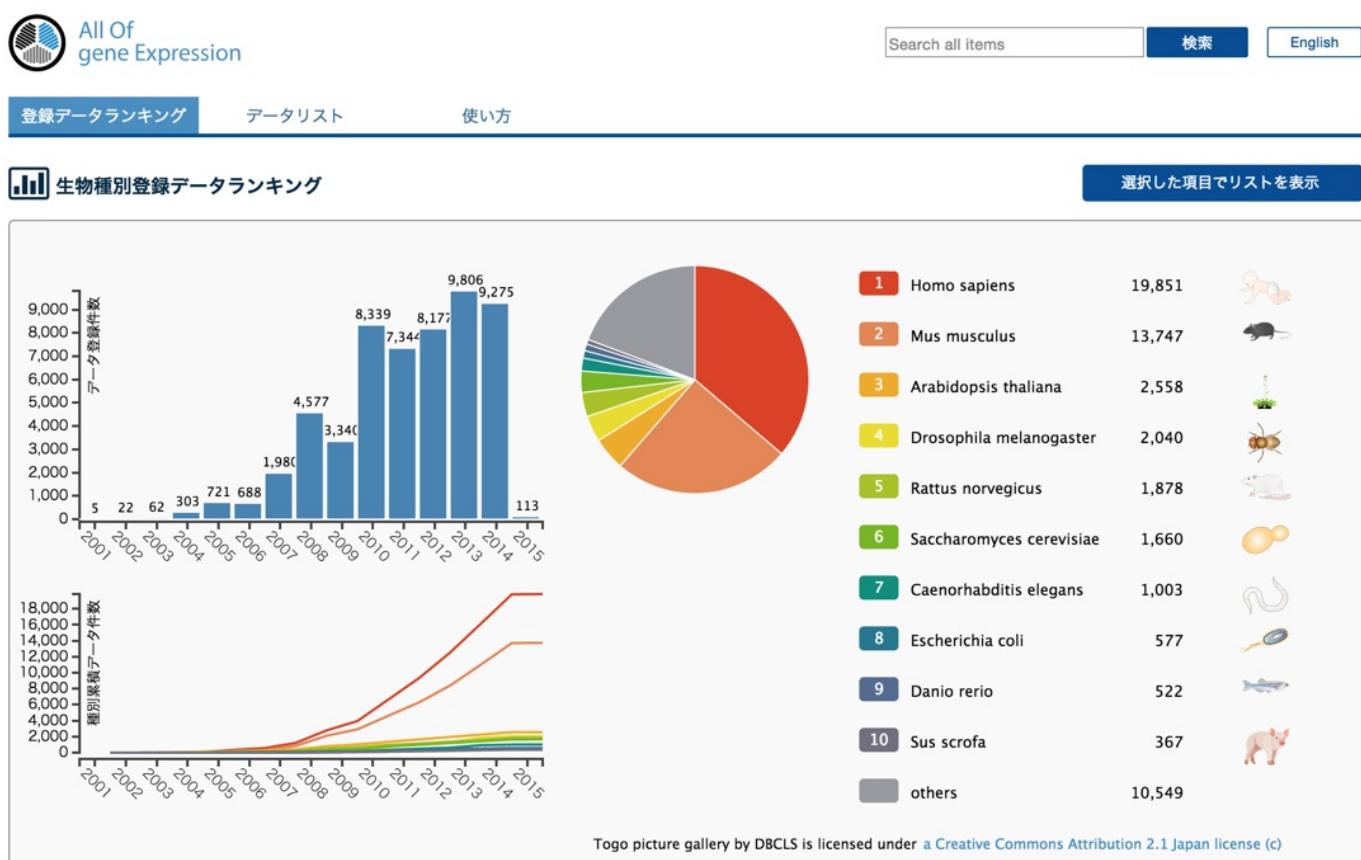
- NCBIが提供・維持管理している遺伝子発現情報（主にマイクロアレイ）のデータベース
- <http://www.ncbi.nlm.nih.gov/geo/> <http://www.ncbi.nlm.nih.gov/geo/>
- 自分の興味のある発現データセットや遺伝子プロファイルを検索することができるだけでなく、それらの生データを自由にダウンロードすることができます。
 - GEOのエントリについて（GEO ID番号の最初の3文字の意味）
 - GPL: Platform — マイクロアレイチップの種類
 - GSM: Sample — 1枚のマイクロアレイチップから得られたサンプルデータ
 - GSE: Series — 1つの実験で得られたGSMのセット
 - GDS: DataSet — NCBIのスタッフが解析可能なデータを集めて再編成したGSMのセット
- 【参考】NCBI GEOの使い方1～マイクロアレイデータの検索・取得～ 2011
<http://doi.org/10.7875/togotv.2011.077>

AOE(All of gene Expression) <http://aoe.dbcls.jp/>

- 遺伝子発現用のデータ目次
 - 年ごと、生物種ごとにヒストグラム表示
- <http://AOE.dbcls.jp/>
- マイクロアレイ(Affymetrix, Agilent, それ以外) + RNA-seq(Illumina, それ以外)
- キーワード検索も

【実習2】 AOEを使って、低酸素関連遺伝子発現データを検索する

1. <http://AOE.dbcls.jp/> <http://AOE.dbcls.jp/> にアクセスします



<http://gyazo.com/263a82cf4448166f348d839e89e58ae1>

2. 「生物種別登録データランク」で生物種のところ(6. *Saccharomyces cerevisiae*)をクリックすると、その生物種だけのデータ登録数に絞りこめます

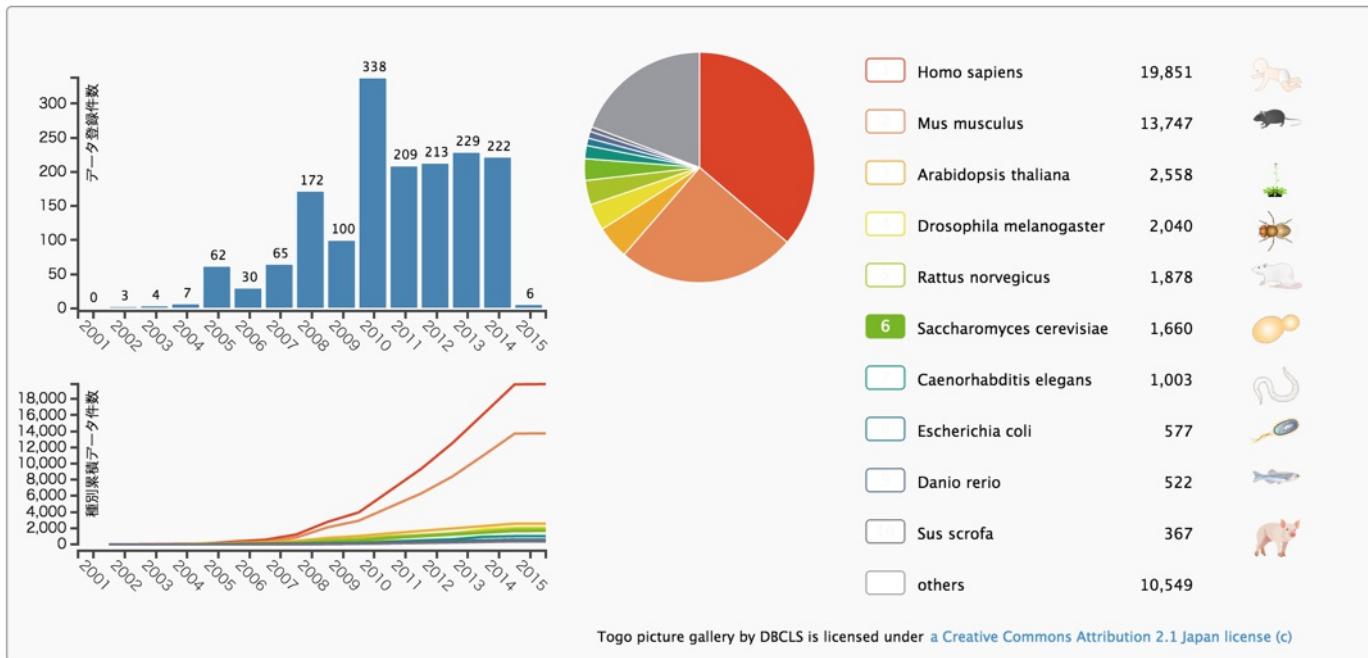
登録データランキング

データリスト

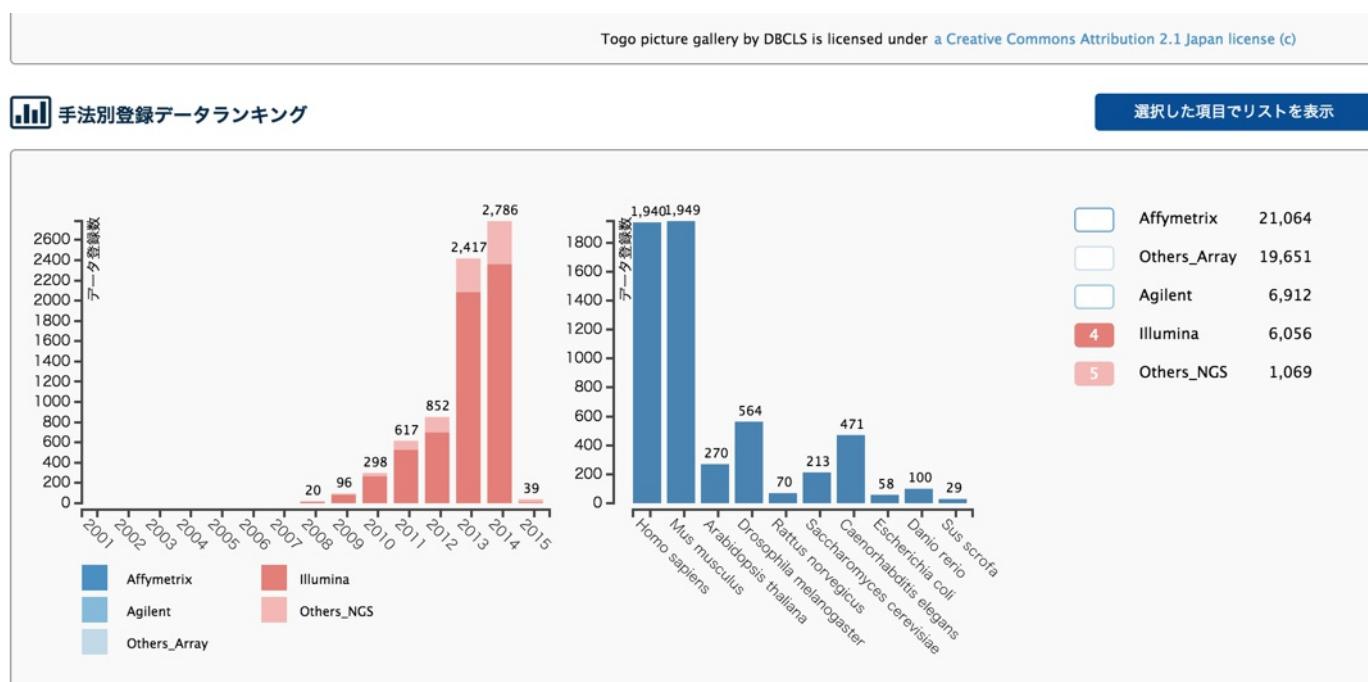
使い方

 生物種別登録データランキング

選択した項目でリストを表示


<http://gyazo.com/a486da624a97278a524cb282b42aeebd>

3. 「手法別登録データランキング」で赤系の色(4. illuminaと5. Other_NGS)だけクリックして残すと、NGSによる遺伝子発現測定のデータのみに絞りこめます


<http://gyazo.com/b69d96059e492c5da246e30504e8d023>

4. 上部の検索窓でキーワード(例えば、``hypoxia)を入れて検索ボタンを押すと、そのキーワード

を実験タイトルに含むエントリだけに絞りこめます



hypoxia

検索

English

登録データランキング

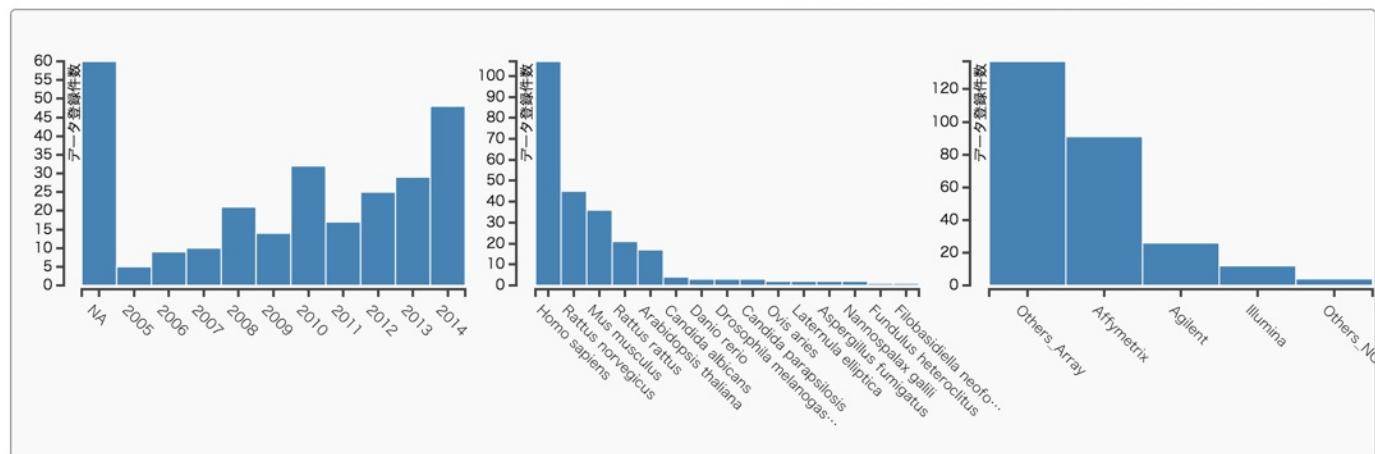
データリスト

使い方

hypoxia の検索結果 (270件)

選択した項目でリストを表示

検索結果をクリアする



生物種別登録データランキング

選択した項目でリストを表示

<http://gyazo.com/e307a5a241dab93eeccc3e9109faa803>

【参考】AOEを使って遺伝子発現データベースの統計を見ながら検索する
<http://doi.org/10.7875/togtv.2014.096>

個々の遺伝子の発現プロファイルを調べる

RefEx (Reference Expression dataset) <http://refex.dbcls.jp/>

- ヒト、マウス、ラットの遺伝子発現情報リファレンスデータセット
- <http://refex.dbcls.jp/>
- 4つの異なる実験手法 (EST、GeneChip、CAGE、RNA-seq) によって得られた正常組織、初代培養細胞、細胞株における遺伝子発現データを検索、閲覧可能
 - 最近新たに、FANTOM5 CAGEデータが追加(ヒト556種、マウス286種)
 - 掲載しているデータやオリジナルデータなどの詳細については、RefExについて <http://refex.dbcls.jp/about.php?lang=ja>

- RefExで掲載されているデータはすべて再利用可能
 - 「RefEx analysis」として論文に引用していただいたケースも
 - Aberrant IDH3 α expression promotes malignant tumor growth by inducing HIF-1-mediated metabolic reprogramming and angiogenesis, *Oncogene*, (22 December 2014) | doi:10.1038/onc.2014.411
<http://www.nature.com/onc/journal/vaop/ncurrent/full/onc2014411a.html>
- このツールでできること
 - 正常組織における遺伝子発現データを調べる
 - 測定手法による遺伝子発現量の差異を比較する
 - 組織特異的遺伝子をワンタッチで検索可能
 - 遺伝子発現解析などで見出された不詳な遺伝子群の機能および関係性を調べる

【実習3】RefExを使って、組織特異的遺伝子を検索する

- 【復習用 統合TV】RefExの使い方 <http://doi.org/10.7875/togotv.2014.009>
1. <http://refex.dbcls.jp/> <http://refex.dbcls.jp/> を開きます。
 2. 画面中央の「組織特異的に発現する遺伝子を見る」の臓器アイコンにカーソルを合わせると、更に詳細な部位のアイコンが出るので、調べたい臓器（例として肝臓）をクリックします。

Reference
Expression
Dataset
[English](#) | [日本語](#)



臓器ごとの発現比較を
4つの実験手法と
ボディーパーツ3Dで。
[▼もっと詳しく](#)



[キーワードで検索](#)

検索

ex) troponin, ALB

組織特異的に発現する
遺伝子を見る



臓器のアイコンをマウスオーバー

遺伝子オントロジー
Gene Ontology

- [cellular process](#)
- [biological regulation](#)
- [metabolic process](#)
- [multicellular organismal process](#)
- [response to stimulus](#)
- [developmental process](#)

他のオントロジーを選ぶ

遺伝子ファミリー
InterPro

- [RNA recognition motif, RNP-1](#)
- [Pleckstrin homology](#)
- [Krueppel-associated box](#)
- [Protein kinase-like domain](#)
- [Zinc finger,C2H2-like](#)
- [GPCR,rhodopsin-like superfamily](#)

他のファミリーを選ぶ

染色体

染色体領域を選ぶ

[Advanced Search](#)

Advanced Search

[ページ上部に戻る](#)

RefExについて

RefExの使い方

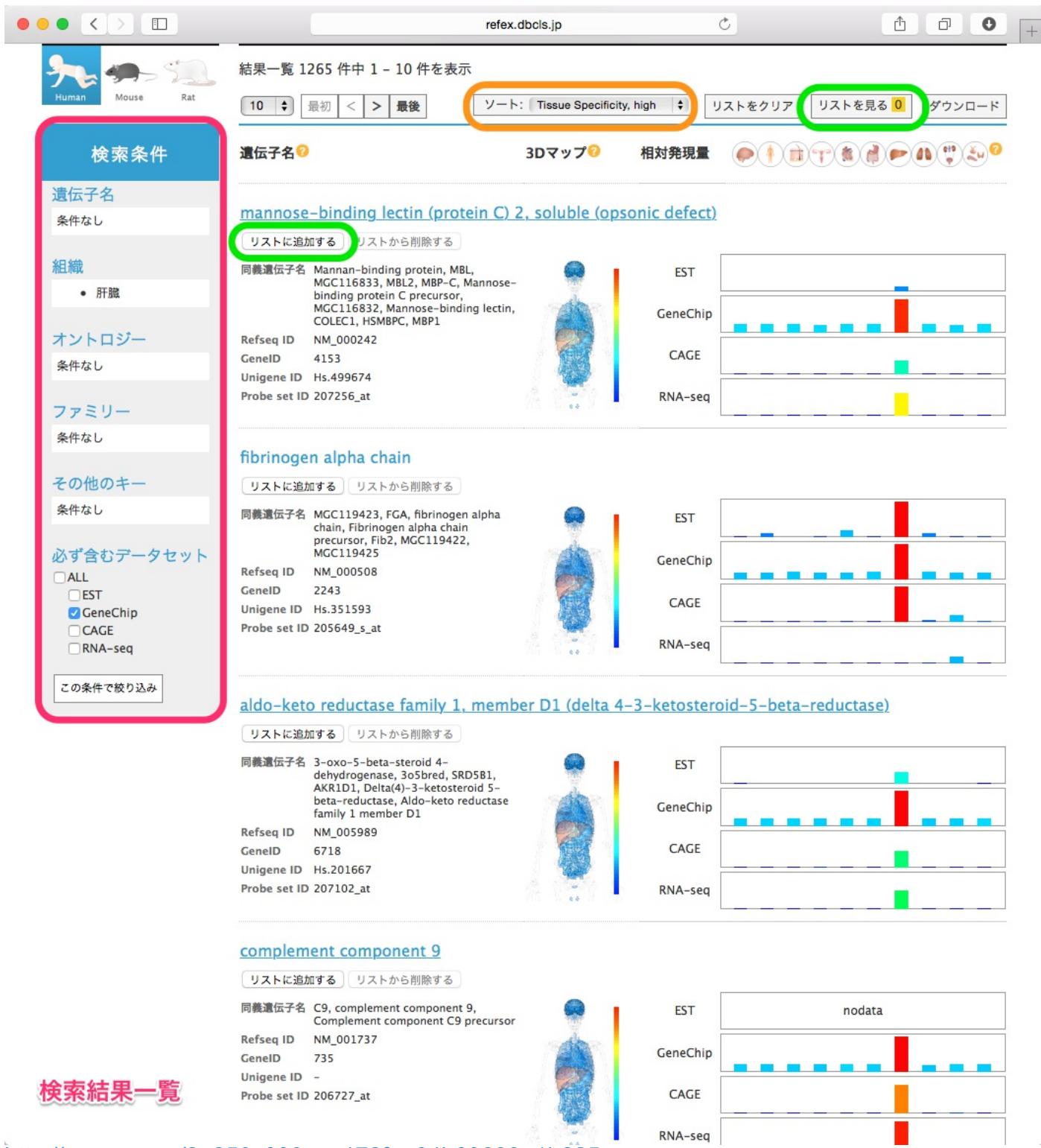
ダウンロード



RefEx はCreative Commons 表示 2.1 日本 Licenseの下でライセンスされています。
原著作者はライフサイエンス統合データベースセンター です。

<http://gyazo.com/fab7f0ba81afbce32061692c344bf03f>

- 検索結果一覧が表示されます。検索結果一覧では、「ソート項目の切り替え」や「絞り込み検索」、「リストへの追加」ができます。(手順11以降で解説します。)
- 各遺伝子の青字の部分(例 fibrinogen alpha chain http://refex.dbcls.jp/gene_info.php?lang=ja&db=human&geneID=2243&refseq=NM_000508&unigene=Hs.351593&probe=205649_s_at)をクリックすると詳細情報を閲覧できます。



5. 「ヒートマップ on Bodyparts3D」では、表示する部位の切り替え（全身・体幹部・頭部）ができます。「皮膚・骨格筋を表示」もしくは「アニメーション表示」にチェックを入れるとどのように表示されるでしょうか。
6. 「組織40分類別データ」では、バーの上にマウスオーバーすると測定部位と発現値が表示されます。

fibrinogen alpha chain 詳細情報を見る

同義遺伝子名 MGC119423, FGA, fibrinogen alpha chain, Fibrinogen alpha chain precursor, Fib2, MGC119422, MGC119425

発現データ

正常組織・臓器 40分類 正常組織(発生期含む)・初代培養細胞 (FANTOM5 CAGE)

発現マップ on BodyParts3D ? 組織40分類別データ ? Download

相対発現量を、人体3D画像にマップしたものです。Genechip組織40分類の発現パターンを使用しています。

EST GeneChip CAGE RNA-seq

バーをマウスオーバー

詳細情報

IDs ? 遺伝子オントロジー (GO ID) ?

Refseq ID NM_000508
Gene ID 2243
Unigene ID Hs.351593
Probe set ID 205649_s_at [HG-U133_Plus_2]
Ensembl ID ENSG00000171560

オーソログ対応遺伝子 ?

マウス[1] NM_001111048, NM_010196
ラット[3] NM_001008724, NM_052797

染色体 ?

4q31.3 [155.504.278 - 155.511.918]

詳細情報ページ

<http://gyazo.com/b60518629c6dd0fe8163776cc7824a3c>

- 「Download」をクリックすると、表示中の遺伝子の組織40分類別の発現データがタブ区切り形式でダウンロードできます。
 - 「Probe set ID」のリンク先をクリックすると、どういう情報が参照できるでしょうか。

refex.dbcls.jp

Reference
Expression
Dataset
English | 日本語

RefEx

検索

fibrinogen alpha chain 詳細情報を見る

同義遺伝子名 MGC119423, FGA, fibrinogen alpha chain, Fibrinogen alpha chain precursor, Fib2, MGC119422, MGC119425

発現データ

正常組織・臓器 40分類 正常組織(発生期含む)・初代培養細胞(FANTOM5 CAGE)

liver 1/2 ↑ ↓ go

FANTOM5 CAGEのサンプル情報一覧は[こちら](#)

liver, adult
11.22:adult tissue

サンプル名
発現値:サンプル分類

ドラッグしながら
移動可能

<http://gyazo.com/78a17e8253cb9ed64f6becf96b5a1e03>

9. 遺伝子オントロジー(GO ID)をクリックすると、そのGO termを持つ他の遺伝子を一括で検索できます。
 - 例として、GO:0007596 <http://refex.dbcls.jp/genelist.php?lang=ja&db=human&idkind=1&ids=GO:0007596> blood coagulation をクリックしてみましょう。
10. 右側のFANTOM5 CAGEのタブをクリックすると、FANTOM5 CAGEデータのビューアに切り替わります。
 - ビューアは上部が拡大図で、下部が全体表示になっています。
 - 検索窓にキーワードを入れるとサンプル名を検索できます。ヒットしたサンプルはオレンジ色で強調されます。
 - 右側に、サンプル名と発現値、サンプル分類が表示されます。
 - RefEx用に整理したサンプル情報一覧 <http://bit.ly/fantom5cagehuman>も閲覧可能です。
11. 検索結果一覧に戻ります。ソート項目を切り替えて、どのように結果が変わるでしょうか。
12. 様々な条件で検索結果を絞り込むことができます。絞り込み検索は左のバーから行えます。

- 遺伝子名に「liver」を含むデータは何件あるでしょうか。
- 「遺伝子名」の下の「条件なし」をクリックして表示されるウインドウに「liver」と入力し、「Include」をクリックし、「この条件で絞り込み」を押します。
- 「遺伝子名」の項目で「Exclude」に「solute」を加えると、検索結果はどう変わるでしょうか。
- 「組織」の項目で、データ元をRNA-seqに変更したり、臓器の指定を追加すると検索結果はどう変わるでしょうか。
- 「必ず含むデータセット」の「ALL」にチェックを入れると、検索結果はどう変わるでしょうか。

13. 個々の遺伝子の詳細情報は、リストに追加することで並列に比較することができます。

- 肝臓特異的遺伝子の検索結果一覧 [http://refex.dbcls.jp/genelist.php?
lang=ja&db=human&roku_valid=1&rk%5B31%5D=31&order_key=score](http://refex.dbcls.jp/genelist.php?lang=ja&db=human&roku_valid=1&rk%5B31%5D=31&order_key=score)に移動して、3つの遺伝子を「リストに追加」してみましょう。
- 追加した件数は「リストを見る」の横に表示されます。
- 「リストを見る」をクリックするとリストに移動します。
- 『並べて表示する』にチェックを入れて、「遺伝子を並べて表示」をクリックします。
- 並列に比較することで見えてくる「違い」はなんでしょうか。

RefEx - 一覧

RefEx - Gene List

Close

遺伝子を並べて表示

Gene Name	fibrinogen alpha chain	aldo-keto reductase family 1, member D1 (delta 4-3-ketosteroid-5-beta-reductase)	mannose-binding lectin (protein C) 2, soluble (opsonic defect)
同義遺伝子名	MGC119423, FGA, fibrinogen alpha chain, Fibrinogen alpha chain precursor, Fib2, MGC119422, MGC119425	3-oxo-5-beta-steroid 4-dehydrogenase, 3o5bred, SRD5B1, AKR1D1, Delta(4)-3-ketosteroid 5-beta-reductase, Aldo-keto reductase family 1 member D1	Mannan-binding protein, MBL, MGC116833, MBL2, MBP-C, Mannose-binding protein C precursor, MGC116832, Mannose-binding lectin, COLEC1, HSMBPC, MBP1

発現マップ on BodyParts3D

EST

GeneChip

CAGE

RNA-seq

相対発現量

EST

GeneChip

CAGE

RNA-seq

同義遺伝子名: Mannan-binding protein, MBL, MGC116833, MBL2, MBP-C, Mannose-binding protein C precursor, MGC116832, Mannose-binding lectin, COLEC1, HSMBPC, MBP1

RefSeq ID: NM_000242

EST

GeneChip

<http://gyazo.com/f832aab525efcbd99854b8c920be0fcf>

<http://gyazo.com/0c604ddeee80bf4adf14ce52876a5744>

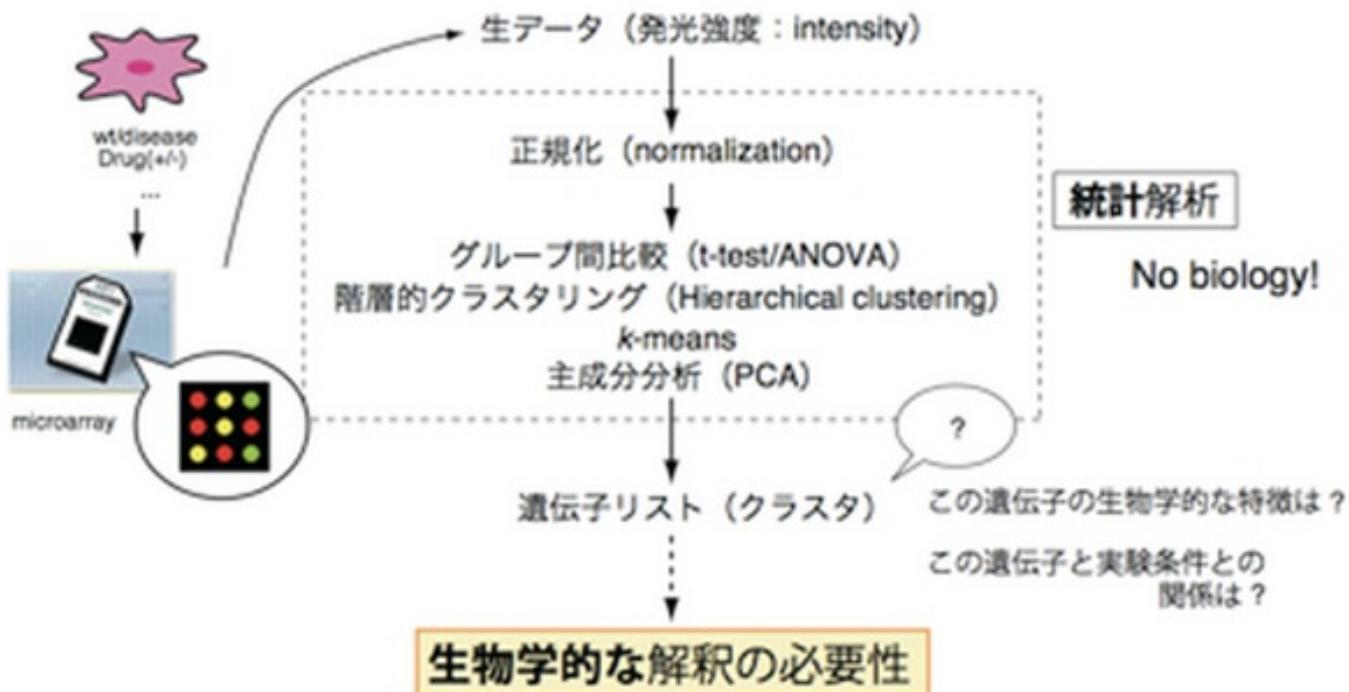
14. 自分の研究テーマに関連する、また興味のある遺伝子について検索してみましょう。

数十～数千の遺伝子群の生物学的解釈

DAVID: The Database for Annotation, Visualization and Integrated Discovery

<http://david.abcc.ncifcrf.gov/>

- マイクロアレイやRNA-seqデータの生物学的な解釈
- マイクロアレイやRNA-seq、すなわち遺伝子発現解析の一般的な目的は、実験条件によって得られた数十～数千の遺伝子群の発現が生物学的にどういう意味を持つかを考えることです。



<http://gyazo.com/52cb4c40b1313a52f8ded6923bdd8ef0>

- 今回は、その方法の一つとして、遺伝子リストにある遺伝子群に機能アノテーション(Gene Ontologyなど)を付与し、生物学的な解釈を行います。

遺伝子リストの準備

- サンプルデータとして、公共遺伝子発現データベースからメタ解析してえた遺伝子リストを用います。この遺伝子リスト **affy10.txt** は、ある刺激前後の2群間で発現増加した実験が10回以上あった遺伝子群のリストです。
→ **affy20.txt** <https://raw.githubusercontent.com/AJACS-training/AJACS60/master/bono1/affy10.txt> (右クリックして「新しいタブで開く」もしくは「名前を付けてリンク先を保存」してください。)
- このデータは、どのような実験から得られたデータなのか、どのように解釈できるのかを DAVIDを使って考察してみましょう！

【実習4】 DAVIDを用いて、発現データの結果を生物学的に解釈する

- 【復習用 統合TV】 DAVIDを使ってマイクロアレイデータを解析する 2012
<http://doi.org/10.7875/togtv.2012.079>
 - 【復習用 統合TV】 DAVIDの使い方 実践編 <http://doi.org/10.7875/togtv.2013.033>
- DAVID6.8beta <https://david-d.ncifcrf.gov/home.jsp>の上部メニューの「Start Analysis」をクリックします。

DAVID Bioinformatics Resources 6.8(Beta)
National Institute of Allergy and Infectious Diseases (NIAID), NIH

Home Start Analysis Shortcut to DAVID Tools Technical Center Downloads & APIs Term of Service Why DAVID? About Us

*** Welcome to DAVID 6.8 Beta with updated Knowledgebase ([more info](#)). ***

Recommending: A paper published in *Nature Protocols* describes step-by-step procedure to use DAVID!

Welcome to DAVID 6.8 Beta

2003 - 2016

What's Important in DAVID?

- [Version 6.7 release note](#)
- [New requirement to cite DAVID](#)
- [IDs of Affy Exon and Gene arrays supported](#)
- [Novel Classification Algorithms](#)
- [Pre-built Affymetrix and Illumina backgrounds](#)
- [User's customized gene background](#)
- [Enhanced calculating speed](#)

Statistics of DAVID

DAVID Bioinformatic Resources Citations
4061

Identify enriched biological themes, particularly GO terms

<https://gyazo.com/35e1864ecb0c4567836b588004707628>

- 画面左側バーで、probe IDリストをコピペ or ファイルを指定します。
- リストのIDの種類タイプを選択します。... 今回は、「AFFYMETRIX_3PRIME_IVT_ID」と「Gene List」
- Submit List をクリックするとリストが読み込まれます。

Upload Gene List[Demolist 1](#) [Demolist 2](#)[Upload Help](#)**Step 1: Enter Gene List**

A: Paste a list

Or

B: Choose From a File

 affy10.txt
 Multi-List File ?**Step 2: Select Identifier**AFFYMETRIX_3PRIME_IVT_ID**Step 3: List Type** Gene List
 Background**Step 4: Submit List**<https://gyazo.com/d6f8a6a7ea659373fd80a1a31d82b2fc>

← Step 1. Submit your gene list through left panel.

An example:

Copy/paste IDs to "**box A**" → Select Identifier as "**Affy_ID**" → List "Submit" button

1007_s_at
1053_at
117_at
121_at
1255_g_at
1294_at
1316_at
1320_at
1405_i_at
1431_at
1438_at
1487_at
1494_f_at
1598_g_at

5. アップロードしたリストは、左側バーの「List Manager」で「affy10」として保存されています。削除やrenameもできます。

DAVID Bioinformatics Database

Analysis Wizard
DAVID Bioinformatics Resources 6.8(Beta), NIAID/NIH

Home Start Analysis Shortcut to DAVID Tools Technical Center Downloads & APIs Term of Service Why DAVID? About Us

*** Welcome to DAVID 6.8 Beta with updated Knowledgebase ([more info](#)). ***

Upload List Background

Gene List Manager

Select to limit annotations by one or more species

[Help](#)

- Use All Species -
Homo sapiens(1042)
Unknown(100)

[Select Species](#)

List Manager [Help](#)

affy10

Select List to:

[Use](#) [Rename](#)
[Remove](#) [Combine](#)
[Show Gene List](#)

[View Unmapped Ids](#)

<https://gyazo.com/624b0b90b35dd24d84433bd691c4e318>

Analysis Wizard

[Tell us how you like the tool](#)
[Contact us for questions](#)

✓ Step 1. Successfully submitted gene list

Current Gene List: affy10

Current Background: Homo sapiens

Step 2. Analyze above gene list with one of DAVID tools



[Which DAVID tools to use?](#)

⊕ Functional Annotation Tool

- [Functional Annotation Clustering](#)
- [Functional Annotation Chart](#)
- [Functional Annotation Table](#)

Gene Functional Classification Tool

6. 解析を続けます。真ん中の「Functional Annotation Tool」をクリックします。

7. 「Gene Ontology」をクリックすると、Gene Ontologyを用いた解析の細かいメニューが表示されます。

Upload List Background

Gene List Manager

Select to limit annotations by one or more species

[Help](#)

- Use All Species -
Homo sapiens(1042)
Unknown(100)

[Select Species](#)

List Manager [Help](#)

affy10

Select List to:

[Use](#) [Rename](#)
[Remove](#) [Combine](#)
[Show Gene List](#)

[View Unmapped Ids](#)

Annotation Summary Results

[Help and Tool Manual](#)

Current Gene List: affy10

Current Background: Homo sapiens

717 DAVID IDs

[Check Defaults](#)

[Clear All](#)

Disease (1 selected)

Functional_Categories (3 selected)

Gene_Ontology (3 selected)



<https://gyazo.com/33cba945d73b6ab63fb5926e31e833f6>

8. 今回は、GOTERM_BP_FAT (BP = Biological Process)に注目します。その右の「Chart」

をクリックすると結果がポップアップされます。

DAVID Bioinformatics Resources 6.8(Beta)
National Institute of Allergy and Infectious Diseases (NIAID), NIH

*** Welcome to DAVID 6.8 Beta with updated Knowledgebase ([more info](#)). ***

Functional Annotation Chart

[Help and Manual](#)

Current Gene List: affy10
Current Background: Homo sapiens
717 DAVID IDs

■ Options

Rerun Using Options Create Sublist

197 chart records

[Download File](#)

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	GOTERM_BP_DIRECT	canonical glycolysis	RT	15	2.1	8.1E-14	2.2E-10	
	GOTERM_BP_DIRECT	glucose metabolic process	RT	25	3.5	3.9E-11	5.3E-8	
	GOTERM_BP_DIRECT	response to hypoxia	RT	28	3.9	3.1E-10	2.8E-7	
	GOTERM_BP_DIRECT	cellular response to hypoxia	RT	22	3.1	1.8E-9	1.2E-6	
	GOTERM_BP_DIRECT	negative regulation of transcription from RNA polymerase II promoter	RT	54	7.5	9.0E-7	4.9E-4	
	GOTERM_BP_DIRECT	carbohydrate metabolic process	RT	33	4.6	1.4E-6	6.3E-4	
	GOTERM_BP_DIRECT	small molecule metabolic process	RT	96	13.4	1.6E-5	6.1E-3	
	GOTERM_BP_DIRECT	gluconeogenesis	RT	10	1.4	4.2E-5	1.4E-2	
	GOTERM_BP_DIRECT	negative regulation of cell proliferation	RT	32	4.5	6.1E-5	1.8E-2	
	GOTERM_BP_DIRECT	positive regulation of transcription from RNA polymerase II promoter	RT	60	8.4	6.6E-5	1.8E-2	
	GOTERM_BP_DIRECT	cholesterol biosynthetic process	RT	9	1.3	6.7E-5	1.6E-2	
	GOTERM_BP_DIRECT	extracellular matrix organization	RT	27	3.8	7.2E-5	1.6E-2	
	GOTERM_BP_DIRECT	positive regulation of transcription, DNA-templated	RT	36	5.0	1.6E-4	3.4E-2	
	GOTERM_BP_DIRECT	glycolytic process	RT	7	1.0	1.8E-4	3.3E-2	

<https://gyazo.com/3bf22e48835906313c5d5dd777c55d91>

9. タイトル行をクリックするとソートできます。
10. さらに、GOTERM_CC_FAT や GOTERM_MF_FAT を見て、上位にリストされた GOTermにどのような共通点・相違点があるでしょうか。
11. Pathways > KEGG_PATHWAY や Tissue Expression > UP_TISSUE なども見てみましょう。
12. さらにMetascapeという別のエンリッチメント解析ツール <http://metascape.org/>を使って同じデータセットで同様にデータ解析してみましょう。

Step 1

Cancel

Select Columns in your Excel file.

at(Type:Affymetrix ID)(Total:1141)

First row used as column header.

Upload File Format

Single List: .xls/.xlsx .csv
Download .txt
Download

Multiple List: .xls/.xlsx .csv
Download .txt
Download

Test Upload

single list
3 gene lists

Step 2

Express Analysis **Custom Analysis**

Please click the button below to see reports

Analysis Report Page

100

News & Updates

- 2016-05-19 Software updated [Release History](#)
- 2016-05-16 Data sources updated ([detailed update report](#)). We commit to keep it fresh!
- 2016-1-4 [Presentation](#) at Host-Pathogen Interactions: Circuits to Systems, SDCSB. Launch of the meta-analysis feature.
- 2015-12-9 [First publication](#): Meta- and Orthogonal Integration of Influenza 'OMICs' Data Defines a Role for UBR4 in Virus Budding. *Cell Host & Microbe*. 2015 18:723-735.
- 2015-10-8 [Keynote presentation](#) at The SDCSB's Quarterly Systems-to-Synthesis Meeting, SDCSB. Launch of metascape.org.

Questions and comments can be sent to: metascape.team@gmail.com



<https://gyazo.com/5151d788f211b25072d3192ef6a3e469>

13. 複数のツールで得られた結果を踏まえ、「ある実験」とはどのような実験であったか考察してみましょう。
- **affy10.txt** は「ある刺激前後の2群間で発現が1.5倍以上上昇した実験が10回以上あった遺伝子群のリスト」
14. 早く終わった方は、さらに20回以上発現上昇があった遺伝子群のリスト、**affy20.txt** <https://raw.githubusercontent.com/AJACS-training/AJACS60/master/bono1/affy20.txt> (右クリックして「新しいタブで開く」もしくは「名前を付けてリンク先を保存」してください) で同じデータ解析をやってみましょう。10回以上のリストと比べてどういった違いがみられるでしょうか?
 - このようにいろいろな閾値を試して、結果を見て考察し、最終的な閾値としておられます(例: Table 4. Gene Set Enrichment Analysis of up/down-regulated genes after UV irradiation. DOI: 10.1371/journal.pone.0116007.t004 <http://dx.doi.org/10.1371/journal.pone.0116007.t004>)