

150827 遺伝子発現データベース・解析ツールの使い方@広島大学

情報・システム研究機構

ライフサイエンス統合データベースセンター

小野 浩雅 hono@dbcls.rois.ac.jp

2015年8月27日(木) 広島大学

これは2015年8月27日の、遺伝子発現データベース・解析ツールの使い方@広島大学 の講習資料です。

この内容の続編として、AJACS御茶ノ水(2015年5月)における応用・実践編 がありますので、こちらもあわせてご活用ください。

概要

本講習は、だれでも自由に使うことができる公共データベース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データ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。

また、自ら行なった大規模発現解析の(あるいは公共データベースから取得・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データベースやツールを知る
 - 統合TV
 - 個々の遺伝子の発現プロファイルを調べる
 - RefEx
 - 【実習1】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2】DAVIDを用いて、発現データの結果を生物学的に解釈する
 - 【実習3】DAVIDを用いて、発現データの結果を生物学的に解釈する その2
-

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

- 本講習内容をスムーズに理解するために押さえておくとよい基礎知識
 - [「遺伝子のDB・ウェブツールの基礎」\(2012年8月AJACS筑波2\)](#)
 - 非モデル生物のデータをモデル生物のデータに見立てるためのID対応表づくりについて
 - [「コマンドラインで遺伝子配列を解析する」\(2012年7月\)](#)
 - 次世代シーケンス解析について
 - [「次世代シーケンサー\(NGS\)と関連するデータベース・ツール」\(2014年7月AJACS信州\)](#)
 - [「galaxyによるNGS解析」\(2015年1月AJACSA三島\)](#)
 - 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](#)
-

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

- 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/ja/>



生命科学系DB・ツール使い倒し系チャンネル

はじめての方へ 番組ランキング ほかの便利な方法 よくある質問 スタッフ 番組リクエスト・お問い合わせ



目的別に検索！

▶ 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

▶ 文献検索・辞書
情報収集
PC環境構築

▶ DBCLSサービス
講演・講習動画

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！

検索窓にキーワードを入れると、入力のたびごとに即座に候補の番組が絞り込まれます。

先頭のタイトル行をクリックすると、昇順・降順で並び替えができます。

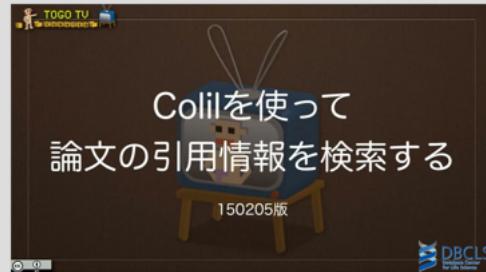
お探しの動画が見つからない or 統合TV未掲載の場合は、[統合TV番組リクエストフォーム](#)へどうぞ!!

10 エントリを表示

検索: 0

番組タイトル（YOUTUBEへのリンク）

番組の概要（画像をクリックすると番組の再生ページへ移動します。）



Colilは、生命科学分野の文献間の引用関係を検索・閲覧できるサービスです。ある論文について、他の論文が本文中でどのように引用し述記しているかを、効率的に調べることができます。

検索例やデータの詳細については、Colil Data Portalにまとめられています。



- YouTube版もあります <http://www.youtube.com/user/togotv/videos>



togotv

チャンネル登録

ホーム 動画 再生リスト チャンネル フリートーク 概要 検索

統合TV プロモーションムービー

simplify sharing
最先端の研究講演・講義も

これら3つのインフラの目的は、科学者間の共有を簡単にしよう、それに伴ってマートンが書いていた科学者の行動規範を

0:43 / 0:59

再生

音量

再生時間

設定

画面

統合TV プロモーションムービー

視聴回数 1,913 回 1年前

本家はこちら <http://togotv.dbcls.jp/2013...>
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。数多くの解説動画の情報をうまく整理し、ユーザーが真する情報に素早く容易にたどりつけるようにまとめています。生命科学研究に使われるサイトを知り、その使い倒し術を知る手段として、日々の実験や学習のためぜひご活用ください。

人気のアップロード

パワーポイントの图形描画機能でイラストをつくる方法

3年前・視聴回数 23,286 回
本家はこちら <http://togotv.dbcls.jp/20110509.html#p01>
スライドやポスター等のプレゼンテーション用資料に「イメージ通りのイラストがみつか...

写真を削除して、

12:14

- 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、[統合TV の発現制御解析 カテゴリー](#)にあります。
 - 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](#)へどうぞ!!
- 統合TVを作つてみたい方、募集中です。

個々の遺伝子の発現プロファイルを調べる

RefEx (Reference Expression dataset)

- 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](#)
- 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 IDH3α expression promotes malignant tumor growth by inducing HIF-1-mediated metabolic reprogramming and angiogenesis. Oncogene. \(22 December 2014\) | doi:10.1038/onc.2014.411](#)
- 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
 - 細胞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1】RefExを使って、組織特異的遺伝子を検索する

- [【復習用】RefExの使い方](#)

1. <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。

The screenshot shows the RefEx homepage. At the top right, there's a logo with three blue horizontal bars and the text "RefEx". To its right, a sidebar says "臓器ごとの発現比較を" (Comparison of gene expression across organs), "4つの実験手法と" (4 experimental methods and), "ボディペーツ3Dで。" (3D Body Parts), and "▼もっと詳しく" (More details). Below the sidebar are links for "Reference", "Expression", "Dataset", and language options "English | 日本語".

In the center, there's a search bar with the placeholder "キーワードで検索" (Search by keyword) and a yellow "検索" (Search) button. To the right of the search bar is an example "ex) troponin, ALB".

On the left, there are icons for "Human" (blue silhouette), "Mouse" (black silhouette), and "Rat" (white silhouette). Below these are sections for "遺伝子オントロジー Gene Ontology" and "遺伝子ファミリー InterPro".

The main feature highlighted is the "組織特異的に発現する遺伝子を見る" (View genes expressed specifically in tissues) section. It shows a grid of icons representing various organs: brain, human figure, skin, uterus, heart, liver, stomach, intestines, lungs, kidneys, and bladder. A red arrow points to the "肝臓" (Liver) icon, which is highlighted with a blue box. To the right of the liver icon, the text "臓器のアイコンをマウスオーバー" (Hover over the organ icon) is written in pink.

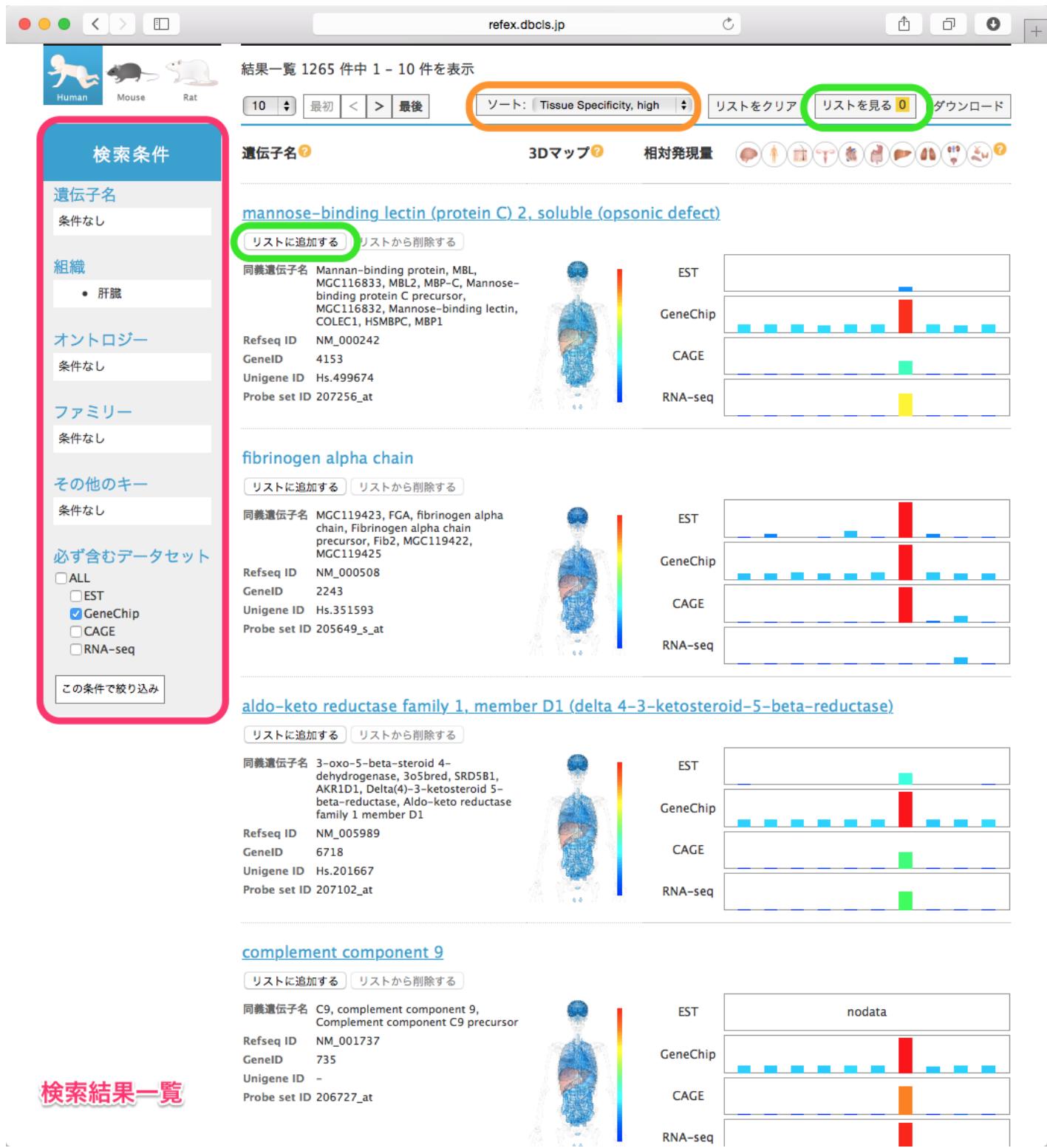
Below the liver icon, there are two columns of biological processes:

- cellular process
- biological regulation
- metabolic process
- multicellular organismal process
- response to stimulus
- developmental process

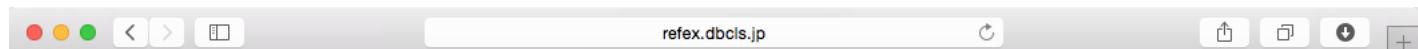
On the far right, there are buttons for "他のオントロジーを選ぶ" (Select other ontologies) and "他のファミリーを選ぶ" (Select other families). At the bottom, there's a "染色体" (Chromosome) section with a "染色体領域を選ぶ" (Select chromosome region) button.

At the very bottom, there are links for "Advanced Search" and "Advanced Search" (repeated), and a "ページ上部に戻る" (Return to top of page) button.

3. 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。)
4. 各遺伝子の青字の部分（例 [fibrinogen alpha chain](#)）をクリックすると詳細情報を閲覧できます。



- 「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。
- 「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。

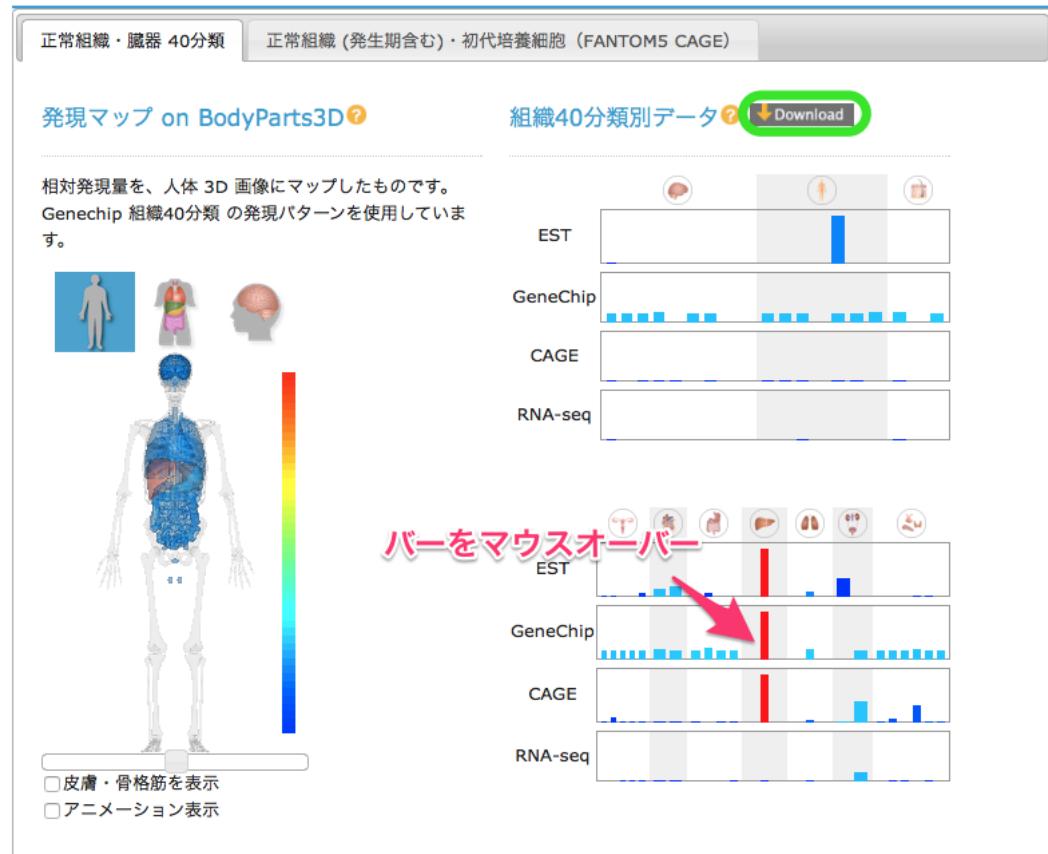


fibrinogen alpha chain

[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



詳細報

IDs?

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_052297

詳細情報ページ

染色体?

4g31.3 [155.504.278 - 155.511.918]

遺伝子オントロジー (GO ID) ?

Biological Process

- GO:0007596 blood coagulation
- GO:0030168 platelet activation
- GO:0002576 platelet degranulation
- GO:0051258 protein polymerization
- GO:0051592 response to calcium ion
- GO:0007165 signal transduction

Cellular Component

- GO:0005938 cell cortex
- GO:0009986 cell surface
- GO:0009897 external side of plasma membrane
- GO:0005576 extracellular region
- GO:0005615 extracellular space
- GO:0005577 fibrinogen complex

7. 「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。
 8. 「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 検索

RefEx

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

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

発見データ

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

9. 遺伝子オントロジー(Gene Ontology: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](#) 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](#)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](#)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.dbcls.jp

RefEx - Gene List

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

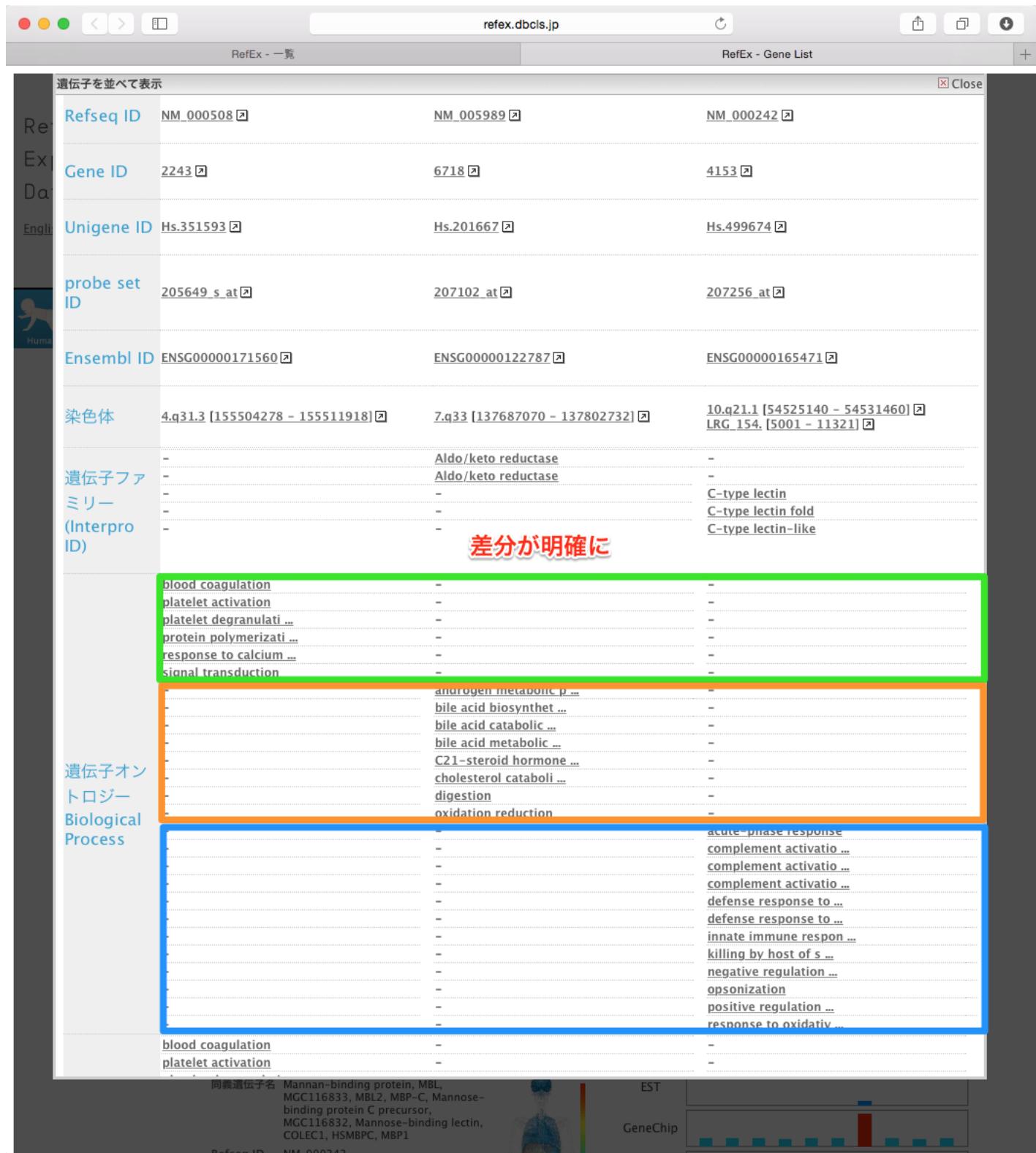
Mannan-binding protein, MBL, MGC116833, MBL2, MBP-C, Mannose-binding protein C precursor, MGC116832, Mannose-binding lectin, COLEC1, HSMBPC, MBP1

RefSeq ID

NM_000242

EST

GeneChip



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。

BioGPS

- ヒト、マウス、ラットのさまざまな組織や細胞(株)における遺伝子発現プロファイルのデータベース
- [BioGPS](#)はAffymetrix社製のマイクロアレイであるGeneChipを用いたさまざまな組織や細胞(株)遺伝子発現プロファイルのデータベース。
- 検索した遺伝子に対して、種々の外部データベースを横断検索することができるだけでなく、それらの設定を保存したり、表示方法を自由にカスタマイズすることができる「Gene annotation portal」。
- 外部データベースには、Wikipedia(Gene Wiki)、著名な試薬会社の検索窓へのリンク集、pathway、Nature系DB、モデル生物DB、文献DBなど多種多様
- マウスのエキソンアレイのデータから遺伝子のスプライシングバリエント(Splicing variant)の発現状況も調べることが可能。最近ではCircadian関係のデータも。
- さらに最近のアップデートで、NCBI Gene Expression Omnibus (GEO)中から選抜されたデータセットに切り替えて発現状況を調べること

が可能に。

【実習(skip)】 BioGPSを使ってある遺伝子の発現プロファイルを調べる

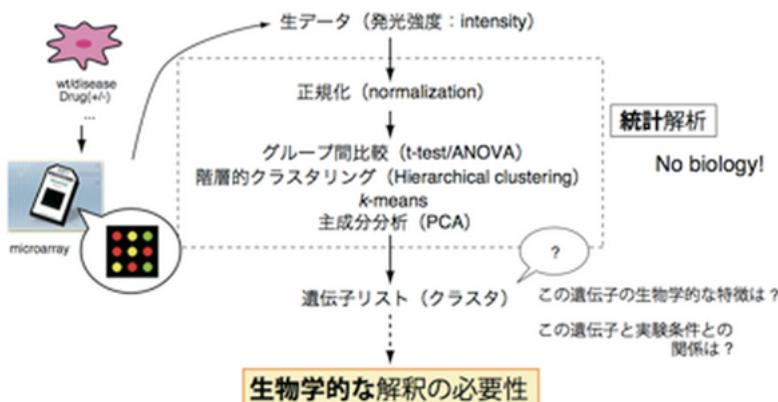
- 【復習用】遺伝子発現プロファイルデータベースBioGPSを使い倒す2012
- 【以前の講習会動画】遺伝子発現データベースの活用法

1. <http://biogps.org/>を開きます。2.骨格筋の分化決定遺伝子であるMyogenic differentiation 1 (MyoD)の発現プロファイルを調べてみましょう。中央の検索窓に「myod」と入力し、「search」を押します。
 2. 表示された検索結果の中から「ID 4654」をクリックします。
 3. 最初はヒトのマイクロアレイデータが表示されます。
 4. 画面左側の"Current Gene List"は右上の<<アイコンをクリックすると非表示にできます。非表示にすることで画面を広く使うことができます。
 5. ページ内のウインドウは通常のウインドウと同じようにドラッグによる移動やサイズの変更などを行うことができます。歯車マークのメニューから"Open in browser"を選択すると、新しいタブで表示できます。
 6. "Search"と書かれた窓に単語(組織名など)を入力すると、その単語の含まれた部分が赤くハイライト表示されます。今回は "Muscle"と入力してみます。
 7. "Zoom"のバーを用いることで、グラフの表示範囲を調整することが出来ます。
 8. 発現量を示すバーをクリックすると発現強度の値が表示されます。
 9. マイクロアレイデータ右上の"Species: Hs"をクリックするとマウスやラットを選択できるので、"M. musculus (Mouse)"をクリックしてマウスのデータを表示できます。
 10. MyoDはどの組織、細胞で強く発現しているでしょうか？
 11. 場合によっては"Probeset"のプルダウンメニューから複数の項目を選択できる場合があります。これはどのようなケースが考えられるでしょうか？
 12. "Static Image"をクリックすると、ズームや検索機能などのついていない、画像だけのグラフで表示されます。低スペックなマシンでは、こちらの方が軽快に動作するでしょう。
 13. "Correlation"タブをクリックして検索すると、発現パターンが似ている他の遺伝子を検索できますが、どのような遺伝子が出てくるでしょうか？
 14. "Downloads"をクリックすると現在表示している遺伝子の発現データを CSV 形式でダウンロードできます。
 15. "Dataset"の右にある'change"をクリックすると、デフォルトで用意されているデータセットやNCBI GEO中のデータセットを検索でき、それらのデータに表示を切り替えることができます。"Species: Hs"に切り替えてから、"change"をクリックしたあと、"Default Datasets"から"Barcode on normal tissues (262 samples)"を選択します。どのようにデータが変わったでしょうか？
 16. さらに"Search"からキーワード検索で、GEOのデータを検索してみましょう。"C2C12"と検索するとどのようなデータが選択できるでしょうか？
 17. 右上の「default rayout」をクリックすると、検索した遺伝子に関して種々の外部データベースを横断検索できますが、どのようなデータが閲覧できるのか調べてみましょう。
 18. 左上の「Search」タグをクリックして検索画面にもどり、自分の興味ある遺伝子について同様に検索してみましょう。すぐに自分の興味ある遺伝子が浮かばない場合は、著名な*IPS細胞*を作るために必要な4因子 (Oct3/4・Sox2・Klf4・c-Myc) がどの組織で発現しているか、またデータを切り替えて検索してみましょう。
- 【余談】[BioGPSのiPhoneアプリ](#)が無料で公開されていますので、「あの遺伝子はどの組織で発現しているのかな？」とふと調べたいときにお手持ちのiPhoneで遺伝子発現を調べられます。

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

- マイクロアレイデータの生物学的な解釈
- マイクロアレイ実験の一般的な目的は、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。



- ・今回は、その方法の一つとして、マイクロアレイの結果にGene Ontologyの用語を付与することで、生物学的な解釈を行います。

マイクロアレイデータの準備

- ・サンプルデータとして、NCBI GEOから取得した公共の遺伝子発現データを用います。このデータは、ある実験の前後の2群間で有意に発現減少した遺伝子群のリストです。
→ [マル秘遺伝子リスト](#)（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2】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】DAVIDを使ってマイクロアレイデータを解析する 2012](#)
- ・[【復習用】DAVIDの使い方 実践編](#)

1. 上部メニューの「Start Analysis」をクリックします。

DAVID Bioinformatics Resources 6.7
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

Shortcut to DAVID Tools

Functional Annotation
Gene-annotation enrichment analysis, functional annotation clustering , BioCarta & KEGG pathway mapping, gene-disease association, homologue match, ID translation, literature match and [more](#)

Gene Functional Classification
Provide a rapid means to reduce large lists of genes into functionally related groups of genes to help unravel the biological content captured by high throughput technologies. [More](#)

Recommending: A [paper](#) published in *Nature Protocols* describes step-by-step procedure to use DAVID!

Welcome to DAVID 6.7

2003 - 2014

Search

What's Important in DAVID?

- [Current \(v 6.7\) release note](#)
- [New requirement to cite DAVID](#)
- [IDs of Affy Exon and Gene arrays supported](#)
- [Novel Classification Algorithms](#)

2. 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。
3. リストのIDの種類タイプを選択します。… 今回は、「AFFY_ID」と「Gene List」
4. Submit List をクリックするとリストが読み込まれます。

[Upload](#) [List](#) [Background](#)

Analysis Wizard

[Upload Gene List](#)

[Demolist 1](#) [Demolist 2](#)

[Upload Help](#)

Step 1: Enter Gene List

A: Paste a list

Clear

Or

B: Choose From a File

ファイルを選択 選択されていません

Multi-List File [?](#)

Step 2: Select Identifier

AFFYMETRIX_3PRIME_IVT_ID

Step 3: List Type

Gene List

Background

Step 4: Submit List

[Submit List](#)

An example:

Copy/paste IDs to "box A" -> Select Identifier as "**Affy_ID**" -> List Type as "**Gene List**" -> Click "**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

5. アップロードしたリストは、左側バーの「List Manager」で「Uploaded List_1」として保存されています。削除やrenameもできます。

[Upload](#) [List](#) [Background](#)

Analysis Wizard

[Gene List Manager](#)

Select to limit annotations by one or more species

[Help](#)

- Use All Species -
Arabidopsis thaliana(2928)
Unknown(1)

[Select Species](#)

[List Manager](#) [Help](#)

List_1

Select List to:

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

[View Unmapped Ids](#)

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

[Contact us for questions](#)

Step 1. Successfully submitted gene list

Current Gene List: List_1

Current Background: Arabidopsis thaliana

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](#)

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。

Gene List Manager

Select to limit annotations by one or more species

[Help](#)

- Use All Species -
Arabidopsis thaliana(2928)
Unknown(1)

[Select Species](#)List Manager [Help](#)

List_1

Select List to:

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

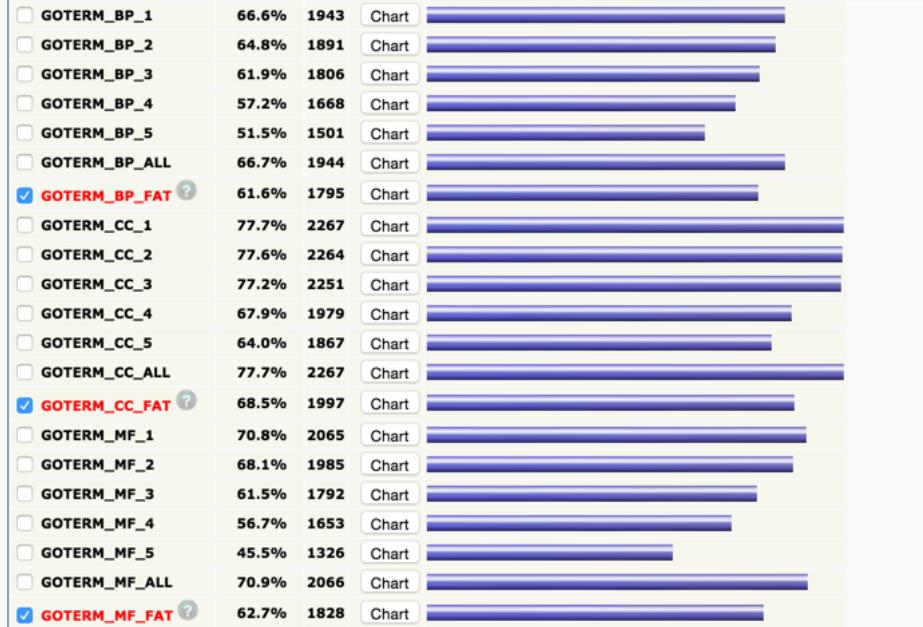
Annotation Summary Results

[Help and Tool Manual](#)

Current Gene List: List_1

Current Background: Arabidopsis thaliana

2916 DAVID IDs

[Check Defaults](#) [Clear All](#) Functional_Categories (3 selected) Gene_Ontology (3 selected) General Annotations (0 selected) Literature (0 selected) Main_Accessions (0 selected) Pathways (1 selected) Protein_Domains (3 selected) Protein_Interactions (0 selected) Tissue_Expression (0 selected)

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。

Functional Annotation Chart

[Help and Manual](#)

Current Gene List: List_1

Current Background: Arabidopsis thaliana

2916 DAVID IDs

 Options[Rerun Using Options](#) [Create Sublist](#)

301 chart records

[Download File](#)

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	photosynthesis	RT	102	3.5	1.5E-45	2.7E-42	
<input type="checkbox"/>	GOTERM_BP_FAT	photosynthesis, light reaction	RT	56	1.9	6.2E-29	5.7E-26	
<input type="checkbox"/>	GOTERM_BP_FAT	pigment metabolic process	RT	53	1.8	5.7E-21	3.5E-18	
<input type="checkbox"/>	GOTERM_BP_FAT	pigment biosynthetic process	RT	47	1.6	2.6E-19	1.2E-16	
<input type="checkbox"/>	GOTERM_BP_FAT	carboxylic acid biosynthetic process	RT	117	4.0	9.0E-18	3.3E-15	
<input type="checkbox"/>	GOTERM_BP_FAT	organic acid biosynthetic process	RT	117	4.0	9.0E-18	3.3E-15	
<input type="checkbox"/>	GOTERM_BP_FAT	generation of precursor metabolites and energy	RT	111	3.8	2.8E-15	8.4E-13	
<input type="checkbox"/>	GOTERM_BP_FAT	response to abiotic stimulus	RT	241	8.3	1.1E-13	2.8E-11	
<input type="checkbox"/>	GOTERM_BP_FAT	plastid organization	RT	41	1.4	2.9E-13	6.7E-11	
<input type="checkbox"/>	GOTERM_BP_FAT	nitrogen compound biosynthetic process	RT	124	4.3	5.1E-13	1.0E-10	
<input type="checkbox"/>	GOTERM_BP_FAT	heterocycle biosynthetic process	RT	50	1.7	9.6E-13	1.8E-10	

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。

- CC = Cellular Component

Functional Annotation Chart

[Help and Manual](#)**Current Gene List: List_1****Current Background: Arabidopsis thaliana****2916 DAVID IDs****Options**[Rerun Using Options](#) [Create Sublist](#)**68 chart records**[Download File](#)

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
<input type="checkbox"/>	GOTERM_CC_FAT	plastid part	RT		515	17.7	9.0E-190	2.9E-187
<input type="checkbox"/>	GOTERM_CC_FAT	chloroplast part	RT		504	17.3	1.1E-187	1.8E-185
<input type="checkbox"/>	GOTERM_CC_FAT	plastid	RT		987	33.8	1.4E-165	1.5E-163
<input type="checkbox"/>	GOTERM_CC_FAT	chloroplast	RT		973	33.4	1.3E-164	1.1E-162
<input type="checkbox"/>	GOTERM_CC_FAT	plastid thylakoid	RT		230	7.9	7.0E-110	4.6E-108
<input type="checkbox"/>	GOTERM_CC_FAT	chloroplast thylakoid	RT		230	7.9	7.0E-110	4.6E-108
<input type="checkbox"/>	GOTERM_CC_FAT	organelle subcompartment	RT		230	7.9	4.5E-109	2.4E-107
<input type="checkbox"/>	GOTERM_CC_FAT	thylakoid	RT		278	9.5	8.0E-109	3.7E-107
<input type="checkbox"/>	GOTERM_CC_FAT	thylakoid part	RT		228	7.8	6.0E-105	2.5E-103
<input type="checkbox"/>	GOTERM_CC_FAT	chloroplast thylakoid membrane	RT		200	6.9	4.4E-100	1.6E-98
<input type="checkbox"/>	GOTERM_CC_FAT	plastid thylakoid membrane	RT		200	6.9	4.4E-100	1.6E-98

- MF = Molecular Function

Functional Annotation Chart

[Help and Manual](#)**Current Gene List: List_1****Current Background: Arabidopsis thaliana****2916 DAVID IDs****Options**[Rerun Using Options](#) [Create Sublist](#)**151 chart records**[Download File](#)

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
<input type="checkbox"/>	GOTERM_MF_FAT	cofactor binding	RT		109	3.7	3.8E-8	4.4E-5
<input type="checkbox"/>	GOTERM_MF_FAT	chlorophyll binding	RT		14	0.5	1.5E-5	8.4E-3
<input type="checkbox"/>	GOTERM_MF_FAT	vitamin B6 binding	RT		32	1.1	3.4E-5	1.3E-2
<input type="checkbox"/>	GOTERM_MF_FAT	pyridoxal phosphate binding	RT		32	1.1	3.4E-5	1.3E-2
<input type="checkbox"/>	GOTERM_MF_FAT	poly(U) RNA binding	RT		11	0.4	4.6E-5	1.3E-2
<input type="checkbox"/>	GOTERM_MF_FAT	poly-pyrimidine tract binding	RT		11	0.4	4.6E-5	1.3E-2
<input type="checkbox"/>	GOTERM_MF_FAT	ATP-dependent peptidase activity	RT		12	0.4	8.3E-5	1.9E-2
<input type="checkbox"/>	GOTERM_MF_FAT	rRNA binding	RT		23	0.8	8.5E-5	1.6E-2
<input type="checkbox"/>	GOTERM_MF_FAT	coenzyme binding	RT		72	2.5	9.7E-5	1.6E-2
<input type="checkbox"/>	GOTERM_MF_FAT	vitamin binding	RT		39	1.3	1.2E-4	1.7E-2

11. Pathways > KEGG_PATHWAY や Tissue Expression > UP_TISSUE なども見てみましょう。

12. DAVID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。

- マル秘遺伝子リストは「ある実験の前後の2群間で有意に発現減少した遺伝子群のリスト」
- 生物種はArabidopsis thaliana (シロイヌナズナ)

答え合わせ

【実習2】DAVIDを用いて、発現データの結果を生物学的に解釈する その2

- DAVID の使い方に慣れてきたところで、実戦的な生物学的解釈に挑戦してみましょう。
- 今回は「正解」はありません。情報分析力と想像力が問われます。
- 例題は、[GSE28619](#) をつかいます。
- 健常者 vs アルコール性肝炎患者 の2群比較です。
- 多重比較法 (Benjamini & Hochberg) を指定して、有意水準1%未満かつ2倍以上発現差のあった遺伝子群のリストをあらかじめ用意しました。
 - 「健常者>AH患者_遺伝子リスト」[GEO2R Ctrl.txt](#)
 - 「AH患者>健常者_遺伝子リスト」[GEO2R AH.txt](#)
 - (この遺伝子リストの作り方は、[AJACS御茶ノ水の回](#) で解説しています。)
- このデータを使った論文 ([Transcriptome analysis identifies TNF superfamily receptors as potential therapeutic targets in alcoholic hepatitis.](#))も出ていますが、答え合わせをするもよし、著者らが見逃している着眼点はないか探すもよし、です。

