TargetMine チュートリアルと実習

(独)産業技術総合研究所 関西センター2012年12月26日

Copyright © 2012 The Mizuguchi laboratory (mizuguchilab.org), NIBIO

● このチュートリアルは TargetMine v0.99 に基づいて作成したものです。

内容

- I. 簡単な検索とレポートページの見方について
- II. リストとリストの解析について
- III. テンプレートクエリ
- IV. クエリビルダー

まずは TargetMine ウェブサイトにアクセスしてください。

http://targetmine.nibio.go.jp

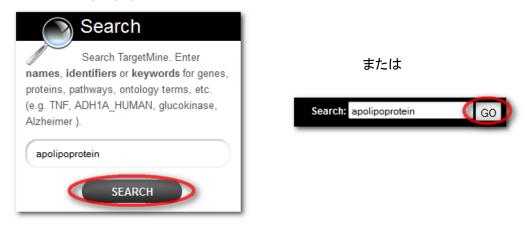
全ての機能は TargetMine のページ上部にあるツールバーから操作できます。



I. 簡単な検索とレポートページの見方について

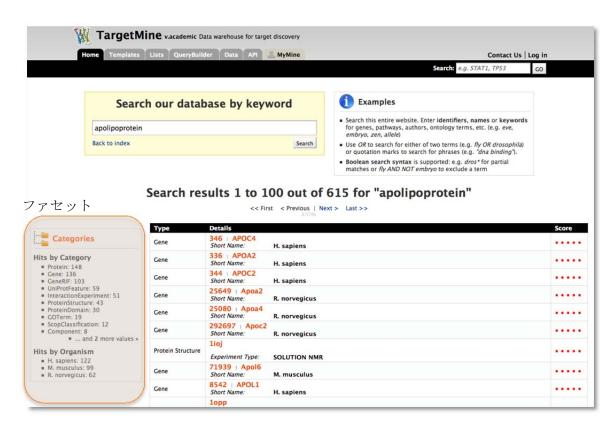
検索機能はトップページ左のサーチボックス(下図左),または,他のどのページでも右上の検索ボックスからアクセスできます。検索語を入力すれば,(例:遺伝子名,ID,キーワード),TargetMine が持つデータに対して全文検索ができます。

それでは、'apolipoprotein' で検索してみてください。



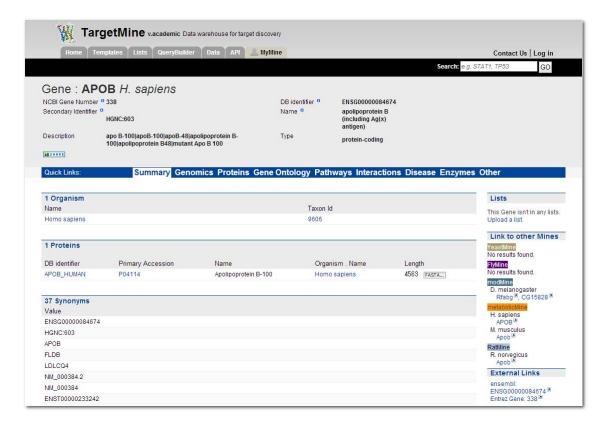
ヒント: ワイルドカード(アスタリスク) "*"を使うと部分一致検索ができます。

検索結果は、左側のファセット(Categoriesと書かれた分類)を用いて絞り込みができます。



● 各検索結果 (オレンジのハイライト) をクリックすると, その "オブジェクト" (特定 の遺伝子やタンパク質など) のレポートページが表示されます。

TargetMine にはいろいろなタイプのオブジェクトが格納されていますが、レポートの表示は全て同じ形式です(下図)。



● ページ上部にはオブジェクトの基本的な情報が表示され、下部にはオブジェクトのカテゴリ(遺伝子やタンパク質、化合物などの分類分け、TargetMine 内では"クラス"と呼ばれる)に応じて、より詳細な情報が表示されます。例えば、関連するクラスの情報や、クラス毎に予め用意された検索例("テンプレート"と呼ばれる)など(これらの詳細は後述)。

II. リストとリストの解析について

リスト機能を使うと自分で指定したリスト (例:遺伝子/タンパク質のリスト)を作成することができます。リストは検索結果の一覧からも作成できますし、ID を使ってアップロードすることもできます。リストは検索の絞り込みや、クエリビルダー(詳細は以下に示します)のテンプレートとして、また、論理演算を使った検索(Union、Intersect、Subtract)にも使用することができます。リストの検索やコピーも可能です。

リストページは上部メニューバーの "Lists" タブから開くことができます。"アップロード" タブはリストを作る専用のページで、ID の入力、またはファイルのアップロードからリストを作ることができます。

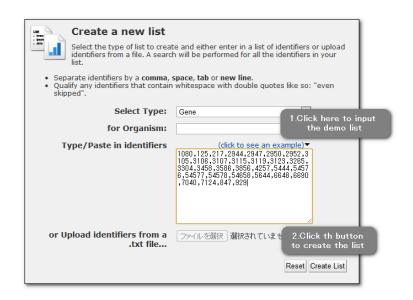


1つのリストには、ドロップダウンメニューから1つのタイプ(例:遺伝子、タンパク質)を指定してください(タイプの混在したリストは作れません)。オプションで生物種を指定することもできます。また、複数のタイプのIDを混在させることは構いません(例:遺伝子記号と遺伝子ID)。

このチュートリアルでは、例として遺伝子のリストを作成し、解析します。'List Analysis page'で得られた解析結果は、クエリビルダーを用いて新たなクエリを作るために保存します。

'click to see an example' をクリックしてください。クリックするとサンプルの ID 群が自動的に入力されます。

次は "Create List" ボタンをクリックしてください。



● リストの確認

"2 Verify identifier matches" では、入力された ID が TargetMine 中に存在するか、または曖昧さが残らないかなどのチェックがなされます。特にメッセージが表示されなければ、そのままリストに名前をつけます。任意の名前でもデフォルトの名前のままでも構いません。

今回はチュートリアルなので 'demo' と入力し, 'Save a list of 31 Genes' ボタンをクリックしてください。



• List Analysis Page

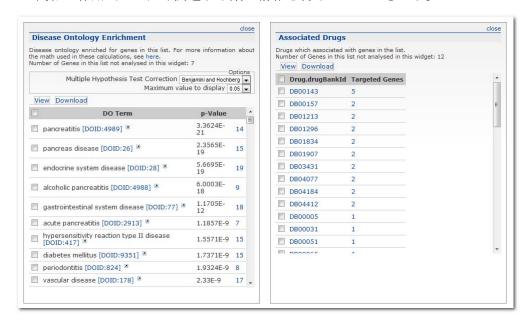
'List Analysis' ページでは、リスト中の各メンバーの詳細など多くの情報を得ることができます。リストが作成されると自動的に 'List Analysis' ページが表示されます。一度作ったリストについては、'Lists' タブの 'View' をクリックすると、いつでもこのページに戻ることができます (下図)。

List Analysis は3つのセクション (List, Widgets, Templates) に分かれています。



● 画面上部のボックス (List) には概要として, リスト中のメンバーの最初の 10 個が表示されます。そこには, リストの内容や情報, 変換ツールやリンク先(存在する場合のみ)が表示されます。これらの表の上部には Export/Edit オプションと検索ボックスがあります。

● 画面中部 (Widgets) にある一連のボックスからはリストについてのより多くの情報を得ることができます。例えば、このリストと特に強く関係付けられる KEGG パスウェイ、遺伝子オントロジー (GO) ターム、創薬や疾患に関するオントロジー、タンパク質間相互作用 (PPIs) や関連する薬の情報を得ることができます。



● 画面下部(Templates)には、このリストについて、あらかじめ実行された各種検索(テ ンプレートクエリ)の結果が表示されます。



III. テンプレートクエリ

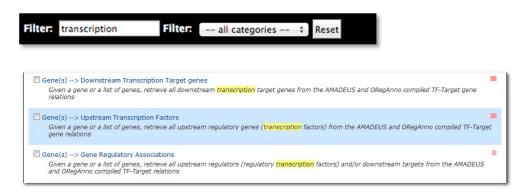
多様な検索(解析)を素早く簡単に実行させるために、TargetMine では事前に定義されたテンプレートクエリのライブラリを持っています。一般に、TargetMine の検索は、あるクラスについて条件を指定して(例えば、"遺伝子名が TGFB1 である")、関連するクラスのデータ(例えば、知られている全ての相互作用タンパク質)を表示する、という形式をとります。テンプレートライブラリは、そのような検索を多数用意したもので、検索例中の一部を変更することもできます。(TargetMine では、条件を指定することを「制限をかける」constrain という言葉で表現します。)

テンプレートクエリのアクセス方法は複数あります:

- テンプレートページには TargetMine のトップページ, もしくは上部のメニューバー のタブからアクセスしてテンプレートの検索や閲覧をすることが可能です。
- テンプレートはレポートページと List Analysis Page にも表示されます。



例としてテンプレートクエリに 'transcription' と入力し、検索します。



この検索は、与えられた遺伝子または遺伝子リストについて、上流の転写因子 (AMADEUS または OregAnno データベースに収納されている情報)をすべて取得するというものです。

オブジェクト(特定の遺伝子やタンパク質など)またはリストを指定する(「特定のオブジェクトまたはリストに制限する」)という型のテンプレートクエリでは、オブジェクトの ID を入力するボックスと共に、リストを制限するためのオプションが提供されます。

そして、テンプレートは 'Gene(s) --> Upstream Transcription Factors' をクリックしてください。

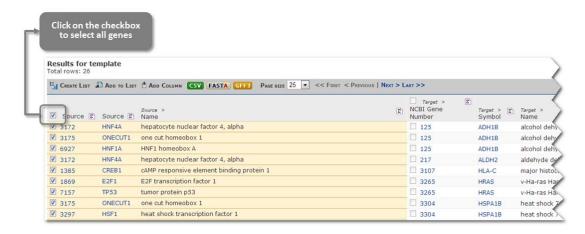


チェックボックスをクリックし、先ほど保存したサンプルリスト "demo" を選択してください。次に、"Show Results" ボタンを押し、検索を実行します。

その結果、先ほど作ったリストの遺伝子をターゲットとする転写因子が表示されます。

結果の一覧から新しいリストを作ることもできます。この検索結果から List Analysis ページや他のテンプレートに使用するための転写因子のリストを保存できます。

"Source" から全ての遺伝子を選択してください。



"CREATE LIST" タブをクリックし,名前をつけてリストを保存してください。(例: 'upstream_TFs')



一群のオブジェクト全体の情報を得るために,リストの使用は大変効果的です。この方法 を使うと以下の様なリストも作成できます。

- 選択した遺伝子(産物)の相互作用タンパク質
- これらの遺伝子の疾患オントロジー (DO) アノテーション

他にも色々と試してみてください。

IV. クエリビルダー

ここからは上級者向け応用編です。 'クエリビルダー'を使用すると,以下のようなことができます。

● TargetMine 上の複数のデータ・ソースを任意に組み合わせた新しいクエリの構築

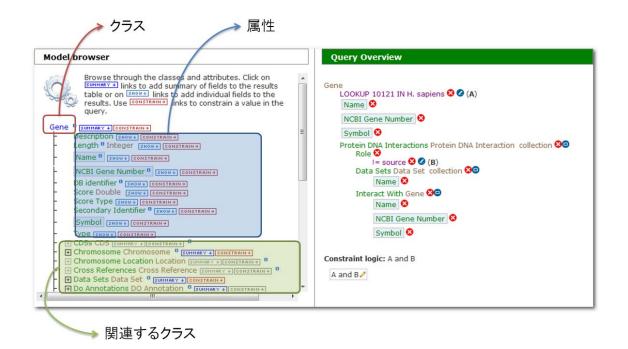
● 既存のテンプレートクエリを修正することで出力を変え、Constraints を加える。 以下のような例から実際に操作をしてみてください。

先ほど使用したテンプレート ' Gene(s) --> Upstream Transcription Factors' を選択してください。 クエリビルダーを実行するために、 'Edit Query' をクリックしてください。

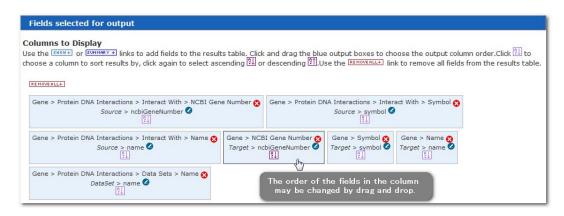


左側の 'Model browser' を用いて、データモデルをユーザーが操作することができます。この例では、'Model browser' 上には、Gene (遺伝子) と名付けられたクラスとその全ての属性 (例:名前、シンボル) および関連するクラス (例:生物種、タンパク質) が表示されます。TargetMine では、オブジェクトベースのデータモデルを使用し、生物学的なタイプは 'クラス'、個々のエントリは 'オブジェクト' と呼んでいます。

関連するクラスと属性を見るためには 出 をクリックしてください。 SHOW タブは出力にフィールドを追加するために使用します。 CONSTRAIN+ タブでは絞り込み(フィルタ)を追加することができます。



Model Browser の下には、ユーザーが選択したフィールドが '**Fields selected for output**' として表示されます。それぞれのフィールドは、検索結果ページの各列に対応します。フィールドの順序はドラッグ・アンド・ドロップで変更することができます。



右側の 'Query Overview'では、クエリに適用された絞り込み条件を示しています。その内の 1 つは、'Gene'で 'LOOKUP 10121 IN H. sapiens'と書かれているものです。これは、ID 10121 を持つヒトの遺伝子のみにクエリが絞り込まれた(制限された)ということを示しています。

もし、この遺伝子をターゲットにするものだけでなく、全ての転写因子について調べたく なった時には、絞り込み条件を外して、もう一度クエリを走らせれば解除できます。

絞り込み条件を解除するためには,'LOOKUP 10121 IN H. sapiens' の隣にある ♡ ボタンをクリックしてください。

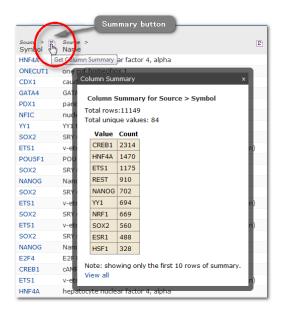


そして、 'Show results' をクリックしてください。



そうすると、(TargetMine に格納されている)全ての転写因子とそのターゲット遺伝子を得ることができます。

がサマリーボタンをクリックするとその列に関する概要の情報を得ることができます。



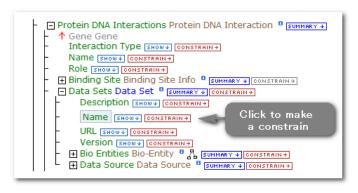
もう少し他の操作もしてみましょう。

'Query' をクリックして、クエリビルダーに戻ります。



それでは、"Gene → Protein DNA Interactions → Data Sets → Name" の順でパスを実行してみてください。次に Name の横の ONSTRAIN ションをクリックするとポップアップが表示されます。'ORegAnno' をドロップダウンリストから選択し 'Add to query' ボタンをクリックします。

そうするとデータセットに 'ORegAnno' データベースの絞り込みをかけることができます。

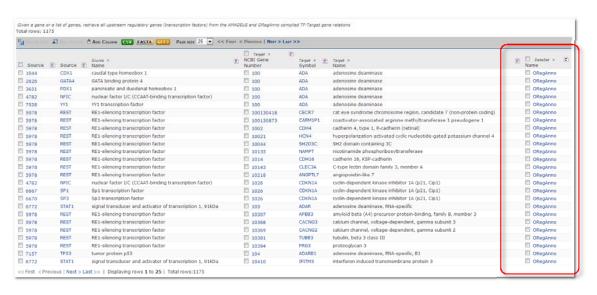




Model browser

Query overview

'Show results' をクリックし、もう一度実行してみましょう。



検索結果には、ORegAnno データベースから取得した転写因子の情報のみが表示されます。

このようにクエリビルダーを使うと細かく条件を指定して調べることが出来るので, ぜひ 利用してみてください。