

H-InvDB 遺伝子リスト特徴抽出ツール: HEAT

今西 規 (産総研・バイオメディカル情報研究センター)

HEAT: A new tool for gene set enrichment analysis using comprehensive annotation of human genes in H-InvDB.

Tadashi Imanishi (BIRC, AIST)

H-InvDB 遺伝子リスト特徴抽出ツール(英語名 H-InvDB Enrichment Analysis Tool、略称 HEAT)は、ヒト遺伝子と転写産物に関する統合データベース H-InvDB が持つ豊富なアノテーションを利用した Gene Set Enrichment Analysis (GSEA) のための新しい解析ツールである。HEAT は、ユーザが指定したヒト遺伝子の集合(遺伝子リスト)の中に有意に高い頻度で出現する H-InvDB のアノテーション項目を探し出すことができる。遺伝子の指定は、HUGO gene symbol、アクセッション番号(DDBJ/EMBL/GenBank)、H-InvDB の HIT および HIX を用いることができる。その他の遺伝子名の ID を使いたい場合は、ID 一括変換システム(<http://biodb.jp/>)によってあらかじめ ID 変換しておくことよ。統計学的な検定にはフィッシャーの正確確率検定を用いている。解析対象となるアノテーション項目は、染色体バンド、遺伝子ファミリー、Gene Ontology (GO)、機能ドメイン(InterPro)、構造ドメイン(SCOP)、KEGG パスウェイ、細胞内局在予測(Wolf PSORTによる)、組織特異的遺伝子発現(H-ANGEL データベースによる)である。HEAT は、他の GSEA 用ツールと比べてより多くのアノテーション項目を扱うことができる点が優れている。ゲノムワイド相関解析の結果解釈やマイクロアレイ実験のデータ解析等の場面で、HEAT は遺伝子リストに共通な特徴を発見することに役立つと考えられる。HEAT は、<http://hin.jp/HEAT/> にて無償で利用できる。

H-InvDB 遺伝子リスト特徴抽出ツール

H-InvDB Enrichment Analysis Tool (HEAT)

入力された遺伝子セットに
共通な特徴を探すツール。

解析対象となるアノテーション項目：
 染色体バンド、
 遺伝子ファミリー、
 Gene Ontology (GO)、
 機能ドメイン(InterPro)、
 立体構造ドメイン(SCOP)、
 KEGG パスウェイ、
 細胞内局在予測(Wolf PSORT)、
 組織特異的遺伝子発現(H-ANGEL)

H-InvDB Enrichment Analysis Tool

H-InvDB 遺伝子リスト特徴抽出ツール

遺伝子リスト投入画面

H-InvDB 遺伝子リスト特徴抽出ツール(H-InvDB Enrichment Analysis Tool (HEAT))は、ヒト遺伝子の集合(遺伝子リスト)に対して、その特徴を体系的に特定するデータマイニングツールです。H-InvDBのさまざまなアノテーション項目について、ユーザが入力した遺伝子リストの中に平均より有意に高い頻度で出現する項目を見つけ出します。この手法は一般にGene Set Enrichment Analysis (GSEA) と呼ばれ、マイクロアレイ実験のデータ解析等によく使われます。統計学的な検定にはフィッシャーの正確確率検定を用いています。<[詳細はこちら](#)>

遺伝子リストの投入 (ヒト遺伝子のIDをスペースまたはカンマ区切りで入力してください)

(ここに遺伝子リストを投入)

アップロード: [テキストファイルからアップロードする](#)

入力の種類: HIT HIX HUGO Gene Symbol Accession Number

実行ボタン →

Copyright(C)2009 AIST and J.BIC. All Rights Reserved.

統合データベースからの
知識発見支援ツール

http://hin.jp/heat/ にて公開中。

図. HEAT システム (<http://hin.jp/HEAT/>) の概要。

引用文献

- 1) Yamasaki C, Murakami K, Takeda J, Sato Y, Noda A, Sakate R, Habara T, Nakaoka H, Todokoro F, Matsuya A, Imanishi T, and Gojobori T (2009) H-InvDB in 2009, extended database and data mining resources for human genes and transcripts. *Nucleic Acids Research* 38 (Database Issue): gkp1020.
- 2) Imanishi T and Nakaoka H (2009) Hyperlink Management System and ID Converter System: enabling maintenance-free hyperlinks among major biological databases. *Nucleic Acids Research* 37 (Web Server Issue): W17-W22 (gkp355).