遺伝子発現データベース　GENEVESTIGATOR

Sample Composition Tools と収載データのアップデート

NEBION

GENEVESTIGATOR は遺伝子発現データベースのオンライン解析ツールです。公共データベースに登録されたマイクロアレイや次世代シーケンサー（NGS）の膨大な遺伝子発現データをキュレートすることで、様々な研究者により登録された大量の実験結果を統合して解析可能にします。また、GENEVESTIGATOR は使いやすいインターフェースと高速な検索エンジンを搭載しているため、研究者が標的遺伝子の探索などの遺伝子発現解析を行う際に、注目する遺伝子の同定や発現変動遺伝子の優先順位付けなどを簡単かつ正確に行うことができます。本稿では新しい解析ツールである Sample Composition Tools や新たに追加された遺伝子発現データについて紹介します。

Sample Composition Tools

GENEVESTIGATOR には多数の実験結果における遺伝子発現を横断的に解析できる COMPENDIUM-WIDE ANALYSIS と、個々の実験における遺伝子発現の解析を行う SINGLE EXPERIMENT ANALYSIS の 2 つの機能が搭載されています。SINGLE EXPERIMENT ANALYSIS では、実験群と対照群の発現変動遺伝子を探索する Differential Expression や多次元データを 2 次元平面上に t 分布型確率的近傍埋め込み法（t-SNE）で可視化を行う Dimension Reduction などの解析ツールを利用できます。

Sample Composition Tools は SINGLE EXPERIMENT ANALYSIS に新しく追加された解析ツールです。Sample Composition Tools はサンプル中の異なる細胞タイプの組成やその相対量、および疾患や治療、発達などの異なるサンプル特性をグラフに表示することができます。そのため異なるサンプル特性同士の相関関係を解析することが可能です。

ヒトの肺胞マクロファージにおいて正常細胞と新型コロナウイルス（SARS-CoV-2）に感染した細胞を single cell RNA-seq で解析したデータセットを使用しました。図 1 は、新型コロナウイルス感染症（COVID-19）患者の重篤度ごとに各年齢の割合を表示した Sample Composition Tools の結果です。60–70 歳が赤色、50–60 歳が青色、40–50 歳が黄色、30–40 歳が緑色、20–30 歳がピンク色で色付けされています。COVID-19 の症状が軽症の人のデータや COVID-19 に罹患していない人のデータでは比較的年齢が若い人が多く、COVID-19 の症状が重篤な人のデータでは比較的高齢な人が多いことが分かります。この結果から、COVID-19 の重篤度と年齢には相関関係があることが分かります。その他にも細胞に発現しているバイオマーカーや性別、他の疾患情報などの複数の特性との関連性を調べることで、さらに分析を拡張できる可能性があります。

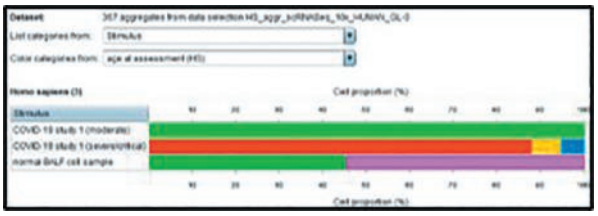

図 1. 年齢と COVID-19 の重症度の Sample Composition Tools

遺伝子発現データの追加収載

GENEVESTIGATOR は公共データベースに登録されたマイクロアレイや RNA-seq などのバルク組織の遺伝子発現データや single cell RNA-seq の遺伝子発現データをキュレートし、統合したデータを利用した解析が可能です。2021 年 10 月現在、バルク組織の遺伝子発現データは、134 研究の合計 11,451 サンプルを利用することが可能です。バルク組織の遺伝子発現データでは、筋萎縮性側索硬化症のデータ、さらに自己免疫疾患や COVID-19、免疫学領域の研究データなどが追加されました（表 1）。特にがん関連の遺伝子発現データベースである The Cancer Genome Atlas（TCGA）の遺伝子発現データでは、がん患者の生存期間だけでなく死因やがん発生中の生存期間、がん再発のイベント、無増悪期間などのパラメーターが追加されています。これらのパラメーターは TCGA の論文[1] を参考に登録されています。生物種は Homo sapiens（HS）や Mus musculus（MM）や Rattus norvegicus（RN）の 3 種類を利用することができます。Single cell RNA-seq の遺伝子発現データは、19 の研究データが追加されて、合計 138 研究の 570 万細胞のデータが利用可能です。

表 1. 追加されたバルク組織の遺伝子発現データ

Focus area	Organisms	# studies*	# samples*
Cardiovascular System	HS	3	148
Developmental and Regenerative Systems	HS	3	145
Digestive and Endocrine Systems	HS	24	2461
Immune System	HS	43	3283
Nervous System	HS	14	1681
Oncology	HS MM	28	1807
Pharmacology and Toxicology	HS MM RN	3	25
Reproductive and Renal Systems	HS	9	552
Respiratory System	HS	20	2503
Various Systems	HS	7	552

ご試用

GENEVESTIGATOR は、無償トライアルにより 30 日間ご評価いただけます。遺伝子発現解析をされる方はぜひ GENEVESTIGATOR をお試し下さい。トライアルをご希望の方は弊社ウェブサイトよりお問い合わせ下さい。

https://www.molsis.co.jp/contact/

1）Liu, J.; et al. *Cell* **2018,** *173* (2), 400–416.