

Working with public gene expression data

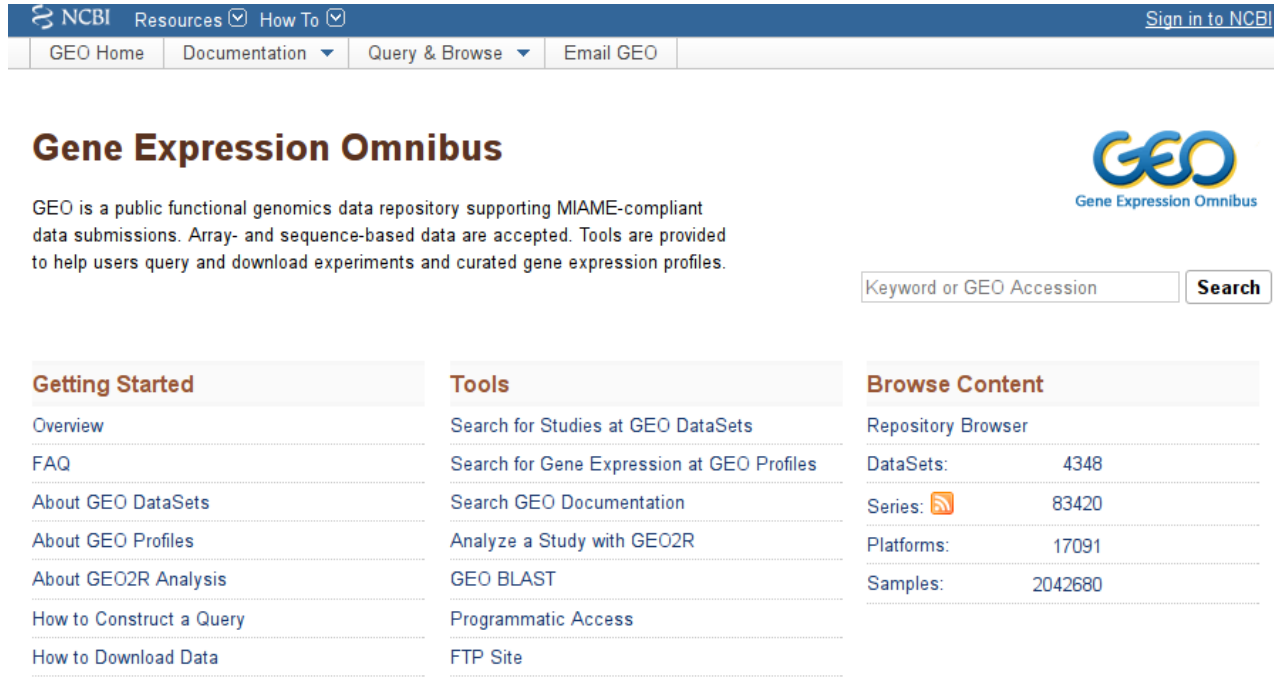
Alexey Sergushichev

Sep 23, Nice

Outline

- ✓ Exploring gene expression datasets
- ✓ Simple analysis methods
- ✓ **Working with public datasets**

Gene Expression Omnibus: the most common repository of gene expression




The screenshot shows the NCBI GEO website. At the top, there is a navigation bar with "NCBI Resources" and "How To" menus, and a "Sign in to NCBI" link. Below this is a secondary navigation bar with "GEO Home", "Documentation", "Query & Browse", and "Email GEO". The main content area features the "Gene Expression Omnibus" title and logo, a brief description of the repository, and a search box. Three columns of links are provided: "Getting Started", "Tools", and "Browse Content".

Gene Expression Omnibus

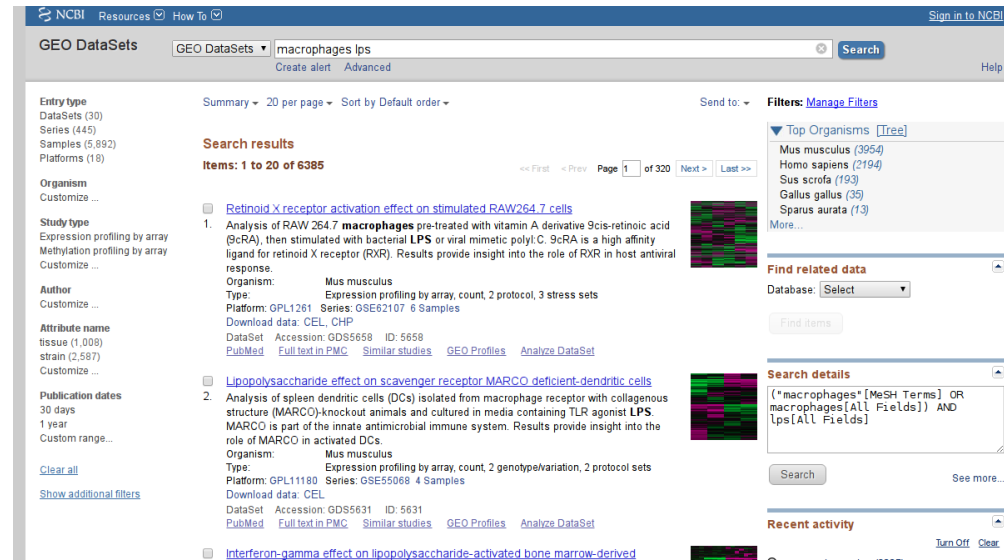
GEO is a public functional genomics data repository supporting MIAME-compliant data submissions. Array- and sequence-based data are accepted. Tools are provided to help users query and download experiments and curated gene expression profiles.

Keyword or GEO Accession

| Getting Started | Tools | Browse Content |
|--|--|---|
| Overview | Search for Studies at GEO DataSets | Repository Browser |
| FAQ | Search for Gene Expression at GEO Profiles | DataSets: 4348 |
| About GEO DataSets | Search GEO Documentation | Series:  83420 |
| About GEO Profiles | Analyze a Study with GEO2R | Platforms: 17091 |
| About GEO2R Analysis | GEO BLAST | Samples: 2042680 |
| How to Construct a Query | Programmatic Access | |
| How to Download Data | FTP Site | |

You can search GEO for datasets

- ✓ Go to GEO datasets:
<https://www.ncbi.nlm.nih.gov/gds>
- ✓ Search for “macrophages LPS”



The screenshot shows the NCBI GEO DataSets search results for the query "macrophages lps". The search results are displayed in a list format, showing the first three items. Each item includes a title, a brief description, and various metadata fields such as Organism, Type, Platform, and Series. The first item is "Retinoid X receptor activation effect on stimulated RAW264.7 cells", the second is "Lipopolysaccharide effect on scavenger receptor MARCO deficient-dendritic cells", and the third is "Interferon-gamma effect on lipopolysaccharide-activated bone marrow-derived". The page also features a sidebar with filters for Entry type, Organism, Study type, Author, Attribute name, and Publication dates. On the right side, there are sections for "Top Organisms", "Find related data", "Search details", and "Recent activity".

Result close view

Bone marrow derived macrophages

- [Interferon- \$\gamma\$ and lipopolysaccharide treatment effect on bone marrow derived macrophages](#)

7. Analysis of bone marrow derived **macrophages** (BMDM) treated with interferon- γ (IFN- γ), lipopolysaccharide (**LPS**) or both. IFN- γ and **LPS** stimulate BMDM production of reactive oxygen species (ROS) in immune defense response. Results provide insight into regulatory factors for ROS production in BMDM.

Organism: Mus musculus

Type: Expression profiling by array, count, 4 protocol sets

Platform: GPL1261 Series: GSE53986 16 Samples

Download data: CEL, CHP

DataSet Accession: GDS5196 ID: 5196

[PubMed](#) [Similar studies](#) [GEO Profiles](#) [Analyze DataSet](#)

LPS-stimulated (and IFN γ)

Affymetrix Mouse
Genome 430 2.0

GSE identifier

GEO search options

NCBI Resources How To Sign in to NCBI


GEO DataSets Search

Summary 20 per page Sort by Default order

Send to: Filters: [Manage Filters](#)

[DataSets \(30\)](#)
[Series \(445\)](#)
[Samples \(5,892\)](#)
[Platforms \(18\)](#)

Search results
 Items: 1 to 2 of 320

Organism 

[Retinoid X receptor activation effect on stimulated RAW264.7 cells](#)
 1. Analysis of RAW 264.7 **macrophages** pre-treated with vitamin A derivative 9cis-retinoic acid (9cRA), then stimulated with bacterial LPS or viral mimetic poly:I:C. 9cRA is a high affinity ligand for retinoid X receptor (RXR). Results provide insight into the role of RXR in host antiviral response.
 Organism: Mus musculus
 Type: Expression profiling by array, count, 2 protocol, 3 stress sets
 Platform: GPL1261 Series: GSE62107 6 Samples
 Download data: CEL, CHP
 DataSet Accession: GDS5658 ID: 5658
[PubMed](#) [Full text in PMC](#) [Similar studies](#) [GEO Profiles](#) [Analyze DataSet](#)

[Lipopolysaccharide effect on scavenger receptor MARCO deficient-dendritic cells](#)
 2. Analysis of spleen dendritic cells (DCs) isolated from macrophage receptor with collagenous structure (MARCO)-knockout animals and cultured in media containing TLR agonist LPS. MARCO is part of the innate antimicrobial immune system. Results provide insight into the role of MARCO in activated DCs.
 Organism: Mus musculus
 Type: Expression profiling by array, count, 2 genotype/variation, 2 protocol sets
 Platform: GPL11180 Series: GSE55068 4 Samples
 Download data: CEL
 DataSet Accession: GDS5631 ID: 5631
[PubMed](#) [Full text in PMC](#) [Similar studies](#) [GEO Profiles](#) [Analyze DataSet](#)

[Interferon-gamma effect on lipopolysaccharide-activated bone marrow-derived](#)

Top Organisms [\[Tree\]](#)
 Mus musculus (3954)
 Homo sapiens (2194)
 Sus scrofa (193)
 Gallus gallus (35)
 Sparus aurata (13)
[More...](#)

Find related data
 Database:

Search details
 {"*macrophages" [MeSH Terms] OR
 macrophages[All Fields]} AND
 lps[All Fields]
 [See more...](#)

Recent activity

 macrophages lps (6385)

GEO profiles for searching by individual genes

- ✓ Go to <https://www.ncbi.nlm.nih.gov/geoprofiles/>
- ✓ Search for Acod1



NCBI Resources How To Sign in to NCBI

GEO Profiles GEO Profiles Acod1 Search

Gene symbol Summary 20 per page Sort by Subgroup effect Send to Filters: Manage Filters

Gene keyword Search results Profile data

Organism Items: 1 to 20 of 2722 << First < Prev Page 1 of 137 Next > Last >> Download profile data

Gene ontology 1. [Acod1 - Macrophage response to viral infection](#) Profile pathways Find pathways

Differential expression 1. Annotation: **Acod1**, aconitate decarboxylase 1
 Organism: Mus musculus
 Reporter: GPL81, 98773_s_at (ID_REF), GDS1271, 16365 (Gene ID)
 DataSet type: Expression profiling by array, transformed count, 4 samples
 ID: 11404830
[GEO DataSets](#) [Gene](#) [UniGene](#) [Profile neighbors](#) [Chromosome neighbors](#) [Homologene neighbors](#)

Up/down genes 2. [Acod1 - MyD88 deficient macrophage response to zymosan](#) Find related data

DataSet keyword 2. Annotation: **Acod1**, aconitate decarboxylase 1
 Organism: Mus musculus
 Reporter: GPL1261, 1427381_at (ID_REF), GDS26864988 (Gene ID), L38281
 DataSet type: Expression profiling by array, transformed count, 4 samples
 ID: 37834587
[GEO DataSets](#) [Gene](#) [UniGene](#) [Profile neighbors](#) [Chromosome neighbors](#) [Homologene neighbors](#)

Customize ... Database: Select

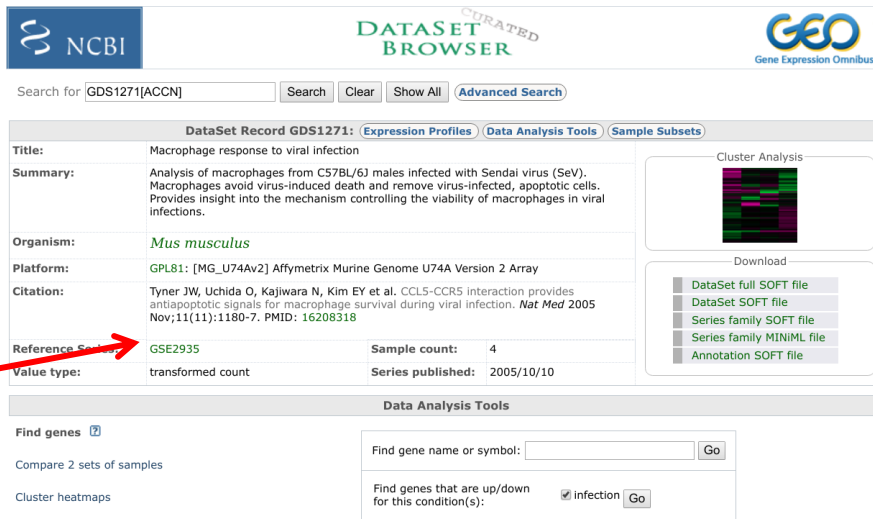
GEO accession Clear all Search details

Customize ... Show additional filters Acod1[All Fields]

GDS ID, not GSE

GDS is a curated GSE

- ✓ Click on any GDS ID on previous screen



NCBI CURATED DATASET BROWSER GEO Gene Expression Omnibus

Search for Search Clear Show All Advanced Search

DataSet Record GDS1271: Expression Profiles Data Analysis Tools Sample Subsets

Title: Macrophage response to viral infection

Summary: Analysis of macrophages from C57BL/6J males infected with Sendai virus (SeV). Macrophages avoid virus-induced death and remove virus-infected, apoptotic cells. Provides insight into the mechanism controlling the viability of macrophages in viral infections.

Organism: *Mus musculus*

Platform: GPL81: [MG_U74Av2] Affymetrix Murine Genome U74A Version 2 Array

Citation: Tyner JW, Uchida O, Kajiwara N, Kim EY et al. CCL5-CCR5 interaction provides antiapoptotic signals for macrophage survival during viral infection. *Nat Med* 2005 Nov;11(11):1180-7. PMID: 16208318

Reference Series: [GSE2935](#) **Sample count:** 4

Value type: transformed count **Series published:** 2005/10/10

Data Analysis Tools

Find genes [?](#)

Compare 2 sets of samples

Cluster heatmaps

Find gene name or symbol: Go

Find genes that are up/down for this condition(s): infection Go

Cluster Analysis

Download

- DataSet full SOFT file
- DataSet SOFT file
- Series family SOFT file
- Series family MINIML file
- Annotation SOFT file

GSE ID

Both GDS and GSE IDs can be used in Phantasus

GeneQuery: expression-based phenotype searching engine

- ✓ Go to <http://ctlab.itmo.ru/genequery/searcher/>
- ✓ <https://artyomovlab.wustl.edu/genequery/searcher/>

GeneQuery^α

Search

Example

Downloads

About

Contacts

Database species:

- Homo Sapiens Mus Musculus
 Rattus Norvegicus

Query species:

- Homo Sapiens Mus Musculus
 Rattus Norvegicus

Gene list (separated by newline/whitespace/tab)

Search

Run example ▾

Searching for datasets in GeneQuery

- ✓ Paste top 250 genes up-regulated on LPS
- ✓ Query species <- mouse
- ✓ Database species <- mouse

GeneQuery^a

Database species: Homo Sapiens Mus Musculus Rattus Norvegicus

Query species: Homo Sapiens Mus Musculus Rattus Norvegicus

Gene list (separated by newline/whitespace/tab)

```

Cd38
H2-Q6///LOC68395///H2-Q8
Isg15///Gm9706
Saa3
Fpr2
Usp18
AW112010
Zbp1
Serpinb2
Rsad2
Clec4e
Ifit1
Slfn4
Adgb
Pydc4
Bst1
Herc6
Gbp2
Il1b
Cd274

```

Search

Run example ▾

GeneQuery: results page

| | | | |
|--|---------|---|--------|
| Modules | 1707 | Detected gene format | SYMBOL |
| Detected groups | 22 | Genes entered | 255 |
| Min log₁₀(adj.p.value) | -120.70 | Unique entrez IDs | 237 |
| | | <input type="button" value="show gene conversion table"/> | |
| | | Apply orthology | no |

| | |
|---|--|
| <input type="button" value="Export result to CSV"/> | <input type="checkbox"/> Group results |
|---|--|

| # | Experiment title | Module | log ₁₀ (adj.p.value) | Overlap | GSE | GMT |
|----|---|--------|---------------------------------|----------|----------|-------------------|
| 1 | SM001: SARS CoV MA15 infection of C57Bl/6 mouse model – Data from 4 viral doses at 1, 2, 4 and 7 days post infection. | 3 | -120.70 | 124/370 | GSE33266 | 🔗 |
| 2 | Transcriptomic definition of curcumin treated microglia | 4 | -101.66 | 116/498 | GSE23639 | 🔗 |
| 3 | ATF7 mediates lipopolysaccharide-induced epigenetic changes in macrophages involved in innate immunological memory | 6 | -95.74 | 113/310 | GSE71113 | 🔗 |
| 4 | Gene expression by Retinoic Acid in mouse Dendritic Cells | 5 | -93.10 | 116/321 | GSE34324 | 🔗 |
| 5 | Gene expression comparison of post-sepsis versus nave mice-derived tumor associated macrophages | 2 | -92.36 | 141/774 | GSE64498 | 🔗 |
| 6 | Expression data from polarized macrophages: effect of p53 | 2 | -89.31 | 187/1569 | GSE53321 | 🔗 |
| 7 | Interferon response genes as immunopathogenic correlates of SARS coronavirus infection in mice (129/S6/SvEv) | 6 | -87.32 | 96/327 | GSE19137 | 🔗 |
| 8 | Transcriptional response in skin of mouse 24 hours after intradermal infection with Trypanosoma cruzi | 3 | -86.69 | 103/418 | GSE13522 | 🔗 |
| 9 | An essential role for the antiviral endoribonuclease, RNase-L, in antibacterial immunity. | 4 | -84.87 | 111/385 | GSE13530 | 🔗 |
| 10 | Mechanisms Establishing TLR4-Responsive Activation States of Inflammatory Response Genes | 4 | -84.10 | 129/575 | GSE23622 | 🔗 |
| 11 | Expression data from bone marrow cultures treated with adiponectin - evidence for lipopolysaccharide contamination | 4 | -83.95 | 106/345 | GSE35466 | 🔗 |
| 12 | SM007 - Infection with SARS MA15 of C57BL6/J mice and CXCR3 knockouts in the same strain of mice. | 2 | -79.16 | 126/752 | GSE50878 | 🔗 |

GeneQuery-DE

- ✔ Similar to genequery, but the modules are based on differential expression
- ✔ Early experimental version
- ✔ (Don't) go to <https://ctlab.itmo.ru/genequery-de/searcher/>
- ✔ Paste LPS signature genes

GeneQueryDE^α

Database species: Homo Sapiens Mus Musculus Rattus Norvegicus

Query species: Homo Sapiens Mus Musculus Rattus Norvegicus

Gene list (separated by newline/whitespace/tab)

```

Mb21d1
Gstt1
Eif2ak2
LOC100862473//Gm16340//Ifi203
Slc2a1
Trim34b//Trim34a
Rab20
Pstpip2
Tcstv3
Tapbp1
Lrrc8a//Phyhd1
Trib3
Spic
Slc16a10
Mcoln2
Ms4a6c
Pyhin1
Slc13a3
Cst7
Slc25a37
  
```

Search

Run example ▾

GeneQuery-DE results

| # | Experiment title | Module | $\log_{10}(\text{adj. } p\text{-value})$ | Overlap | GSE | Dif exprs |
|----|---|--|--|---------|----------|-------------------|
| 1 | NRROS negatively regulates ROS in phagocytes during host defense and autoimmunity | up in 'treatment.LPS' compared to 'treatment.Untreated' | -130.48 | 114/200 | GSE53986 | 🔗 |
| 2 | Dysregulation of Macrophage Activation Profiles by Engineered Nanoparticles | up in 'treatment.Silica.LPS' compared to 'treatment.control' | -110.17 | 99/200 | GSE44294 | 🔗 |
| 4 | Host cell gene expression in Human respiratory syncytial virus (HRSV) infected mouse macrophage cells at 4, 24 hours post infection | up in 'time.24.hpi' compared to 'time.0.hpi' | -108.57 | 101/200 | GSE31378 | 🔗 |
| 3 | Host cells infected by various types of viruses | up in 'time.24.hpi' compared to 'time.0.hpi' | -108.57 | 101/200 | GSE31524 | 🔗 |
| 6 | Dysregulation of Macrophage Activation Profiles by Engineered Nanoparticles | up in 'treatment.LPS' compared to 'treatment.control' | -108.24 | 98/200 | GSE44294 | 🔗 |
| 5 | Dysregulation of Macrophage Activation Profiles by Engineered Nanoparticles | up in 'treatment.SPIO.LPS' compared to 'treatment.SPIO' | -108.24 | 98/200 | GSE44294 | 🔗 |
| 8 | Dysregulation of Macrophage Activation Profiles by Engineered Nanoparticles | up in 'treatment.Silica.LPS' compared to 'treatment.SPIO' | -106.32 | 97/200 | GSE44294 | 🔗 |
| 7 | Dysregulation of Macrophage Activation Profiles by Engineered Nanoparticles | up in 'treatment.LPS' compared to 'treatment.Silica' | -106.32 | 97/200 | GSE44294 | 🔗 |
| 10 | Dysregulation of Macrophage Activation Profiles by Engineered Nanoparticles | up in 'treatment.Silica.LPS' compared to 'treatment.Silica' | -106.32 | 97/200 | GSE44294 | 🔗 |
| 9 | Dysregulation of Macrophage Activation Profiles by Engineered Nanoparticles | up in 'treatment.LPS' compared to 'treatment.SPIO' | -106.32 | 97/200 | GSE44294 | 🔗 |
| 11 | IM001: Influenza A/VN/1203/04 infection of C57Bl/6 mouse model - | up in 'treatment.Influenza.A/VN/1203/04.infected.with.10*3.PFU' compared to 'treatment.Time.matched.mock' on 'time.2.day' background | -105.57 | 94/200 | GSE33263 | 🔗 |

RNA-seq data deposition is not as standardized as microarrays

- ✓ Different protocols: 3', full-length, 5'
- ✓ Different quantification methods:
 - alignment: expectation minimization vs counting
 - alignment-free quantification
 - ...

Let's look at GSE110749 in GEO

Series GSE110749 [Query DataSets for GSE110749](#)

| | |
|-------------------|---|
| Status | Public on Apr 21, 2018 |
| Title | Dimethyl itaconate inhibits secondary wave of NF-κB signaling in macrophage activation |
| Organism | Mus musculus |
| Experiment type | Expression profiling by high throughput sequencing |
| Summary | Natural metabolite itaconate and its membrane permeable derivative dimethyl itaconate (DI) selectively inhibit a subset of cytokines during macrophage activation (e.g. IL-1β, IL-6, IL-12 but not TNF-α). Selectivity of DI action stems from the inhibitory effects of secondary, but not primary, wave of NF-κB signaling. |
| Overall design | Bone marrow-derived macrophages (BMDMs) from C57BL/6 mice were stimulated with LPS + IFNγ for indicated timepoints. Some samples were treated with dimethyl itaconate for 12h prior to LPS stimulation. |
| Contributor(s) | Bambouskova M , Lampropoulou V , Sergushichev A , Artyomov MN |
| Citation(s) | Bambouskova M, Gorvel L, Lampropoulou V, Sergushichev A et al. Electrophilic properties of itaconate and derivatives regulate the IκBζ-ATF3 inflammatory axis. <i>Nature</i> 2018 Apr;556(7702):501-504. PMID: 29670287 |
| Submission date | Feb 16, 2018 |
| Last update date | Nov 30, 2018 |
| Contact name | Maxim N. Artyomov |
| E-mail | martyomov@pathology.wustl.edu |
| Organization name | Washington University in St.Louis |
| Department | Immunology&Pathology |
| Street address | 660 S. Euclid Avenue, Campus Box 8118 |
| City | St.Louis |
| State/province | MO |
| ZIP/Postal code | 63110 |
| Country | USA |

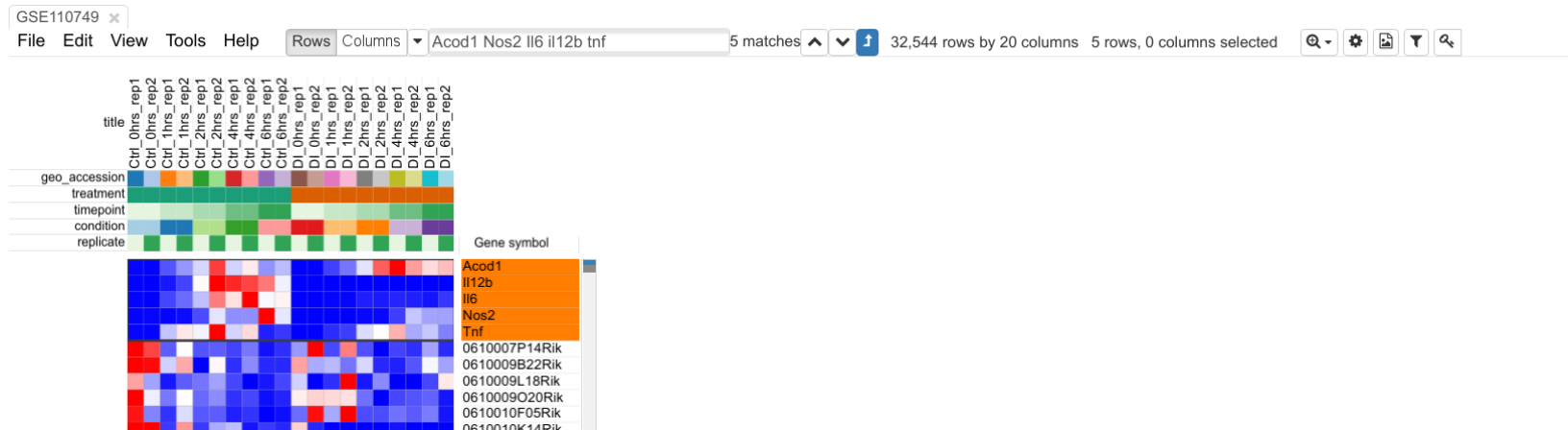
Platforms (1) [GPL17021](#) Illumina HiSeq 2500 (Mus musculus)

Samples (20) [Less...](#)

- [GSM3015930](#) Ctrl_0hrs_rep1
- [GSM3015931](#) Ctrl_0hrs_rep2
- [GSM3015932](#) Ctrl_1hrs_rep1
- [GSM3015933](#) Ctrl_1hrs_rep2
- [GSM3015934](#) Ctrl_2hrs_rep1
- [GSM3015935](#) Ctrl_2hrs_rep2
- [GSM3015936](#) Ctrl_4hrs_rep1
- [GSM3015937](#) Ctrl_4hrs_rep2
- [GSM3015938](#) Ctrl_6hrs_rep1
- [GSM3015939](#) Ctrl_6hrs_rep2
- [GSM3015940](#) DI_0hrs_rep1
- [GSM3015941](#) DI_0hrs_rep2
- [GSM3015942](#) DI_1hrs_rep1
- [GSM3015943](#) DI_1hrs_rep2
- [GSM3015944](#) DI_2hrs_rep1
- [GSM3015945](#) DI_2hrs_rep2
- [GSM3015946](#) DI_4hrs_rep1
- [GSM3015947](#) DI_4hrs_rep2
- [GSM3015948](#) DI_6hrs_rep1
- [GSM3015949](#) DI_6hrs_rep2

Open GSE110749 in Phantasus

✓ Search for markers (e.g., Acod1, Nos2, Tnf, Il6, Il12b)



What should be done after loading data?

What should be done after loading data?

- ✓ Check log-scale or linear
- ✓ Filter probes and genes
- ✓ Checking for outliers with k-means, PCA and markers

Let's look again at GSE110749 in GEO

| Series GSE110749 | | Query DataSets for GSE110749 |
|-------------------|---|------------------------------|
| Status | Public on Apr 21, 2018 | |
| Title | Dimethyl itaconate inhibits secondary wave of NF-κB signaling in macrophage activation | |
| Organism | Mus musculus | |
| Experiment type | Expression profiling by high throughput sequencing | |
| Summary | Natural metabolite itaconate and its membrane permeable derivative dimethyl itaconate (DI) selectively inhibit a subset of cytokines during macrophage activation (e.g. IL-1β, IL-6, IL-12 but not TNF-α). Selectivity of DI action stems from the inhibitory effects of secondary, but not primary, wave of NF-κB signaling. | |
| Overall design | Bone marrow-derived macrophages (BMDMs) from C57BL/6 mice were stimulated with LPS + IFNγ for indicated timepoints. Some samples were treated with dimethyl itaconate for 12h prior to LPS stimulation. | |
| Contributor(s) | Bambouskova M , Lampropoulou V , Sergushichev A , Artyomov MN | |
| Citation(s) | Bambouskova M, Gorvel L, Lampropoulou V, Sergushichev A et al. Electrophilic properties of itaconate and derivatives regulate the IκBζ-ATF3 inflammatory axis. <i>Nature</i> 2018 Apr;556(7702):501-504. PMID: 29670287 | |
| Submission date | Feb 16, 2018 | |
| Last update date | Nov 30, 2018 | |
| Contact name | Maxim N. Artyomov | |
| E-mail | martyomov@pathology.wustl.edu | |
| Organization name | Washington University in St.Louis | |
| Department | Immunology&Pathology | |
| Street address | 660 S. Euclid Avenue, Campus Box 8118 | |
| City | St.Louis | |
| State/province | MO | |
| ZIP/Postal code | 63110 | |
| Country | USA | |

Relations

BioProject [PRJNA434457](#)
 SRA [SRP133013](#)

| Download family | Format |
|---|--------------------------|
| SOFT formatted family file(s) | SOFT ? |
| MINiML formatted family file(s) | MINiML ? |
| Series Matrix File(s) | TXT ? |

| Supplementary file | Size | Download | File type/resource |
|---|----------|--|--------------------|
| GSE110749_counts.txt.gz | 419.1 Kb | (ftp) (http) | TXT |

Raw data are available in SRA

Processed data are available on Series record

Some datasets aren't processed in ARCHS4 but authors-provided table can be in supplemental

Working with text expression table

- ✓ Download counts file from GEO and unpack, or download GSE110749_counts.txt from Google Drive
- ✓ Open in Excel/Calc
- ✓ Add “treatment” line manually
- ✓ Save in text format as GSE110749_counts_ann.txt
 - or as XLSX file GSE110749_counts_ann.xlsx

| 1 | title | Ctrl_0hrs_rep1 | Ctrl_0hrs_rep2 | Ctrl_1hrs_rep1 | Ctrl_1hrs_rep2 | Ctrl_2hrs_rep1 | Ctrl_2hrs_rep2 | Ctrl_4hrs_rep1 | Ctrl_4hrs_rep2 | Ctrl_6hrs_rep1 | Ctrl_6hrs_rep2 | DI_0hrs_rep1 | DI_0hrs_rep2 | DI_1hrs_rep1 | DI_1hrs_rep2 |
|----|---------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|--------------|--------------|--------------|--------------|
| 2 | treatment | Ctrl | Ctrl | Ctrl | Ctrl | Ctrl | Ctrl | Ctrl | Ctrl | Ctrl | Ctrl | DI | DI | DI | DI |
| 3 | ENSMUSG000000000001 | 033 | 627 | 305 | 473 | 336 | 578 | 533 | 613 | 347 | 475 | 463 | 714 | 355 | 0 |
| 4 | ENSMUSG000000000003 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | ENSMUSG000000000028 | 29 | 26 | 13 | 30 | 6 | 10 | 2 | 3 | 0 | 0 | 6 | 27 | 8 | 0 |
| 6 | ENSMUSG000000000037 | 5 | 1 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | ENSMUSG000000000049 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | ENSMUSG000000000056 | 435 | 316 | 119 | 212 | 54 | 105 | 12 | 23 | 39 | 60 | 188 | 168 | 163 | 0 |
| 9 | ENSMUSG000000000058 | 148 | 142 | 82 | 186 | 60 | 95 | 31 | 49 | 40 | 36 | 122 | 214 | 110 | 0 |
| 10 | ENSMUSG000000000078 | 12109 | 8717 | 27631 | 52634 | 24103 | 51526 | 22181 | 24186 | 18669 | 20819 | 3523 | 8595 | 21196 | 0 |
| 11 | ENSMUSG000000000085 | 83 | 68 | 16 | 30 | 13 | 15 | 16 | 16 | 14 | 19 | 62 | 49 | 28 | 0 |
| 12 | ENSMUSG000000000088 | 1277 | 1069 | 586 | 824 | 384 | 938 | 523 | 722 | 379 | 495 | 730 | 526 | 413 | 0 |

Open the table in Phantasus

Open

Click the table cell containing the first data row and column.

Transpose

Data Matrix

Column Annotations

Row Annotations

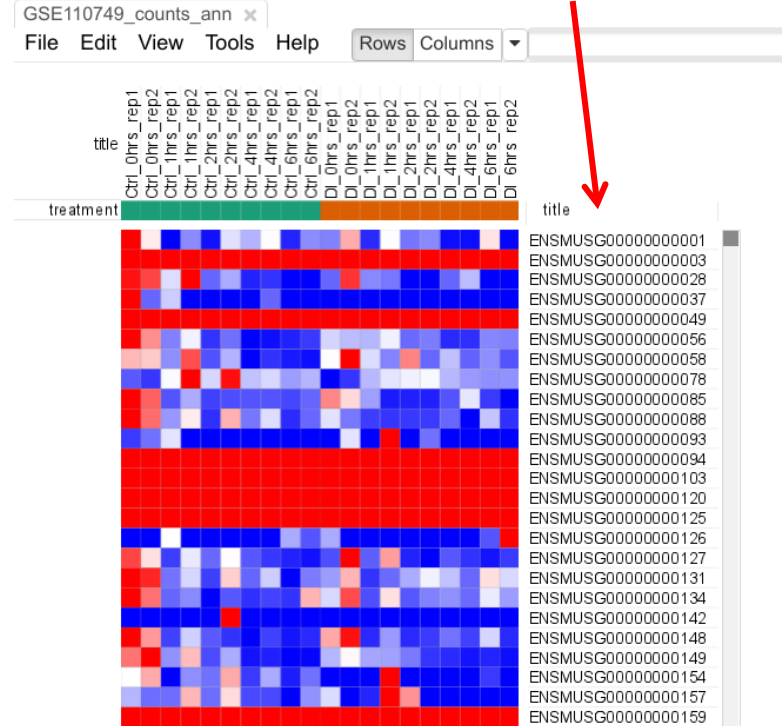
| title | Ctrl_0hrs_rep | Ctrl_0hrs_rep | Ctrl_1hrs_rep | Ctrl_1hrs_rep | Ctrl_2hrs_rep | Ctrl_2hrs_ |
|----------------|---------------|---------------|---------------|---------------|---------------|------------|
| treatment | Ctrl | Ctrl | Ctrl | Ctrl | Ctrl | Ctrl |
| ENSMUSG0(923 | 637 | 305 | 473 | 326 | 578 | |
| ENSMUSG0(0 | 0 | 0 | 0 | 0 | 0 | |
| ENSMUSG0(29 | 26 | 13 | 30 | 6 | 10 | |
| ENSMUSG0(5 | 1 | 2 | 0 | 0 | 0 | |
| ENSMUSG0(0 | 0 | 0 | 0 | 0 | 0 | |
| ENSMUSG0(435 | 316 | 119 | 212 | 54 | 105 | |
| ENSMUSG0(148 | 142 | 82 | 186 | 60 | 95 | |
| ENSMUSG0(12109 | 8717 | 27631 | 52634 | 24103 | 51526 | |
| ENSMUSG0(83 | 68 | 16 | 30 | 13 | 15 | |
| ENSMUSG0(1277 | 1069 | 586 | 824 | 384 | 938 | |

Click at the left-top corner of expression values

OK

Cancel

Need Ensembl to Gene symbol mapping



Annotating rows

- ✔ Going back to Phantasus
- ✔ Tools/Annotate/Annotate rows/From database

Annotate from AnnotationDB ✕

Specimen DB

org.Mm.eg.sqlite - Mus musculus ▼

Source column

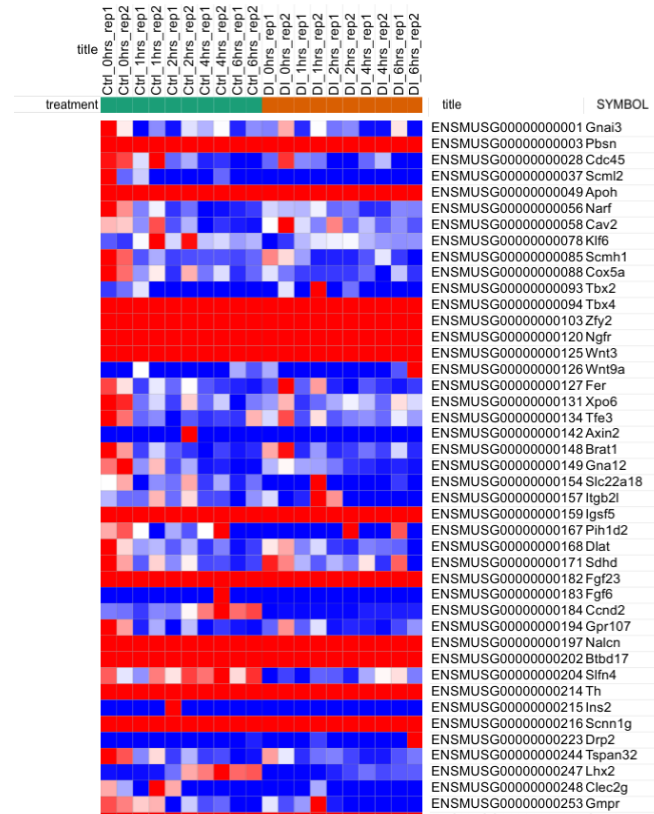
title ▼

Source column type

ENSEMBL - ENSMUSG00000030359;ENSMUSG00000020804 ▼

Result column type

SYMBOL - Pzp;Aanat ▼



Summary

- ✓ Datasets can be found:
 - in papers,
 - using GEO datasets,
 - using GEO profiles,
 - using GeneQuery
 - ...
- ✓ RNA-seq analysis can require some manual work

Alternatives for Phantastus

- ✓ GEO2R
- ✓ GENE-E (deprecated)
- ✓ Morpheus
- ✓ GENE Pattern
- ✓ shinyGEO (<https://gdancik.github.io/shinyGEO>)
- ✓ GeoDiver (<https://www.geodiver.co.uk>)
- ✓ Degust (<http://degust.erc.monash.edu/>)
- ✓ iDep (<http://bioinformatics.sdstate.edu/idep/>)
- ✓ ...

Feedback

- ✔ Bug reports, feature requests and success stories for Phantasmus are very welcome!
- ✔ alsergbox@gmail.com
- ✔ alserg@itmo.ru
- ✔ <https://github.com/ctlab/phantasmus/issues>

- ✔ Cite us as: Zenkova D, Kamenev V, Sablina R, Artyomov M, Sergushichev A. phantasmus: Visual and interactive gene expression analysis. <https://ctlab.itmo.ru/phantasmus> doi: 10.18129/B9.bioc.phantasmus