

GeneQuery: Making non-trivial statements

Going beyond making obvious statements..



Systems Biology can actually be useful tool

**Part I: getting comfortable
“jumping” between the datasets**

Going beyond making obvious statements..



So, you've got your RNA-seq....

Going beyond making obvious statements..



So, you've got your RNA-seq....

So, you've got your
differentially expressed genes...

Going beyond making obvious statements..



So, you've got your RNA-seq....

So, you've got your
differentially expressed genes...

So, you've got your
differentially regulated pathways...

Going beyond making obvious statements..



So, you've got your RNA-seq....

So, you've got your
differentially expressed genes...

So, you've got your
differentially regulated pathways...

What's next?

Single gene follow-up: GEO profiles tool

<https://www.ncbi.nlm.nih.gov/geoprofiles/>

The screenshot displays the NCBI GEO Profiles tool interface. At the top, the search bar contains 'GEO Profiles' and 'IL12b'. The main content area shows search results for 'IL12b' with 4285 items. The first result is 'IL12b - Macrophage-restricted interleukin-10 receptor deficiency effect on colonic resident macrophages'. To the right of the text is a bar chart showing expression levels across different samples, with a blue arrow pointing to a specific data point. Below this are three more results, each with its own bar chart. The right sidebar contains sections for 'Profile data', 'Profile pathways', 'Find related data', 'Search details', and 'Recent activity'. The 'Search details' section shows the search criteria: 'IL12b [All Fields]'. The 'Recent activity' section shows a search for 'IL12b (4285)'.

NCBI Resources How To martymov My NCBI Sign Out

GEO Profiles GEO Profiles IL12b Search

Create alert Advanced Help

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

Gene keyword Customize ... Search results

Organism Customize ... Items: 1 to 20 of 4285 << First < Prev Page 1 of 215 Next > Last >>

Gene ontology Customize ...

Differential expression Up/down genes

DataSet keyword Customize ...

GEO accession Customize ...

Clear all

Show additional filters

1. [IL12b - Macrophage-restricted interleukin-10 receptor deficiency effect on colonic resident macrophages](#)

Annotation: IL12b, interleukin 12b
Organism: Mus musculus
Reporter: GPL6246, 10375331 (ID_REF), GDS5668, AF128214, AF128215, BC103608, BC103609, BC103610, BC103614, M86671, chr11:44400063-44414033 (SPOT ID)
DataSet type: Expression profiling by array, transformed count, 5 samples
ID: 130497480
[GEO DataSets](#) [Gene](#) [UniGene](#) [Profile neighbors](#) [Chromosome neighbors](#) [Homologene neighbors](#)

2. [IL12B - PTPN1 depletion effect on Hodgkin lymphoma derived cell line](#)

Annotation: IL12B, interleukin 12B
Organism: Homo sapiens
Reporter: GPL570, 207901_at (ID_REF), GDS5288, 3593 (Gene ID), NM_002187
DataSet type: Expression profiling by array, transformed count, 4 samples
ID: 120821040
[GEO DataSets](#) [Gene](#) [UniGene](#) [Profile neighbors](#) [Chromosome neighbors](#) [Homologene neighbors](#)

3. [IL12B - Hypoxia effect on monocyte-derived, mature dendritic cells](#)

Annotation: IL12B, interleukin 12B
Organism: Homo sapiens
Reporter: GPL570, 207901_at (ID_REF), GDS3858, 3593 (Gene ID), NM_002187
DataSet type: Expression profiling by array, count, 6 samples
ID: 70396740
[GEO DataSets](#) [Gene](#) [UniGene](#) [Profile neighbors](#) [Chromosome neighbors](#) [Homologene neighbors](#)

4. [IL12b - Type II activated macrophage](#)

Annotation: IL12b, interleukin 12b
Organism: Mus musculus
Reporter: GPL1261, 1449497_at (ID_REF), GDS2041, 16160 (Gene ID), AF128214

Profile data Download profile data

Profile pathways Find pathways

Find related data Database: Select Find items

Search details IL12b [All Fields] Search See more...

Recent activity Turn Off Clear

IL12b (4285) GEO Profiles

Expression of specific inflammasome gene modules stratifies older individuals in PubMed

Single gene follow-up: GEO profiles tool

<https://www.ncbi.nlm.nih.gov/geoprofiles/>

Premise is that if the same gene is changing in two different datasets then there might be some common biology

The screenshot shows the GEO Profiles tool interface. The search term is "IL12b". The search results are displayed in a list format. The first result is "IL12b - Hypoxia effect on monocyte-derived, mature dendritic cells" with the following details:

- Annotation: IL12b, interleukin 12b
- Organism: Mus musculus
- Reporter: GPL570, 1034231 (ID_REF), GDS5668, AF122215, BC123698
- DataSet type: Expression profiling by array, transformed count, 4 samples
- ID: 130497480

The second result is "IL12b - Type II activated macrophage" with the following details:

- Annotation: IL12b, interleukin 12b
- Organism: Mus musculus
- Reporter: GPL1261, 1449497_at (ID_REF), GDS2041, 16160 (Gene ID), AF128214

The interface also includes a sidebar with filters for Gene symbol, Gene keyword, Organism, Differential, DataSet keyword, and GEO accession. The main content area shows a list of search results with details for each, including gene symbols, annotations, and reporter information. A blue arrow points from the red text overlay to the first search result.

This is not good enough

- Single genes can have weird behavior with outliers
- Single genes do not necessarily define the phenotype

Working with gene sets is much more reliable!!

Electrophilic properties of itaconate and derivatives regulate the $\text{I}\kappa\text{B}\zeta$ – ATF3 inflammatory axis

Monika Bambouskova¹, Laurent Gorvel¹, Vicky Lampropoulou¹, Alexey Sergushichev², Ekaterina Loginicheva¹, Kendall Johnson³, Daniel Korenfeld¹, Mary Elizabeth Mathyer⁴, Hyeryun Kim³, Li-Hao Huang¹, Dustin Duncan⁵, Howard Bregman³, Abdurrahman Keskin⁶, Andrea Santeford⁷, Rajendra S. Apte⁷, Raghav Sehgal⁸, Britney Johnson¹, Gaya K. Amarasinghe¹, Miguel P. Soares⁹, Takashi Satoh¹⁰, Shizuo Akira¹⁰, Tsonwin Hai¹¹, Cristina de Guzman Strong⁴, Karine Auclair⁵, Thomas P. Roddy³, Scott A. Biller³, Marko Jovanovic⁶, Eynav Klechevsky¹, Kelly M. Stewart³, Gwendalyn J. Randolph¹ & Maxim N. Artyomov^{1*}

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Data availability. The raw and processed RNA-seq data have been deposited in the Gene Expression Omnibus with accession numbers GSE102190 and GSE110749. The original mass spectra may be downloaded from MassIVE (<http://massive.ucsd.edu>) using the identifier MSV000082101. Source data for the graphical representations found in all figures and Extended Data figures are provided. Source data for western blots (uncropped and unprocessed scans with size marker indications) are presented in Supplementary Fig. 1. All other data that support the findings of this study are available from the corresponding author upon reasonable request.

acc.cgi?acc=GSE110749

Submission date Feb 16, 2018
Last update date Feb 01, 2019
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

[1586-018-0052-z](#)

LET

Electroregulation of regulation

Monika Bambouskova¹,
Kendall Johnson³,
Howard Bregman³,
Gaya K. Amarasinghe⁴,
Karine Auclair⁵, Tamas
Gwendalyn J. Randolph¹

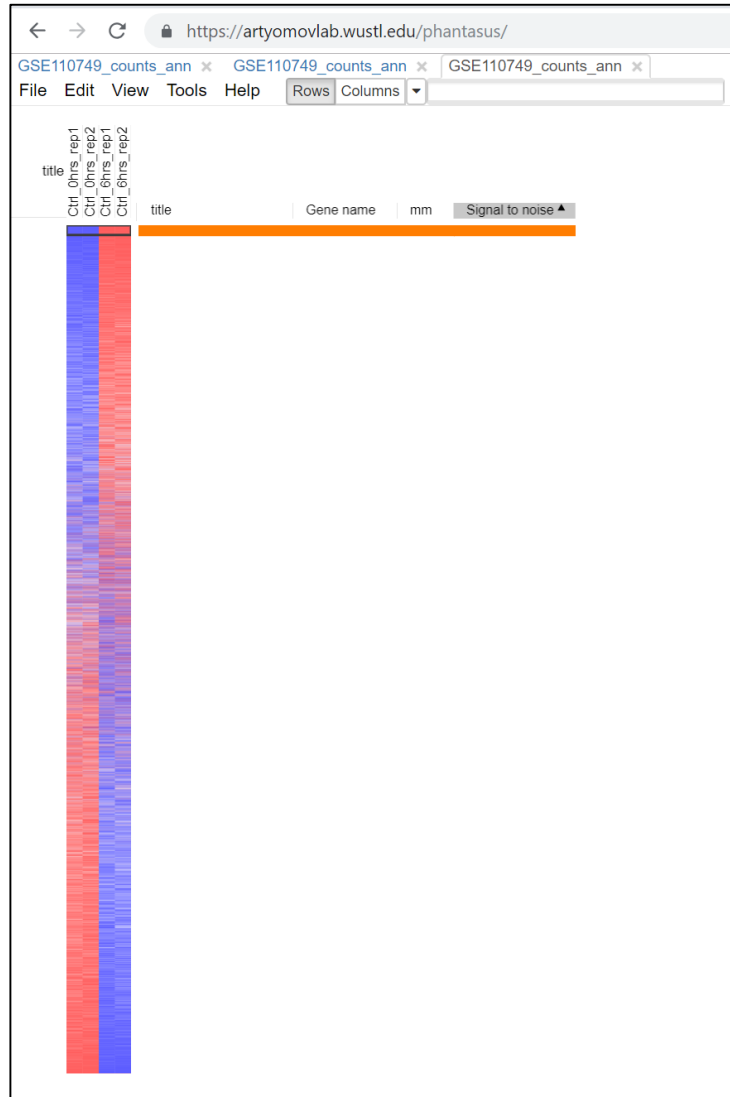
atives

n¹,
ng⁴,

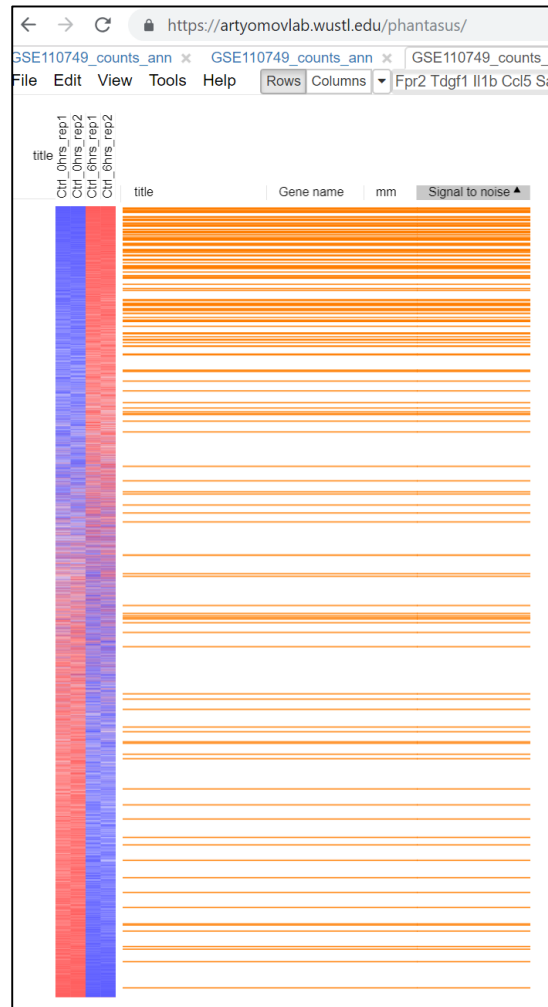
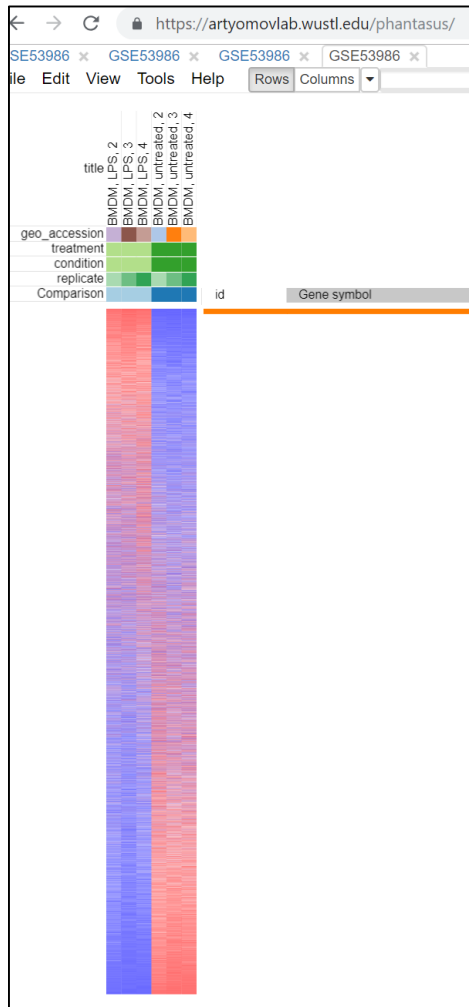
You know what to do!

Counts are
available in folder

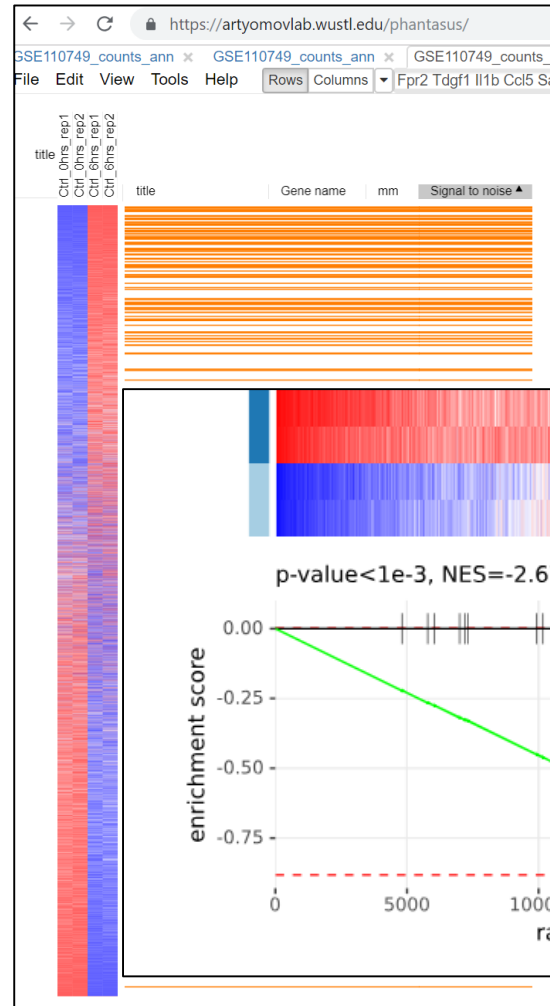
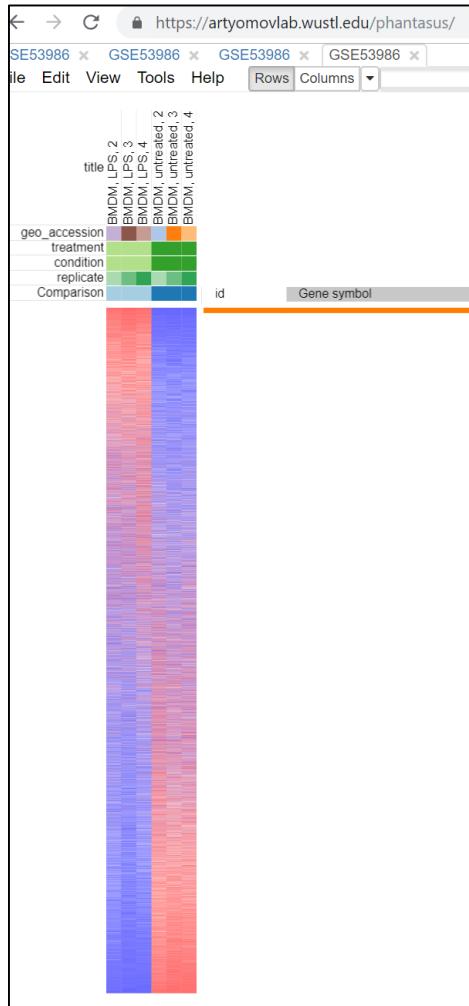
Log2->
qnorm->
Filter top 10k->
DE of 0h vs 6h LPS



The two datasets are connected!



The two datasets are connected!



But you kind of knew the answer!!!

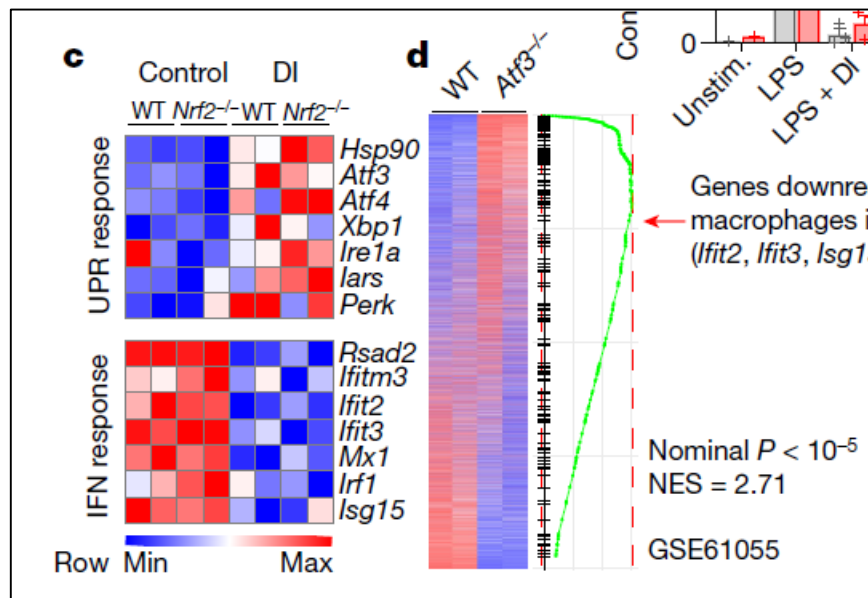
Do people do it in real science? – YES!!

LETTER

<https://doi.org/10.1038/s41586-018-0052-z>

Electrophilic properties of itaconate and derivatives regulate the $\text{I}\kappa\text{B}\zeta$ – ATF3 inflammatory axis

Monika Bambouskova¹, Laurent Gorvel¹, Vicky Lampropoulou¹, Alexey Sergushichev², Ekaterina Loginicheva¹, Kendall Johnson³, Daniel Korenfeld¹, Mary Elizabeth Mather⁴, Hyeryun Kim³, Li-Hao Huang¹, Dustin Duncan⁵, Howard Bregman³, Abdurrahman Keskin⁶, Andrea Santeford⁷, Rajendra S. Apte⁷, Raghav Sehgal⁸, Britney Johnson¹, Gaya K. Amarasinghe¹, Miguel P. Soares⁹, Takashi Satoh¹⁰, Shizuo Akira¹⁰, Tsonwin Hai¹¹, Cristina de Guzman Strong⁴, Karine Auclair⁵, Thomas P. Roddy³, Scott A. Biller³, Marko Jovanovic⁶, Eynav Klechevsky¹, Kelly M. Stewart³, Gwendalyn J. Randolph¹ & Maxim N. Artyomov^{1*}



What if we make database?

**Try to do diff expression between
all conditions in all published datasets...**

MsigDB c7

GSEA
Gene Set Enrichment Analysis

GSEA Home Downloads **Molecular Signatures Database** Documentation Contact

MSigDB Home
About Collections
Browse Gene Sets
Search Gene Sets
Investigate Gene Sets
View Gene Families
Help

Investigate Gene Sets

Gain further insight into the biology behind a gene set by using the following tools:

- ▶ **compute overlaps** with other gene sets in MSigDB ([more...](#))
- ▶ **display the gene set expression profile** based on a selected compendium of expression data ([more...](#))
- ▶ **categorize** members of the gene set by gene families ([more...](#))

Gene Identifiers

- Irgm1
- Nrp2
- Xkr8
- Man2a1
- Rela
- N4bp1
- Sp100
- Itga1
- Helz2
- Col11a1
- RP23-148L8.3
- Aida
- Ptges
- Sp110
- Itga5
- Gnb1
- Sh3bp5
- Stat1
- Id2
- Gstm1
- Kpna4
- Mnda
- Psmc2
- Ak4
- Clic4
- Gng12
- Ehd1

Compute Overlaps

- H: hallmark gene sets [?](#)
- C1: positional gene sets [?](#)
- C2: curated gene sets [?](#)
 - CGP: chemical and genetic perturbations [?](#)
 - CP: Canonical pathways [?](#)
 - CP: BIOCARTA: BioCarta gene sets [?](#)
 - CP: KEGG: KEGG gene sets [?](#)
 - CP: REACTOME: Reactome gene sets [?](#)
 - C3: motif gene sets [?](#)
 - MIR: microRNA targets [?](#)
 - TFT: transcription factor targets [?](#)
 - C4: computational gene sets [?](#)
 - CGN: cancer gene neighborhoods [?](#)
 - CM: cancer modules [?](#)
 - C5: GO gene sets [?](#)
 - BP: GO biological process [?](#)
 - CC: GO cellular component [?](#)
 - MF: GO molecular function [?](#)
 - C6: oncogenic signatures [?](#)
 - C7: immunologic signatures [?](#)

show **top 50** genesets

with FDR q-value below

[compute overlaps](#)

Compendia expression profiles

- Human tissue compendium (Novartis)
- NCI-60 cell lines (National Cancer Institute)



[display expression profile](#)

Gene families

[show gene families](#)

MSigDB.C7 – attempt to make such database

secure | software.broadinstitute.org/gsea/msigdb/genesets.jsp?collection=C7

- ▶ C6 (oncogenic signatures, 189 gene sets) 
- ▶ **C7** (immunologic signatures, 4872 gene sets) 

Click on a gene set name to view its gene set page.

| | | |
|-------------------------------------|-------------------------------------|-------------------------------------|
| GOLDRATH_EFF_VS_MEMORY_CD8_TCELL_DN | GSE21546_SAP1A_KO_VS_SAP1A_KO_AND_E | GSE36476_YOUNG_VS_OLD_DONOR_MEMORY_ |
| GOLDRATH_EFF_VS_MEMORY_CD8_TCELL_UP | LK1_KO_DP_THYMOCYTES_UP | CD4_TCELL_72H_TSST_ACT_UP |
| GOLDRATH_NAIVE_VS_EFF_CD8_TCELL_DN | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_DP | GSE36476_YOUNG_VS_OLD_DONOR_MEMORY_ |
| GOLDRATH_NAIVE_VS_EFF_CD8_TCELL_UP | _THYMOCYTES_DN | CD4_TCELL_DN |
| GOLDRATH_NAIVE_VS_MEMORY_CD8_TCELL_ | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_DP | GSE36476_YOUNG_VS_OLD_DONOR_MEMORY_ |
| DN | _THYMOCYTES_UP | CD4_TCELL_UP |
| GOLDRATH_NAIVE_VS_MEMORY_CD8_TCELL_ | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_EL | GSE36527_CD62L_HIGH_CD69_NEG_VS_CD6 |
| UP | K1_KO_DP_THYMOCYTES_DN | 2L_LOW_CD69_POS_TREG_KLRG1_NEG_DN |
| GSE10094_LCMV_VS_LISTERIA_IND_EFF_C | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_EL | GSE36527_CD62L_HIGH_CD69_NEG_VS_CD6 |
| D4_TCELL_DN | K1_KO_DP_THYMOCYTES_UP | 2L_LOW_CD69_POS_TREG_KLRG1_NEG_UP |
| GSE10094_LCMV_VS_LISTERIA_IND_EFF_C | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_SA | GSE36527_CD62L_HIGH_VS_CD62L_LOW_TR |
| D4_TCELL_UP | P1A_KO_AND_ELK1_KO_DP_THYMOCYTES_DN | EG_CD69_NEG_KLRG1_NEG_DN |
| GSE10147_IL3_AND_HIVP17_VS_IL3_AND_ | | GSE36527_CD62L_HIGH_VS_CD62L_LOW_TR |
| CPG_STIM_PDC_DN | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_SA | EG_CD69_NEG_KLRG1_NEG_UP |
| GSE10147_IL3_AND_HIVP17_VS_IL3_AND_ | P1A_KO_AND_ELK1_KO_DP_THYMOCYTES_UP | GSE36527_CD69_NEG_VS_POS_TREG_CD62L |
| CPG_STIM_PDC_UP | | _LOS_KLRG1_NEG_DN |
| GSE10147_IL3_VS_IL3_AND_CPG_STIM_PD | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_SA | GSE36527_CD69_NEG_VS_POS_TREG_CD62L |
| C_DN | P1A_KO_DP_THYMOCYTES_DN | _LOS_KLRG1_NEG_UP |
| GSE10147_IL3_VS_IL3_AND_CPG_STIM_PD | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_SA | GSE36826_NORMAL_VS_STAPH_AUREUS_INF |
| C_UP | P1A_KO_DP_THYMOCYTES_UP | _IL1R_KO_SKIN_DN |
| GSE10147_IL3_VS_IL3_AND_HIVP17_STIM | GSE21546_WT_VS_ELK1_KO_ANTI_CD3_STI | GSE36826_NORMAL_VS_STAPH_AUREUS_INF |
| _PDC_DN | M_DP_THYMOCYTES_DN | _IL1R_KO_SKIN_UP |
| GSE10147_IL3_VS_IL3_AND_HIVP17_STIM | GSE21546_WT_VS_ELK1_KO_ANTI_CD3_STI | GSE36826_NORMAL_VS_STAPH_AUREUS_INF |
| _PDC_UP | M_DP_THYMOCYTES_UP | _SKIN_DN |
| GSE10211_UV_INACT_SENDAI_VS_LIVE_SE | GSE21546_WT_VS_ELK1_KO_DP_THYMOCYTE | GSE36826_NORMAL_VS_STAPH_AUREUS_INF |
| NDAI_VIRUS_TRACHEAL_EPITHELIAL_CELL | S_DN | _SKIN_UP |
| S_DN | GSE21546_WT_VS_ELK1_KO_DP_THYMOCYTE | GSE36826_WT_VS_IL1R_KO_SKIN_DN |
| GSE10211_UV_INACT_SENDAI_VS_LIVE_SE | S_UP | GSE36826_WT_VS_IL1R_KO_SKIN_STAPH_A |
| NDAI_VIRUS_TRACHEAL_EPITHELIAL_CELL | | UREUS_INF_DN |
| S_UP | GSE21546_WT_VS_SAP1A_KO_AND_ELK1_KO | GSE36826_WT_VS_IL1R_KO_SKIN_STAPH_A |
| GSE10239_KLRG1INT_VS_KLRG1HIGH_EFF_ | _ANTI_CD3_STIM_DP_THYMOCYTES_DN | UREUS_INF_UP |
| CD8_TCELL_DN | GSE21546_WT_VS_SAP1A_KO_AND_ELK1_KO | GSE36826_WT_VS_IL1R_KO_SKIN_UP |
| GSE10239_KLRG1INT_VS_KLRG1HIGH_EFF_ | _ANTI_CD3_STIM_DP_THYMOCYTES_UP | GSE36888_STAT5_AB_KNOCKIN_VS_WT_TCE |
| CD8_TCELL_UP | GSE21546_WT_VS_SAP1A_KO_AND_ELK1_KO | LL_IL2_TREATED_17H_DN |
| GSE10239_MEMORY_VS_DAY4.5_EFF_CD8_T | _DP_THYMOCYTES_DN | GSE36888_STAT5_AB_KNOCKIN_VS_WT_TCE |
| CELL_DN | GSE21546_WT_VS_SAP1A_KO_AND_ELK1_KO | LL_IL2_TREATED_17H_UP |
| GSE10239_MEMORY_VS_DAY4.5_EFF_CD8_T | _DP_THYMOCYTES_UP | GSE36888_STAT5_AB_KNOCKIN_VS_WT_TCE |
| CELL_UP | GSE21546_WT_VS_SAP1A_KO_ANTI_CD3_ST | LL_IL2_TREATED_2H_DN |
| GSE10239_MEMORY_VS_KLRG1HIGH_EFF_CD | IM_DP_THYMOCYTES_DN | GSE36888_STAT5_AB_KNOCKIN_VS_WT_TCE |
| | GSE21546_WT_VS_SAP1A_KO_ANTI_CD3_ST | |

Macrophage activation is found!

software.broadinstitute.org/gsea/msigdb/compute_overlaps.jsp

GSEA
Gene Set Enrichment Analysis

GSEA Home Downloads **Molecular Signatures Database** Documentation Contact

MSigDB Home
About Collections
Browse Gene Sets
Search Gene Sets
Investigate Gene Sets
View Gene Families
Help

Compute Overlaps for Selected Genes

Converted 154 submitted identifiers into 127 entrez genes. [click here for details.](#)


| Collections | # Overlaps Shown | # Gene Sets in Collections | # Genes in Comparison (n) | # Genes in Universe (N) |
|-------------|------------------|----------------------------|---------------------------|-------------------------|
| C7 | 50 | 4872 | 127 | 45956 |

Click the gene set name to see the gene set page. Click the number of genes [in brackets] to download the list of genes.

Color bar shading from light green to black, where lighter colors indicate more significant FDR q-values (< 0.05) and black indicates less significant FDR q-values (>= 0.05).

Save to: [Excel](#) | [GenomeSpace](#)

| Gene Set Name [# Genes (K)] | Description | # Genes in Overlap (k) | k/K | p-value | FDR q-value |
|--------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|-----|-----------------------|-----------------------|
| GSE18281_SUBCAPSULAR_VS_CENTRAL_CORTICAL_REGION_OF_THYMUS_DN [200] | Genes down-regulated in thymus cortical regions: subcapsular versus central cortical. | 33 | | 1.69 e ⁻⁴⁵ | 8.21 e ⁻⁴⁶ |
| GSE9316_IL6_KO_VS_IFNG_KO_INVIVO_EXPANDED_CD4_TCELL_DN [200] | Genes down-regulated in CD4 [GeneID=920] T cells with in vivo expansion: IL6 [GeneID=3569] versus IFNG [GeneID=3458]. | 32 | | 1.59 e ⁻⁴⁷ | 3.88 e ⁻⁴⁴ |
| GSE7348_UNSTIM_VS_LPS_STIM_MACROPHAGE_GE_DN [167] | Genes down-regulated in macrophages: untreated versus LPS. | 29 | | 3.69 e ⁻⁴⁴ | 6 e ⁻⁴¹ |
| GSE32986_CURDLAN_HIGHDOSE_VS_GMCSF_AND_CURDLAN_HIGHDOSE_STIM_DC_DN [200] | Genes down-regulated in bone marrow-derived dendritic cells: high dose of 1,3-beta-D-oligoglucan [PubChem=11375554] versus CSF2 [GeneID=1437] and high dose of 1,3-beta-D-oligoglucan [PubChem=11375554]. | 30 | | 1.23 e ⁻⁴³ | 1.5 e ⁻⁴⁰ |
| GSE21360_SECONDARY_VS_QUATERNARY_MEMORY_CD8_TCELL_UP [177] | Genes up-regulated in memory CD8 T cells: 2' versus 4'. | 29 | | 2.28 e ⁻⁴³ | 2.22 e ⁻⁴⁰ |
| GSE14000_UNSTIM_VS_4H_LPS_DC_TRANSLATED_RNA_DN [200] | Genes down-regulated in comparison of polysome bound (translated) mRNA before and 4 h after LPS (TLR4 agonist) stimulation. | 28 | | 7.84 e ⁻⁴⁰ | 4.78 e ⁻³⁷ |
| GSE19888_ADENOSINE_A3R_INH_PRETREAT_AND_ACT_BY_A3R_VS_TCELL_MEMBRANE_ACT | Genes up-regulated in UMC1 (Genes) | 28 | | 7.84 e ⁻⁴⁰ | 4.78 e ⁻³⁷ |



You know what to do!

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7348>



| | |
|-------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Status | Public on Apr 22, 2007 |
| Title | Gene Expression in Naive and Tolerant Macrophages stimulated with LPS |
| Organism | Mus musculus |
| Experiment type | Expression profiling by array |
| Summary | <p>The inflammatory response initiated by microbial products signaling through Toll-like receptors (TLRs) of the innate immune system is essential for host defense against infection. Because inflammation can be harmful to host tissues, the innate response is highly regulated. Negative regulation of TLR4, the receptor for bacterial lipopolysaccharide (LPS), results in LPS tolerance, defined as hyporesponsiveness to repeated stimulation with LPS. LPS tolerance is thought to protect the host from excessive inflammation by turning off TLR4 signal, which then shuts down TLR-induced genes. However, TLR signaling induces hundreds of genes with very different functions. We reasoned that genes with different functions should have different requirements for regulation. Specifically, genes encoding proinflammatory mediators should be transiently inactivated to limit tissue damage, while genes encoding antimicrobial effectors, which directly target pathogens, should remain inducible in tolerant cells to protect the host from infection. Using an in vitro system of LPS tolerance in macrophages, here we show that TLR-induced genes may indeed be divided into two distinct categories based on their functions and regulatory requirements. Further, we show these distinct groups are regulated by gene-specific, and not signal-specific mechanisms.</p> <p>Keywords: Treatment Comparison</p> |
| Overall design | We examined gene expression using affymetrix genechips in 3 groups of murine bone-marrow derived macrophages: Naive (untreated), Naive stimulated with LPS, and Tolerant stimulated with LPS. Two biological replicates were performed for each group. |
| Contributor(s) | Foster SL , Hargreaves DC , Medzhitov R |
| Citation(s) | Foster SL, Hargreaves DC, Medzhitov R. Gene-specific control of inflammation by TLR-induced chromatin modifications. <i>Nature</i> 2007 Jun 21;447(7147):972-8. PMID: 17538624 |
| Submission date | Mar 22, 2007 |
| Last update date | Jan 08, 2019 |
| Contact name | Simmie L Foster |
| E-mail | simmie.foster@yale.edu |
| Phone | (203)785-6090 |
| Fax | (203)785-4461 |
| Organization name | Yale University School of Medicine |
| Department | Immunobiology |
| Lab | Medzhitov |
| Street address | 300 Cedar St |
| City | New Haven |
| State/province | CT |

ARTICLES

Gene-specific control of inflammation by TLR-induced chromatin modifications

Simmie L. Foster^{1*}, Diana C. Hargreaves^{1*} & Ruslan Medzhitov¹

Toll-like receptors (TLRs) induce a multi-component inflammatory response that must be tightly regulated to avoid tissue damage. Most known regulatory mechanisms target TLR signalling pathways and thus broadly inhibit multiple aspects of the inflammatory response. Given the functional diversity of TLR-induced genes, we proposed that additional, gene-specific regulatory mechanisms exist to allow individual aspects of the TLR-induced response to be differentially regulated. Using an *in vitro* system of lipopolysaccharide tolerance in murine macrophages, we show that TLR-induced genes fall into two categories on the basis of their functions and regulatory requirements. We demonstrate that representatives from the two classes acquire distinct patterns of TLR-induced chromatin modifications. These gene-specific chromatin modifications are associated with transient silencing of one class of genes, which includes pro-inflammatory mediators, and priming of the second class, which includes antimicrobial effectors. These findings illustrate an adaptive response in macrophages and reveal component-specific regulation of inflammation.

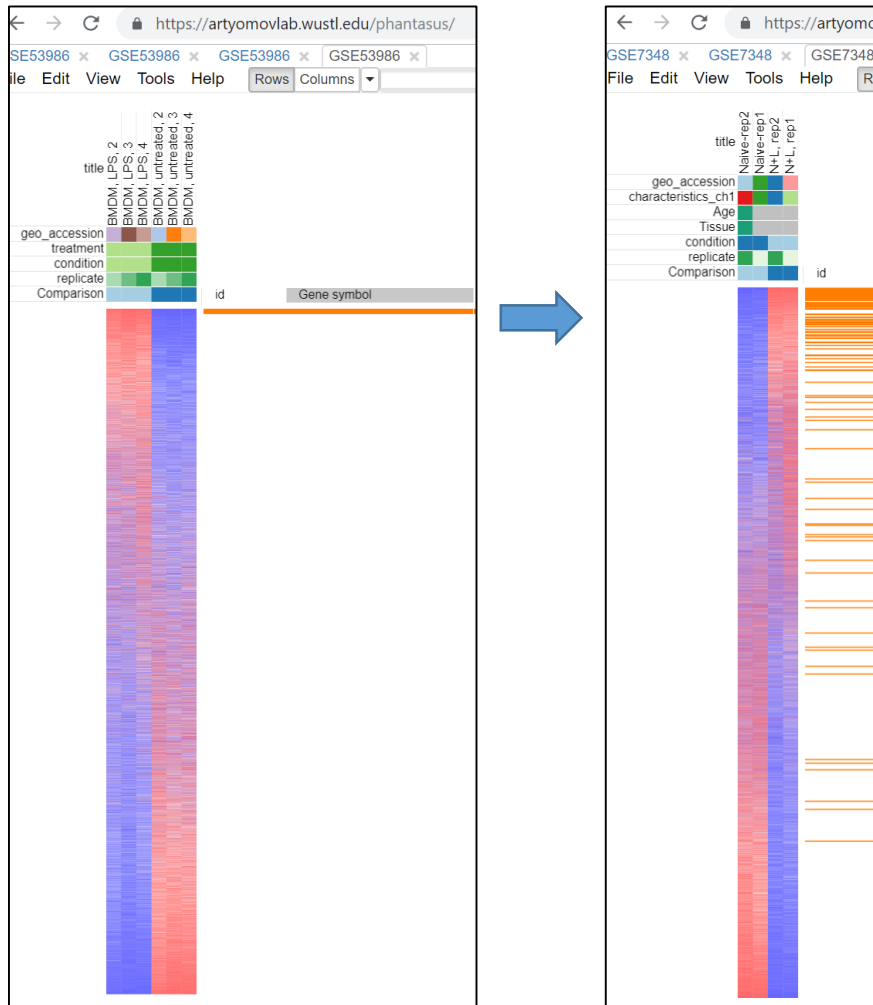
← → ↻ <https://artyomovlab.wustl.edu/phantasus/>

GSE7348 x GSE7348 x GSE7348 x

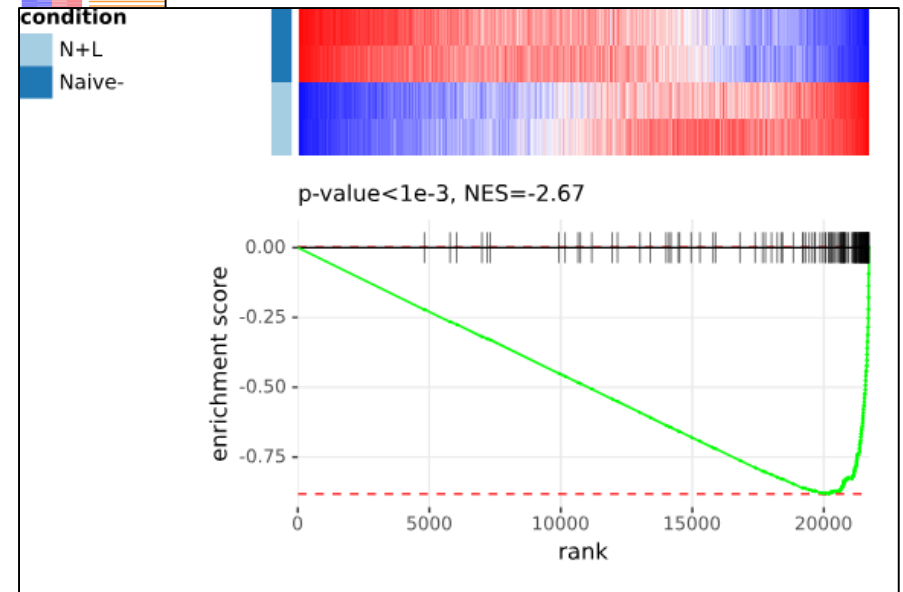
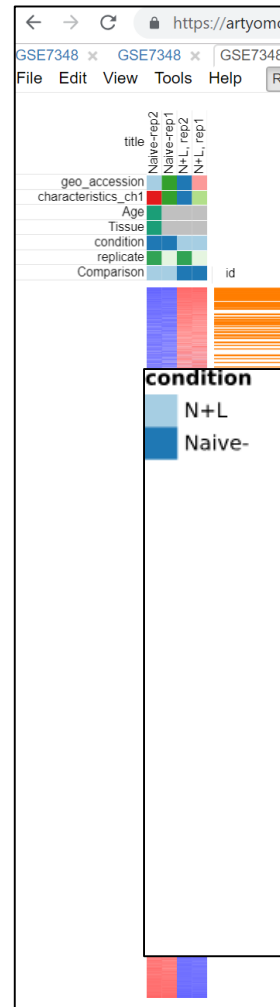
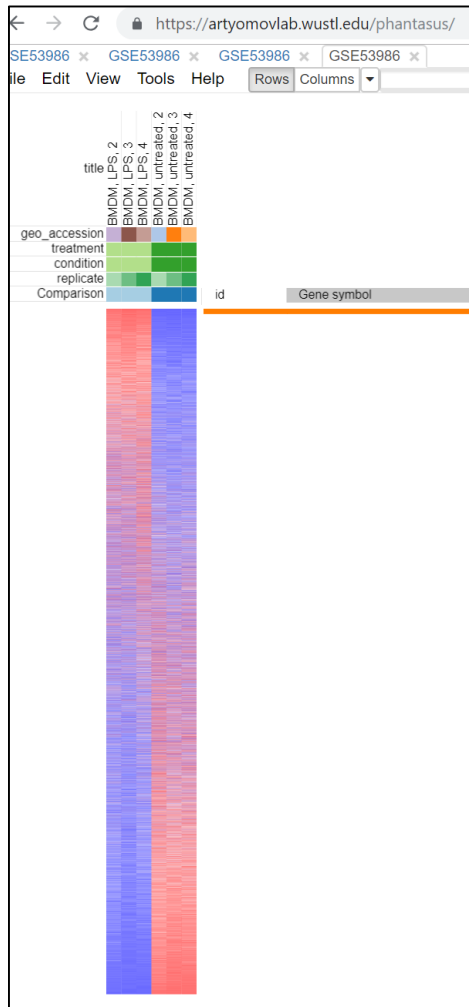
File Edit View Tools Help Rows Columns 21,723 rows by 4 columns

| title | Naive-rep2 | Naive-rep1 | N+L_rep2 | N+L_rep1 |
|---------------------|------------|------------|----------|----------|
| geo_accession | | | | |
| characteristics_ch1 | | | | |
| Age | | | | |
| Tissue | | | | |
| condition | | | | |
| replicate | | | | |
| Comparison | | | | |
| id | | | | |
| Gene symbol | | | | |
| Gene ID | | | | |
| Signal to noise ▲ | | | | |

Differentially expressed genes behave very similar between two datasets!





Differentially expressed genes behave very similar between two datasets!



MSigDB.C7 – attempt to make such database

secure | software.broadinstitute.org/gsea/msigdb/genesets.jsp?collection=C7

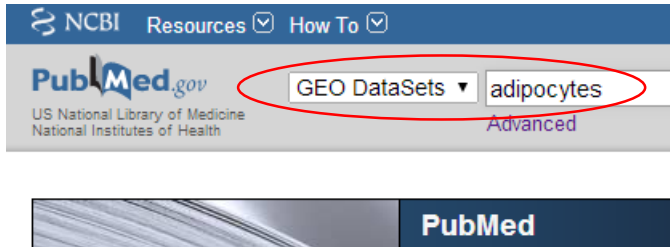
- ▶ C6 (oncogenic signatures, 189 gene sets) 
- ▶ **C7 (immunologic signatures, 4872 gene sets) **

Click on a gene set name to view its gene set page.

| | | |
|-------------------------------------|-------------------------------------|-------------------------------------|
| GOLDRATH_EFF_VS_MEMORY_CD8_TCELL_DN | GSE21546_SAP1A_KO_VS_SAP1A_KO_AND_E | GSE36476_YOUNG_VS_OLD_DONOR_MEMORY_ |
| GOLDRATH_EFF_VS_MEMORY_CD8_TCELL_UP | LK1_KO_DP_THYMOCYTES_UP | CD4_TCELL_72H_TSST_ACT_UP |
| GOLDRATH_NAIVE_VS_EFF_CD8_TCELL_DN | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_DP | GSE36476_YOUNG_VS_OLD_DONOR_MEMORY_ |
| GOLDRATH_NAIVE_VS_EFF_CD8_TCELL_UP | _THYMOCYTES_DN | CD4_TCELL_DN |
| GOLDRATH_NAIVE_VS_MEMORY_CD8_TCELL_ | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_DP | GSE36476_YOUNG_VS_OLD_DONOR_MEMORY_ |
| DN | _THYMOCYTES_UP | CD4_TCELL_UP |
| GOLDRATH_NAIVE_VS_MEMORY_CD8_TCELL_ | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_EL | GSE36527_CD62L_HIGH_CD69_NEG_VS_CD6 |
| UP | K1_KO_DP_THYMOCYTES_DN | 2L_LOW_CD69_POS_TREG_KLRG1_NEG_DN |
| GSE10094_LCMV_VS_LISTERIA_IND_EFF_C | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_EL | GSE36527_CD62L_HIGH_CD69_NEG_VS_CD6 |
| ELL_DN | K1_KO_DP_THYMOCYTES_UP | GSE36527_CD62L_HIGH_VS_CD62L_LOW_TR |
| GSE10147_IL3_AND_HI | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_SA | EG_CD69_NEG_KLRG1_NEG_UP |
| CPG_STIM_PDC_DN | P1A_KO_AND_ELK1_KO_DP_THYMOCYTES_DN | EG_CD69_NEG_KLRG1_NEG_DN |
| GSE10147_IL3_AND_HI | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_SA | GSE36527_CD62L_HIGH_VS_CD62L_LOW_TR |
| CPG_STIM_PDC_UP | P1A_KO_AND_ELK1_KO_DP_THYMOCYTES_UP | EG_CD69_NEG_KLRG1_NEG_UP |
| GSE10147_IL3_VS_IL3_AND_CPG_STIM_PD | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_SA | GSE36527_CD69_NEG_VS_POS_TREG_CD62L |
| C_DN | P1A_KO_DP_THYMOCYTES_DN | _LOS_KLRG1_NEG_DN |
| GSE10147_IL3_VS_IL3_AND_CPG_STIM_PD | GSE21546_UNSTIM_VS_ANTI_CD3_STIM_SA | GSE36527_CD69_NEG_VS_POS_TREG_CD62L |
| C_UP | P1A_KO_DP_THYMOCYTES_UP | _LOS_KLRG1_NEG_UP |
| GSE10147_IL3_VS_IL3_AND_HI | GSE21546_WT_VS_ELK1_KO_ANTI_CD3_STI | GSE36826_NORMAL_VS_STAPH_AUREUS_INF |
| _PDC_DN | M_DP_THYMOCYTES_DN | _IL1R_KO_SKIN_DN |
| GSE10147_IL3_VS_IL3_AND_HI | GSE21546_WT_VS_ELK1_KO_ANTI_CD3_STI | GSE36826_NORMAL_VS_STAPH_AUREUS_INF |
| _PDC_UP | M_DP_THYMOCYTES_UP | _SKIN_DN |
| GSE10211_UV_INACT_SENDAI_VS_LIVE_SE | GSE21546_WT_VS_ELK1_KO_DP_THYMOCYTE | GSE36826_NORMAL_VS_STAPH_AUREUS_INF |
| NDAI_VIRUS_TRACHEAL_EPITHELIAL_CELL | S_DN | _SKIN_UP |
| S_DN | GSE21546_WT_VS_ELK1_KO_DP_THYMOCYTE | GSE36826_WT_VS_IL1R_KO_SKIN_DN |
| GSE10211_UV_INACT_SENDAI_VS_LIVE_SE | S_UP | GSE36826_WT_VS_IL1R_KO_SKIN_STAPH_A |
| NDAI_VIRUS_TRACHEAL_EPITHELIAL_CELL | GSE21546_WT_VS_SAP1A_KO_AND_ELK1_KO | UREUS_INF_DN |
| S_UP | _ANTI_CD3_STIM_DP_THYMOCYTES_DN | GSE36826_WT_VS_IL1R_KO_SKIN_STAPH_A |
| GSE10239_KLRG1INT_VS_KLRG1HIGH_EFF_ | GSE21546_WT_VS_SAP1A_KO_AND_ELK1_KO | UREUS_INF_UP |
| CD8_TCELL_DN | _ANTI_CD3_STIM_DP_THYMOCYTES_UP | GSE36826_WT_VS_IL1R_KO_SKIN_UP |
| GSE10239_KLRG1INT_VS_KLRG1HIGH_EFF_ | GSE21546_WT_VS_SAP1A_KO_AND_ELK1_KO | GSE36888_STAT5_AB_KNOCKIN_VS_WT_TCE |
| CD8_TCELL_UP | _DP_THYMOCYTES_DN | LL_IL2_TREATED_17H_DN |
| GSE10239_MEMORY_VS_DAY4.5_EFF_CD8_T | GSE21546_WT_VS_SAP1A_KO_AND_ELK1_KO | GSE36888_STAT5_AB_KNOCKIN_VS_WT_TCE |
| CELL_DN | _DP_THYMOCYTES_UP | LL_IL2_TREATED_17H_UP |
| GSE10239_MEMORY_VS_DAY4.5_EFF_CD8_T | GSE21546_WT_VS_SAP1A_KO_ANTI_CD3_ST | GSE36888_STAT5_AB_KNOCKIN_VS_WT_TCE |
| CELL_UP | IM_DP_THYMOCYTES_DN | LL_IL2_TREATED_2H_DN |
| GSE10239_MEMORY_VS_KLRG1HIGH_EFF_CD | GSE21546_WT_VS_SAP1A_KO_ANTI_CD3_ST | GSE36888_STAT5_AB_KNOCKIN_VS_WT_TCE |

4872 gene sets from only ~100 datasets!!

GEO contains A LOT of datasets



A screenshot of the NCBI GEO DataSets search results page for 'adipocytes'. The search bar at the top shows 'GEO DataSets' and 'adipocytes'. The results section shows 'Results: 1 to 20 of 3095'. The first result is 'Subcutaneous and visceral adipose tissue from lean or obese children'. The 'Entry type' sidebar shows 'DataSets (39)'. The 'Organism' sidebar shows 'Homo sapiens'. The 'Study type' sidebar shows 'Expression profiling by array, transformed count, 2 disease state, 2 gender, 2 tissue sets'. The 'Top Organisms' sidebar shows 'Mus musculus (1546)', 'Homo sapiens (1471)', 'Sus scrofa (51)', 'Rattus norvegicus (26)', and 'Bos taurus (7)'. The 'Find related data' sidebar shows a 'Database' dropdown menu set to 'Select' and a 'Find items' button.

However, they are poorly annotated

Samples (24)
[Less...](#)

| | |
|-----------|--------------------------------------|
| GSM795602 | Epididymal adipose GPR120KO HFD #427 |
| GSM795603 | Liver GPR120KO HFD #427 |
| GSM795604 | Epididymal adipose GPR120KO HFD #428 |
| GSM795605 | Liver GPR120KO HFD #428 |
| GSM795606 | Epididymal adipose GPR120KO HFD #429 |
| GSM795607 | Liver GPR120KO HFD #429 |
| GSM795608 | Epididymal adipose WT HFD #546 |
| GSM795609 | Liver WT HFD #546 |
| GSM795610 | Epididymal adipose WT HFD #547 |
| GSM795611 | Liver WT HFD #547 |
| GSM795612 | Epididymal adipose WT HFD #560 |
| GSM795613 | Liver WT HFD #560 |
| GSM795614 | Liver WT ND #686 |
| GSM795615 | Epididymal adipose WT ND #687 |
| GSM795616 | Liver WT ND #687 |
| GSM795617 | Epididymal adipose WT ND #690 |
| GSM795618 | Liver WT ND #690 |
| GSM795619 | Epididymal adipose GPR120KO ND #718 |
| GSM795620 | Liver GPR120KO ND #718 |
| GSM795621 | Epididymal adipose GPR120KO ND #719 |
| GSM795622 | Liver GPR120KO ND #719 |
| GSM795623 | Epididymal adipose GPR120KO ND #720 |
| GSM795624 | Liver GPR120KO ND #720 |
| GSM795625 | Epididymal adipose WT ND #686 |

However, they are poorly annotated

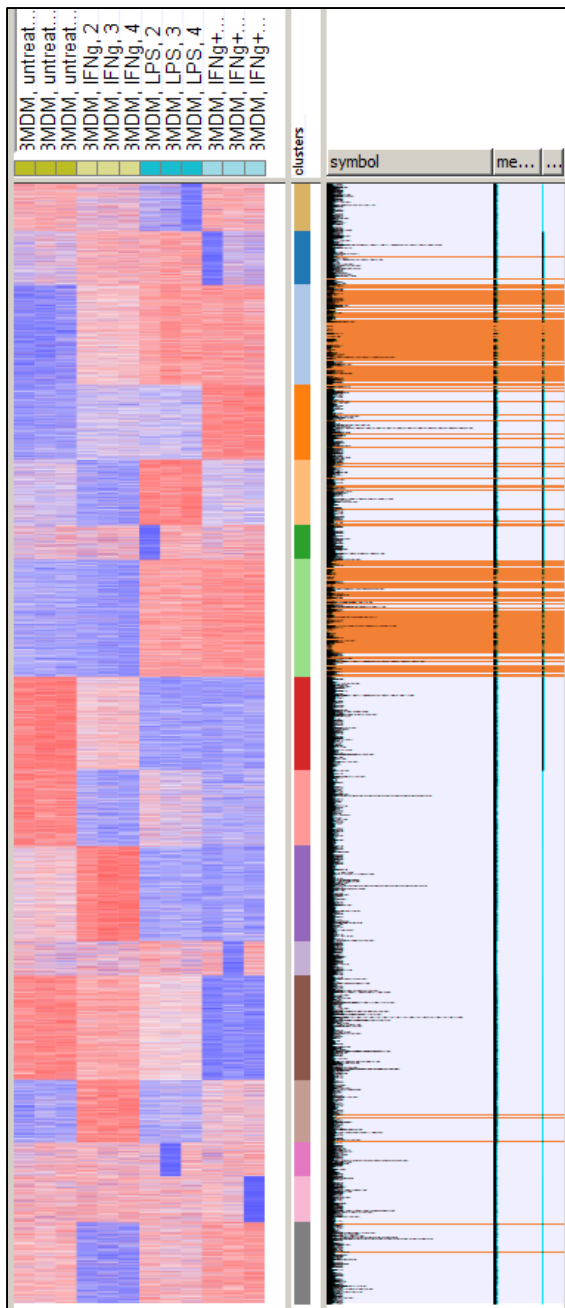
It is virtually impossible
to parse such entry
without manual curation

Samples (24)
[Less...](#)

| | |
|-----------|--------------------------------------|
| GSM795602 | Epididymal adipose GPR120KO HFD #427 |
| GSM795603 | Liver GPR120KO HFD #427 |
| GSM795604 | Epididymal adipose GPR120KO HFD #428 |
| GSM795605 | Liver GPR120KO HFD #428 |
| GSM795606 | Epididymal adipose GPR120KO HFD #429 |
| GSM795607 | Liver GPR120KO HFD #429 |
| GSM795608 | Epididymal adipose WT HFD #546 |
| GSM795609 | Liver WT HFD #546 |
| GSM795610 | Epididymal adipose WT HFD #547 |
| GSM795611 | Liver WT HFD #547 |
| GSM795612 | Epididymal adipose WT HFD #560 |
| GSM795613 | Liver WT HFD #560 |
| GSM795614 | Liver WT ND #686 |
| GSM795615 | Epididymal adipose WT ND #687 |
| GSM795616 | Liver WT ND #687 |
| GSM795617 | Epididymal adipose WT ND #690 |
| GSM795618 | Liver WT ND #690 |
| GSM795619 | Epididymal adipose GPR120KO ND #718 |
| GSM795620 | Liver GPR120KO ND #718 |
| GSM795621 | Epididymal adipose GPR120KO ND #719 |
| GSM795622 | Liver GPR120KO ND #719 |
| GSM795623 | Epididymal adipose GPR120KO ND #720 |
| GSM795624 | Liver GPR120KO ND #720 |
| GSM795625 | Epididymal adipose WT ND #686 |

**Part II: how to get somewhere
you haven't been before**

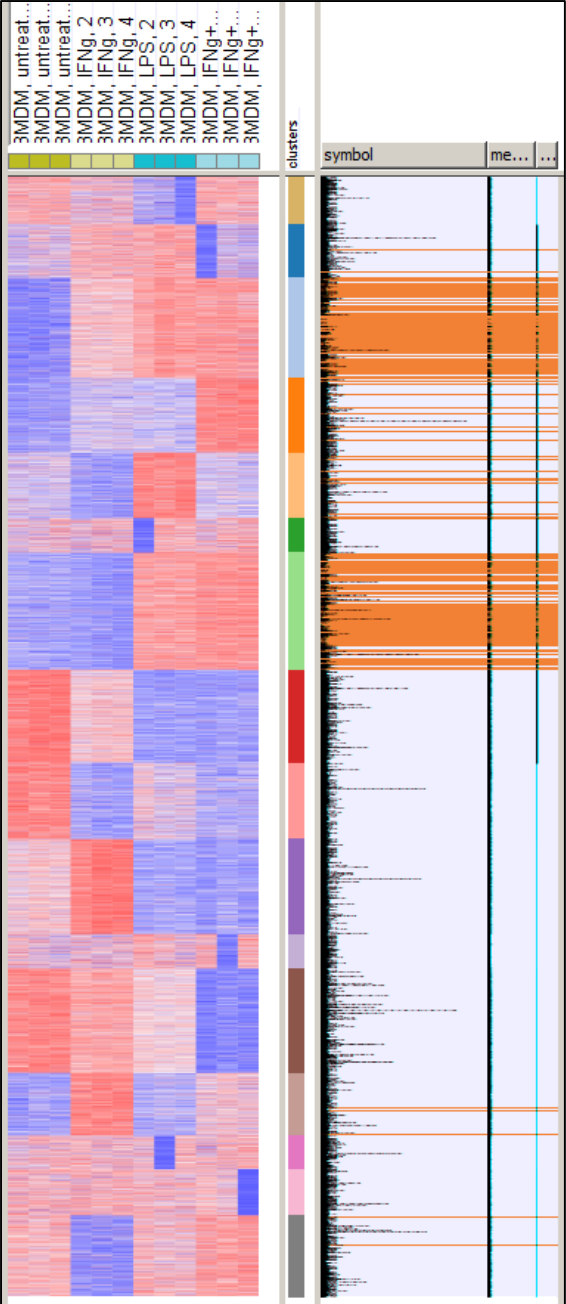
How it looks like in full dataset...



Genes upregulated by LPS in clustered heatmap

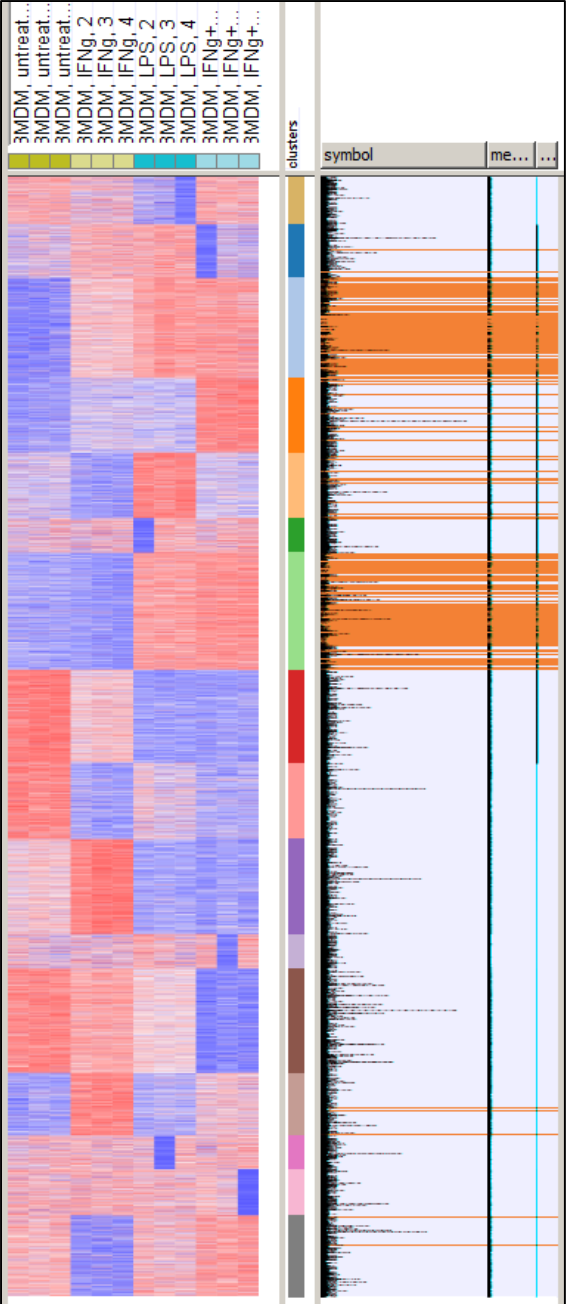
GSE53986
microarray data

Genes upregulated by LPS in your clustered heatmap

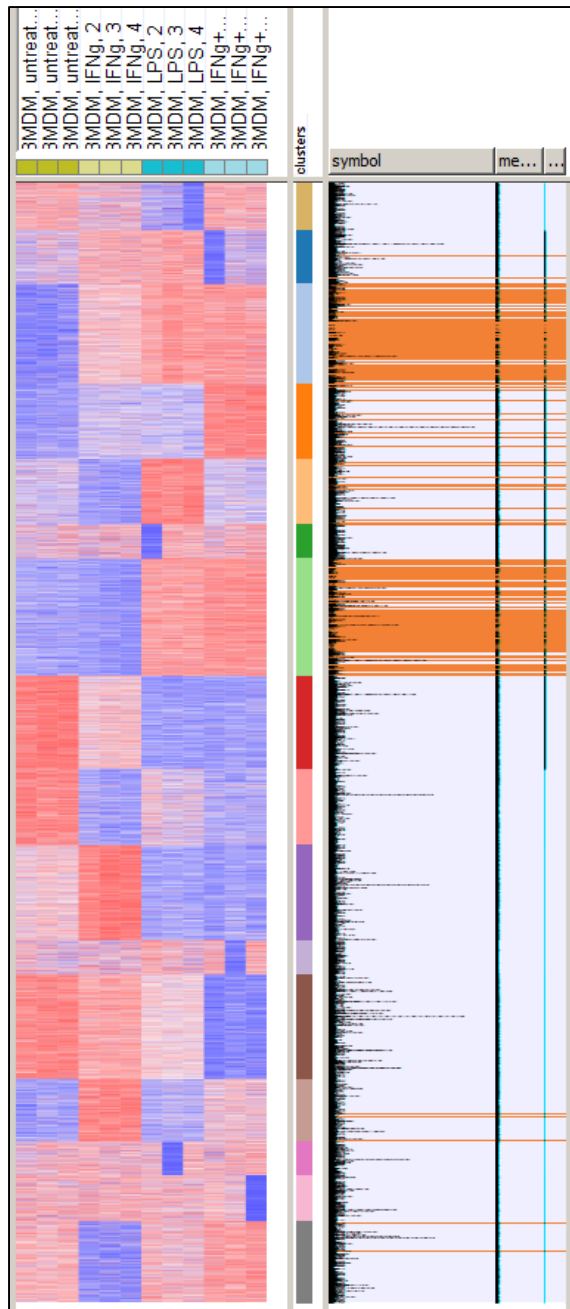


| Gene Set Name [# Genes (K)] | Description | # Genes in Overlap (k) | k/K | p-value | FDR q-value |
|-------------------------------------------------------------------------|-------------------------------------------------------------------------------|------------------------|-----|------------------------|-----------------------|
| HALLMARK_INTERFERON_GAMMA_RESPONSE [200] | Genes up-regulated in response to IFNG [GeneID=3458]. | 100 | | 1.02 e ⁻¹²⁰ | 1.4 e ⁻¹¹⁷ |
| HALLMARK_INTERFERON_ALPHA_RESPONSE [97] | Genes up-regulated in response to alpha interferon proteins. | 51 | | 3.23 e ⁻⁶³ | 2.22 e ⁻⁶⁰ |
| REACTOME_IMMUNE_SYSTEM [933] | Genes involved in Immune System | 102 | | 9.25 e ⁻⁵⁰ | 4.25 e ⁻⁴⁷ |
| REACTOME_CYTOKINE_SIGNALING_IN_IMMUNE_NE_SYSTEM [270] | Genes involved in Cytokine Signaling in Immune system | 41 | | 5.26 e ⁻²⁶ | 1.81 e ⁻²³ |
| REACTOME_INTERFERON_SIGNALING [159] | Genes involved in Interferon Signaling | 33 | | 1.09 e ⁻²⁵ | 3.02 e ⁻²³ |
| REACTOME_CLASS_I_MHC_MEDIATED_ANTIGEN_EN_PROCESSING_PRESENTATION [251] | Genes involved in Class I MHC mediated antigen processing & presentation | 39 | | 3.53 e ⁻²⁵ | 8.11 e ⁻²³ |
| REACTOME_INTERFERON_ALPHA_BETA_SIGNALING [64] | Genes involved in Interferon alpha/beta signaling | 23 | | 2.3 e ⁻²⁴ | 4.53 e ⁻²² |
| REACTOME_ADAPTIVE_IMMUNE_SYSTEM [539] | Genes involved in Adaptive Immune System | 53 | | 5.85 e ⁻²⁴ | 1.01 e ⁻²¹ |
| REACTOME_ANTIGEN_PROCESSING_UBIQUITINATION_PROTEASOME_DEGRADATION [212] | Genes involved in Antigen processing; Ubiquitination & Proteasome degradation | 32 | | 1.66 e ⁻²⁰ | 2.55 e ⁻¹⁸ |
| REACTOME_ER_PHAGOSOME_PATHWAY [61] | Genes involved | 20 | | 1.96 e ⁻²⁰ | 2.7 e ⁻¹⁸ |

Genes upregulated by LPS in your clustered heatmap



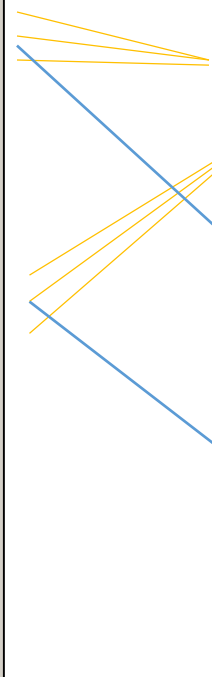
| Gene Set Name [# Genes (K)] | Description | # Genes in k/K | p-value ? | FDR q-value ? |
|-------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------|----------------|-----------------------|-----------------------|
| HALLMARK_TNFA_SIGNALING_VIA_NFKB [200] | Genes regulated by NF-kB in response to TNF [GeneID=7124]. | 45 | 8.11 e ⁻³⁴ | 1.12 e ⁻³⁰ |
| HALLMARK_INFLAMMATORY_RESPONSE [200] | Genes defining inflammatory response. | 44 | 1.21 e ⁻³² | 8.31 e ⁻³⁰ |
| REACTOME_IMMUNE_SYSTEM [933] | Genes involved in Immune System | 86 | 4.15 e ⁻³² | 1.91 e ⁻²⁹ |
| HALLMARK_MTORC1_SIGNALING [200] | Genes up-regulated through activation of mTORC1 complex. | 36 | 9.36 e ⁻²⁴ | 3.23 e ⁻²¹ |
| REACTOME_CYTOKINE_SIGNALING_IN_IMMUNE_NE_SYSTEM [270] | Genes involved in Cytokine Signaling in Immune system | 37 | 3.79 e ⁻²⁰ | 1.05 e ⁻¹⁷ |
| HALLMARK_HYPOXIA [200] | Genes up-regulated in response to low oxygen levels (hypoxia). | 32 | 1.16 e ⁻¹⁹ | 2.67 e ⁻¹⁷ |
| REACTOME_SIGNALING_BY_ILS [107] | Genes involved in Signaling by Interleukins | 24 | 1.24 e ⁻¹⁸ | 2.44 e ⁻¹⁶ |
| REACTOME_ADAPTIVE_IMMUNE_SYSTEM [539] | Genes involved in Adaptive Immune System | 49 | 1.73 e ⁻¹⁸ | 2.99 e ⁻¹⁶ |
| HALLMARK_GLYCOLYSIS [200] | Genes encoding proteins involved in glycolysis and gluconeogenesis. | 27 | 6.49 e ⁻¹⁵ | 9.94 e ⁻¹³ |
| HALLMARK_UV_RESPONSE_UP [158] | Genes up-regulated in response to ultraviolet (UV) radiation. | 23 | 1.27 e ⁻¹³ | 1.76 e ⁻¹¹ |
| HALLMARK_HEME_METABOLISM [200] | Genes involved in metabolism of heme (a cofactor consisting of iron and porphyrin) and erythroblast differentiation. | 25 | 3.84 e ⁻¹³ | 4.81 e ⁻¹¹ |



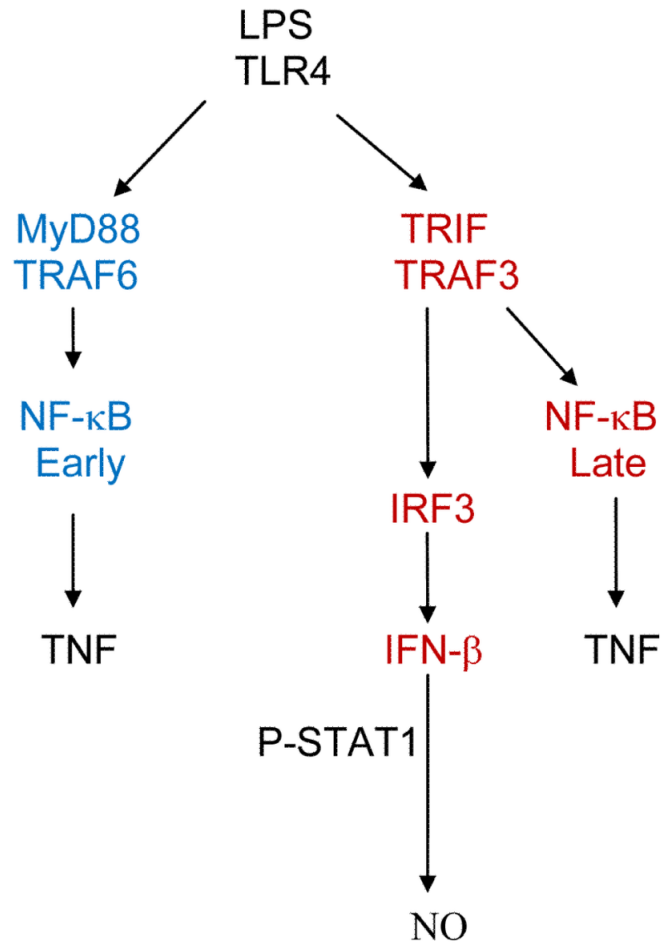
Genes upregulated by LPS in clustered heatmap

Ifng responsive inflammatory genes

Ifng non-responsive inflammatory genes

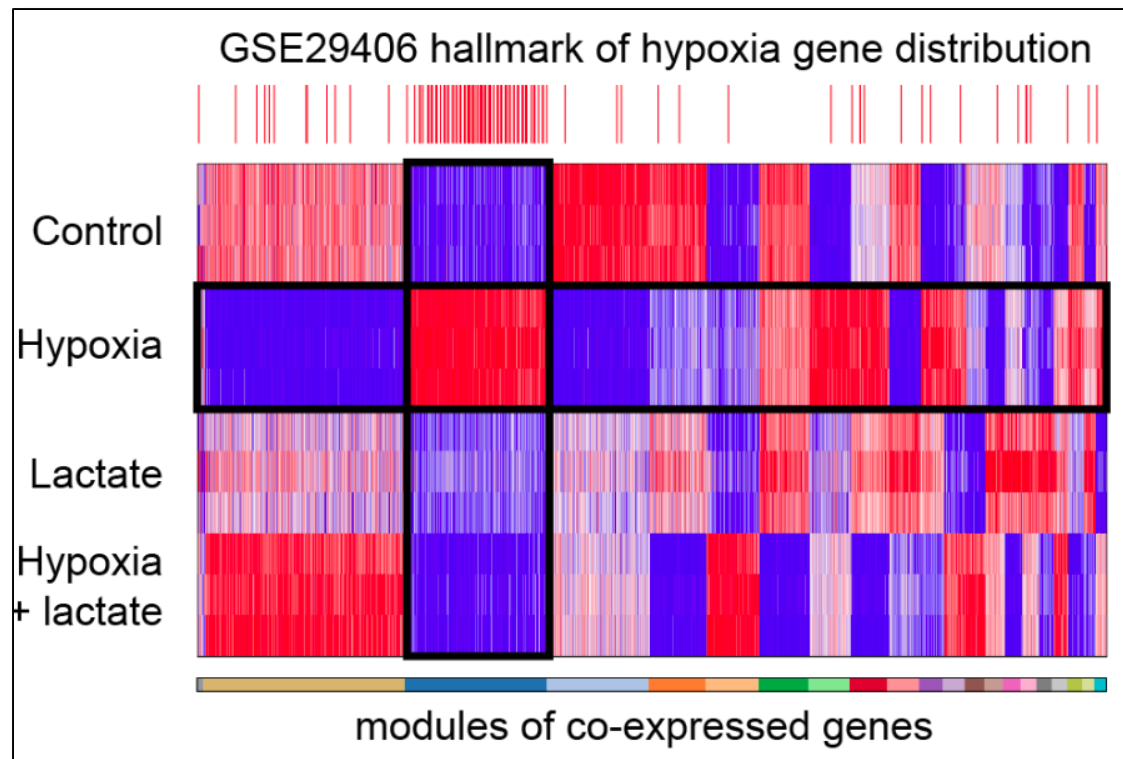


These two transcriptional modules are due to two regulatory modes downstream TLR4



Observation:
biologically meaningful gene sets
are typically co-expressed in relevant datasets!!

Observation: biologically meaningful gene sets are typically co-expressed in relevant datasets!!



Why is it important: Co-expression clustering does NOT require knowledge of conditions



Why is it important: Co-expression clustering does NOT require knowledge of conditions



You can cluster whole GEO!!

NCBI Resources How To

GEO DataSets [Create alert](#) [Advanced](#)

Entry type: clear Summary 20 per page Sort by Number of Samples (High to Low) Send to:

DataSets (0)
Series (2,502)
Samples (0)
Platforms (0)

Organism
Customize ...

Study type
Expression profiling by array
Methylation profiling by array
Customize ...

Author
Customize ...

Attribute name
tissue (688)
strain (835)
Customize ...

Publication dates
30 days
1 year
Custom range...

[Clear all](#)
[Show additional filters](#)

Search results
Items: 1 to 20 of 2502 << First < Prev Page 1 of 126 Next > Last >>

Filters activated: Series. [Clear all](#) to show 30070 items.

[SEQC Project](#)

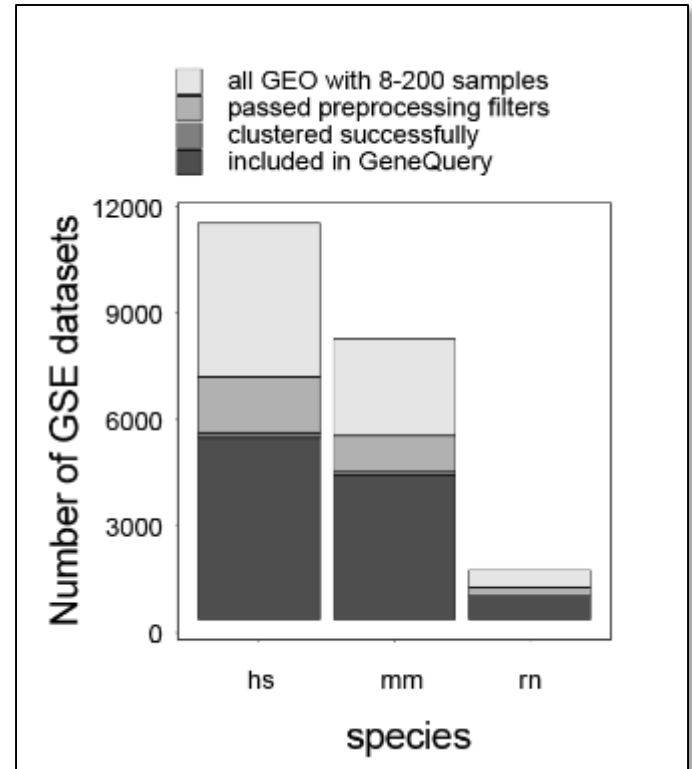
1. (Submitter supplied) This SuperSeries is composed of the SubSeries listed below.
Organism: Rattus norvegicus; synthetic construct; Homo sapiens
Type: Expression profiling by array; Expression profiling by high throughput sequencing
15 related Platforms 4521 Samples
Download data: CEL, TXT
Series Accession: GSE47792 ID: 200047792
[PubMed](#) [Full text in PMC](#) [Similar studies](#) [Analyze with GEO2R](#)

[Expression Project for Oncology \(expO\)](#)

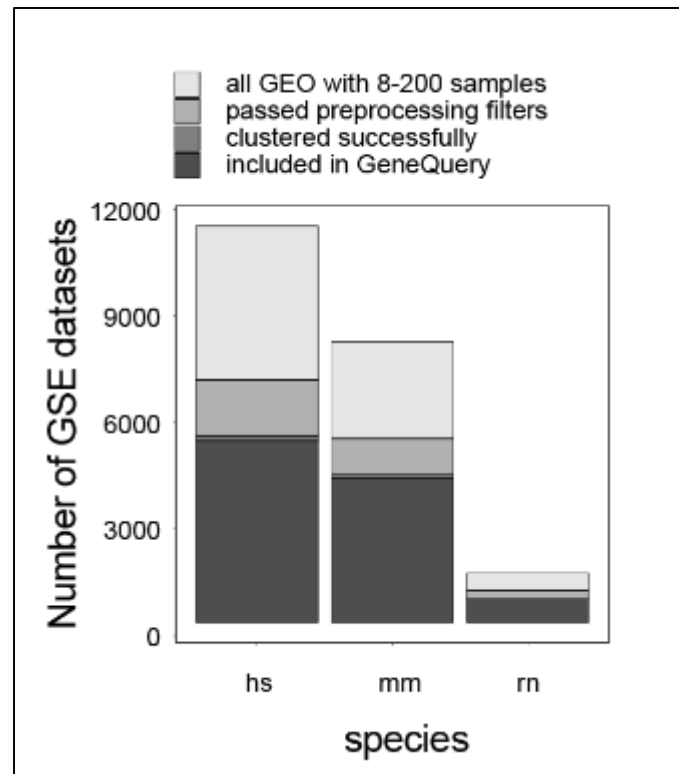
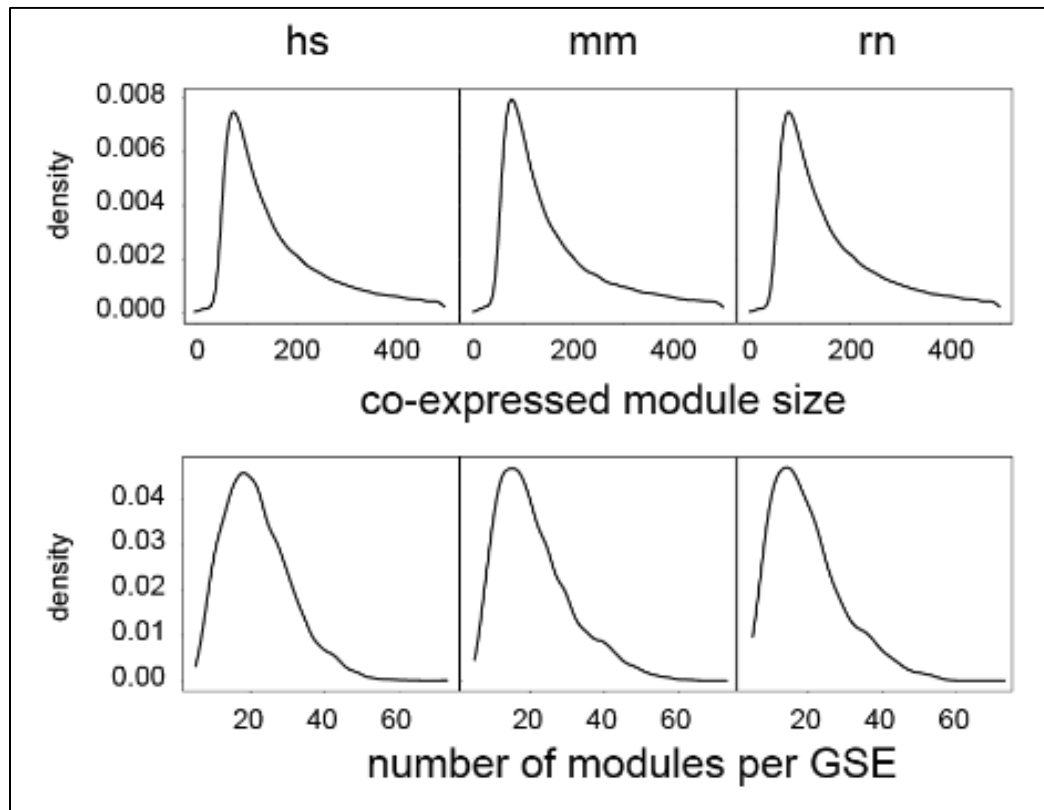
2. (Submitter supplied) The mission of expO is to build on the technologies and outcomes of the Human Genome Project to accelerate improved clinical management of cancer patients. IGC's Expression Project for Oncology (expO) seeks to integrate longitudinal clinical annotation with gene expression data for a unique and powerful portrait of human malignancies, providing critical perspective on diagnostic markers, prognostic indicators, and therapeutic targets. [more...](#)
Organism: Homo sapiens
Type: Expression profiling by array
Platform: GPL570 2158 Samples
Download data: CEL
Series Accession: GSE2109 ID: 200002109
[Analyze with GEO2R](#)

[Utilizing single-cell variations to reveal environment-dependent tuning of network connectivity in human macrophages](#)

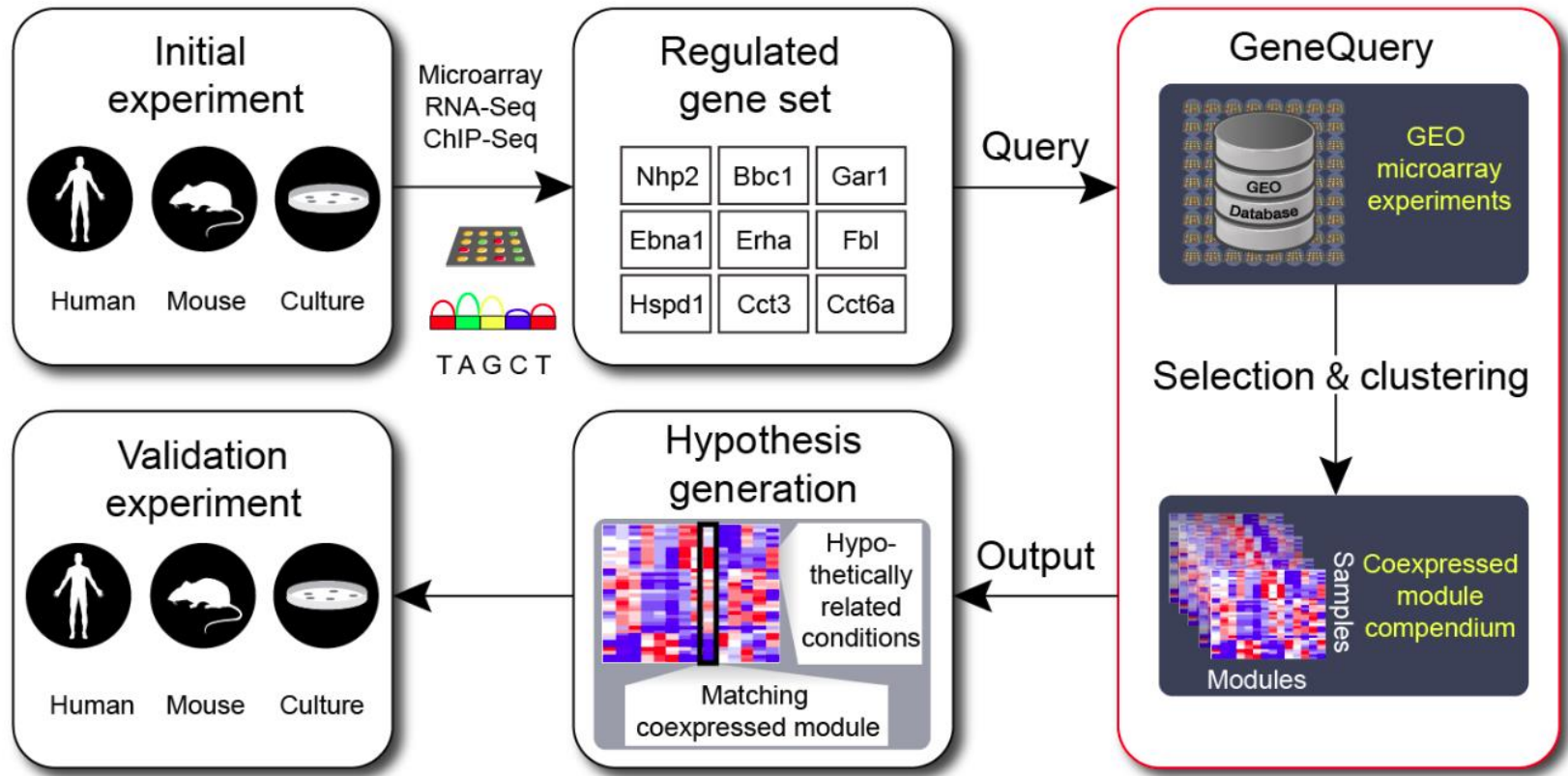
3.



You can cluster whole GEO!!



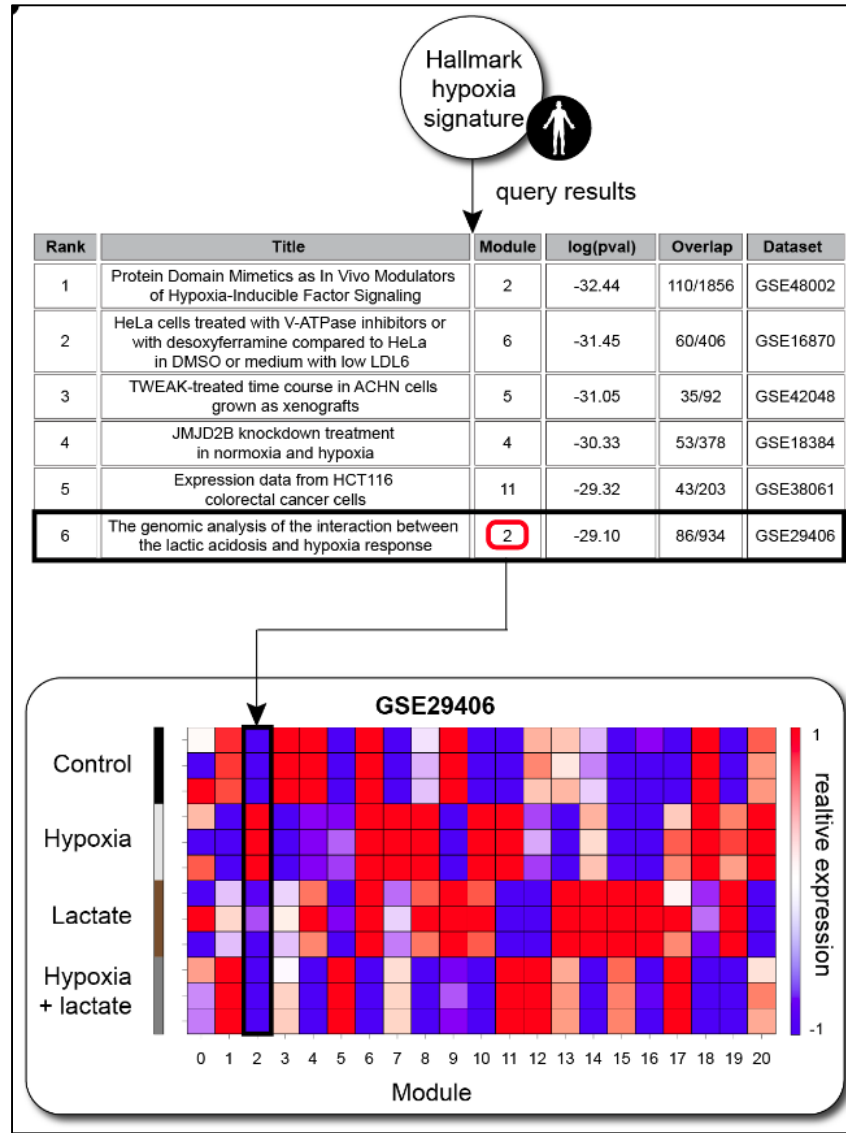
Get the idea for your next experiment!



Get the idea for your next experiment!

<http://artyomovlab.wustl.edu/genequery/searcher/>

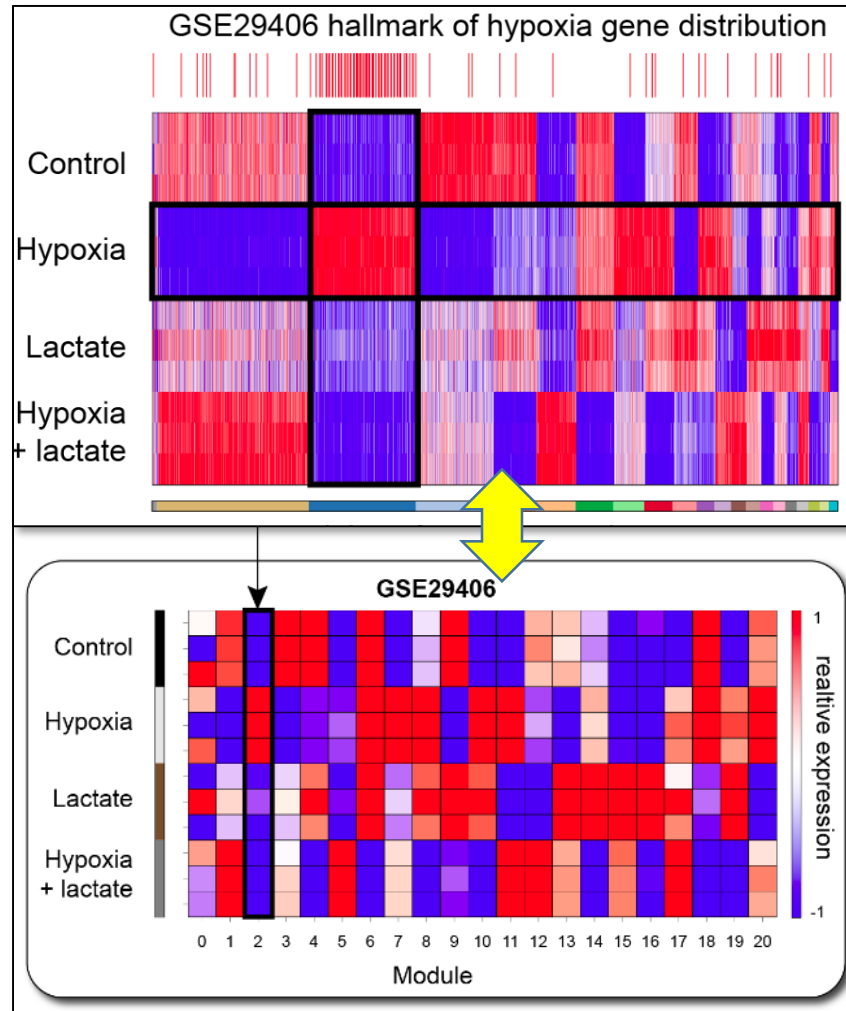
<http://genome.ifmo.ru/genequery/searcher/>



Get the idea for your next experiment!

<http://artyomovlab.wustl.edu/genequery/searcher/>

<http://genome.ifmo.ru/genequery/searcher/>



Eigengene is averaged representation of the cluster behavior

Let's practice with our inflammatory macrophage datasets!

Explore gmt

Let's practice with some hallmark signatures!

