

Genomic set enrichment analysis enhanced through integration of chromatin long-range interactions

Michael M. Hoffman @michaelhoffman



Annie Lu
@zhiyuanlu_annie



Linh Huynh
@vietlinh_huynh



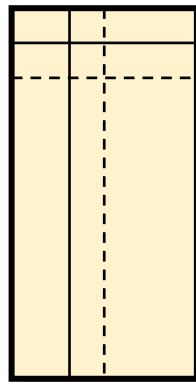
Davide Chicco
@davidechicco_it



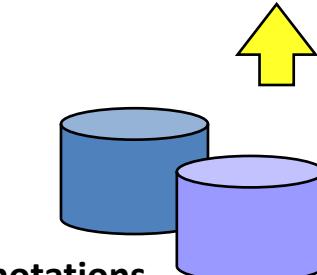
Sarah Bi
@h_s_b_i

Gene set enrichment

Gene expression
data (from RNA-seq
or microarray)



Enrichment
method



Gene annotations
Gene ontology

Enrichment Table

Spindle	0.00001
Apoptosis	0.00025

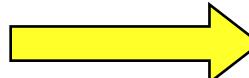
Adapted from Quaid Morris

Enrichment for arbitrary genomic regions

**Genomic
region
(including
noncoding)
list**

chr1:32169580-32169730
chr1:20656800-20656950
chr1:20992300-20992450
chr1:21011700-21011850
chr1:21103160-21103310
chr1:21900720-21900870
...

Gene list

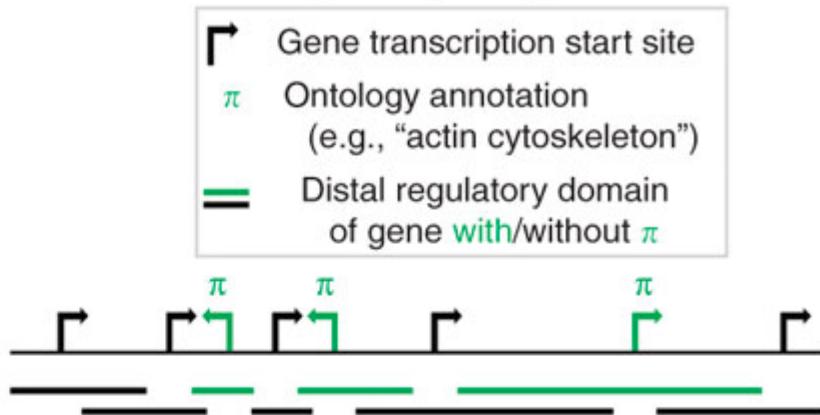


DDX5
KLHL41
MEF2C
MEF2D
MYF6
MYOD1
MYOG
PAX3
PAX7
SOX8

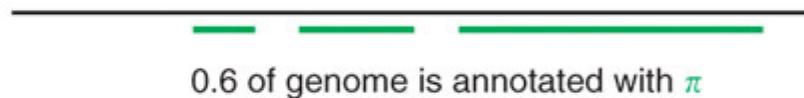
Enrichment in a non-gene context

Genomic Regions Enrichment of Annotations Tool (GREAT)

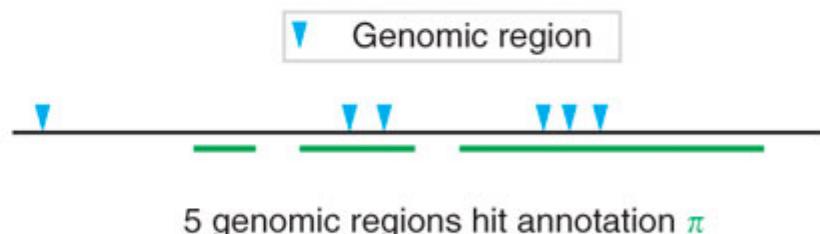
Step 1: Infer distal gene regulatory domains



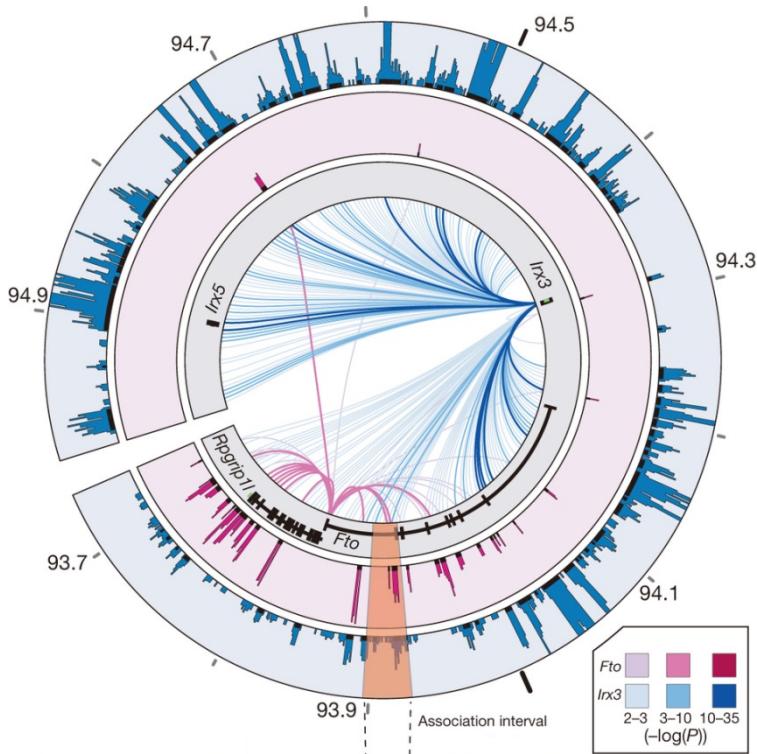
Step 2: Calculate annotated fraction of genome



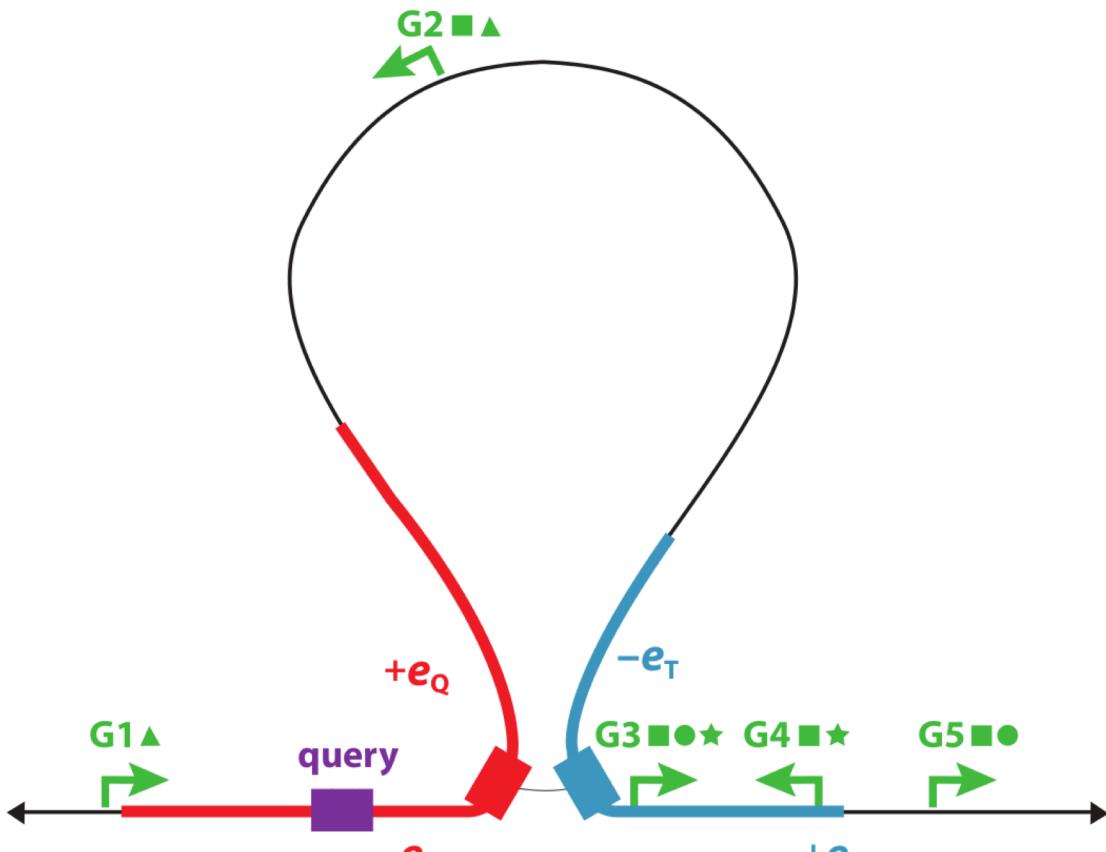
Step 3: Count genomic regions associated with the annotation



Regulatory elements & adjacent genes



Genome-wide association studies (GWAS) have reproducibly associated variants within introns of *FTO* with increased risk for obesity and type 2 diabetes (T2D)^{1–3}. Although the molecular mechanisms linking these noncoding variants with obesity are not immediately obvious, subsequent studies in mice demonstrated that *FTO* expression levels influence body mass and composition phenotypes^{4–6}. However, no direct connection between the obesity-associated variants and *FTO* expression or function has been made^{7–9}. Here we show that the obesity-associated noncoding sequences within *FTO* are functionally connected, at megabase distances, with the homeobox gene *IRX3*. The obesity-associated *FTO* region directly interacts with the promoters of *IRX3* as well as *FTO* in the human, mouse and zebrafish genomes. Furthermore, long-range enhancers within this region recapitulate aspects of *IRX3* expression, suggesting that the obesity-associated interval belongs to the regulatory landscape of *IRX3*. Consistent with this, obesity-associated single nucleotide polymorphisms are associated with expression of *IRX3*, but not *FTO*, in human brains. A direct link between *IRX3* expression and regula-



target genes
and annotation terms

$G_3 \blacksquare \bullet \star$
 $G_4 \blacksquare \star$



enriched
terms

Legend

- query
- transcription start site
- ▲★ annotation terms
- query extension
- target extension

BEST

Biological Enrichment of Sequence Targets

BEHST

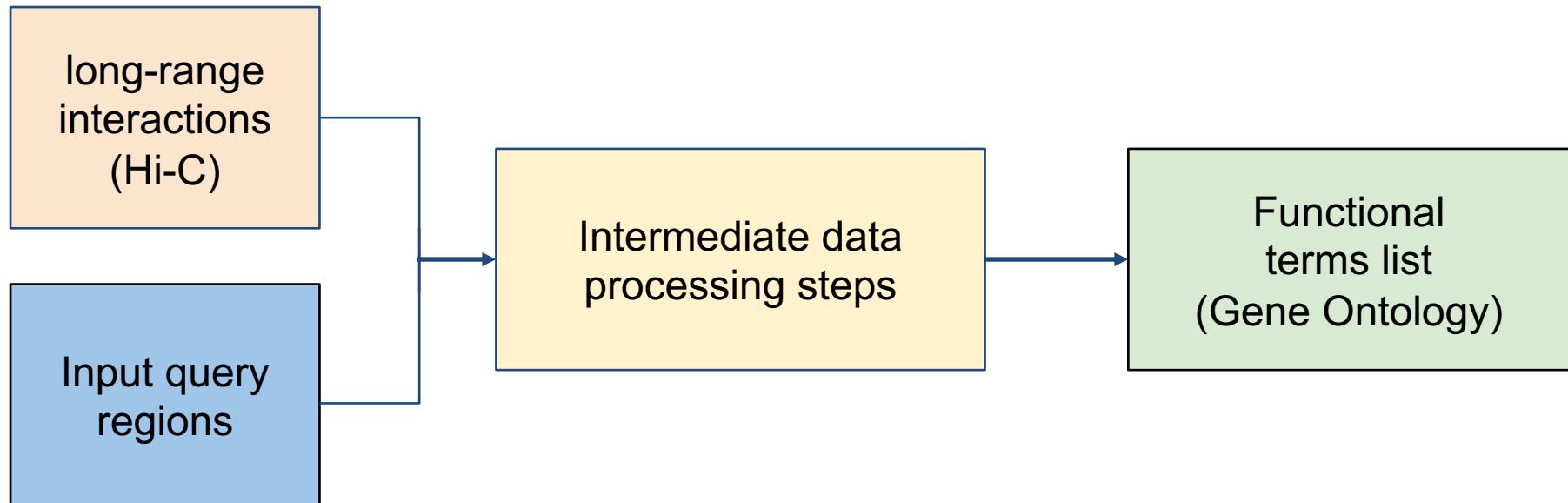
Biological Enrichment of Hidden Sequence Targets

BEHST

Biological Enrichment of Hidden Sequence Targets

Made in Canada! 

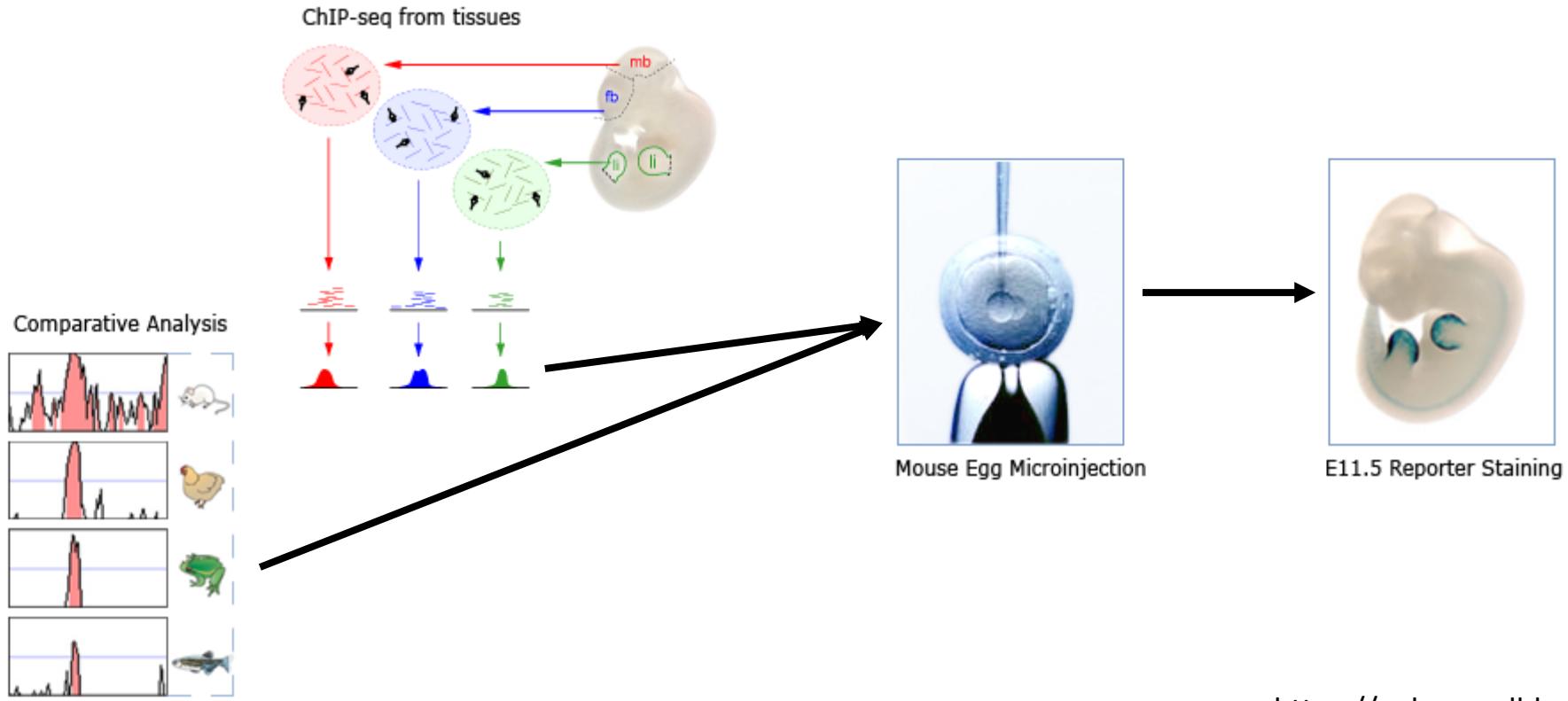
BEHST workflow



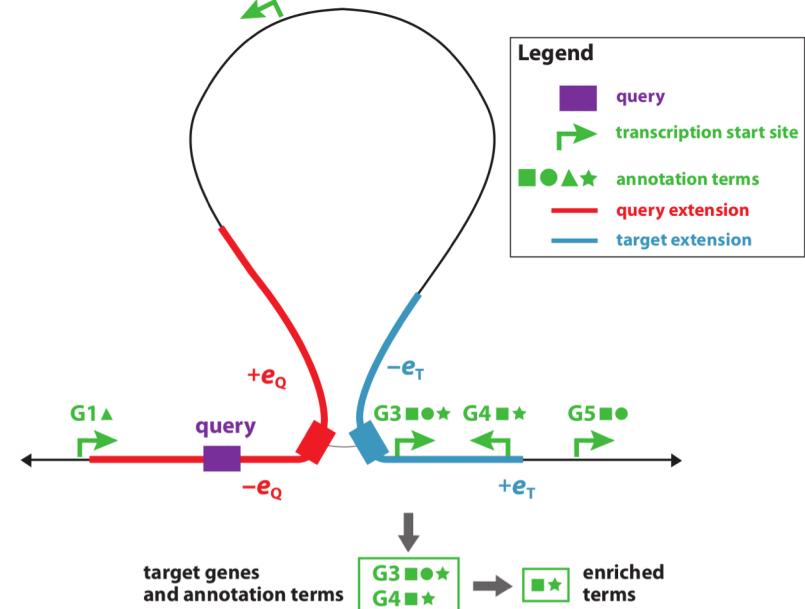
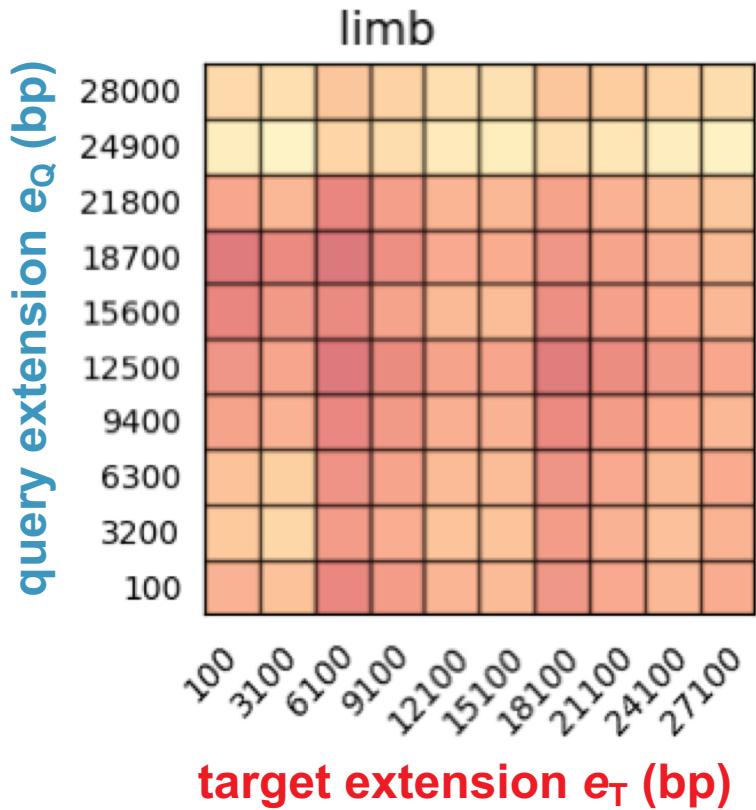
Hi-C datasets

cell type	description	# Hi-C interactions	mean interaction resolution (bp)
GM12878	B-lymphocyte lymphoblastoid	9 448	1 173 831
HeLa-S3	epithelioid cervical carcinoma	3 094	1 435 018
HMEC	mammary epithelial cell	5 152	215 167
HUVEC	umbilical vein endothelial cells	3 865	389 545
IMR90	fetal lung fibroblasts	8 040	416 673
K562	immortalized myelogenous leukemia	6 057	656 974
KBM7	chronic myelogenous leukemia	2 634	487 749
NHEK	normal epidermal keratinocytes	4 929	434 663
Union	union of 8 cell types, excluding duplicates	34 367	742 691

Use case: E11.5 mouse enhancers



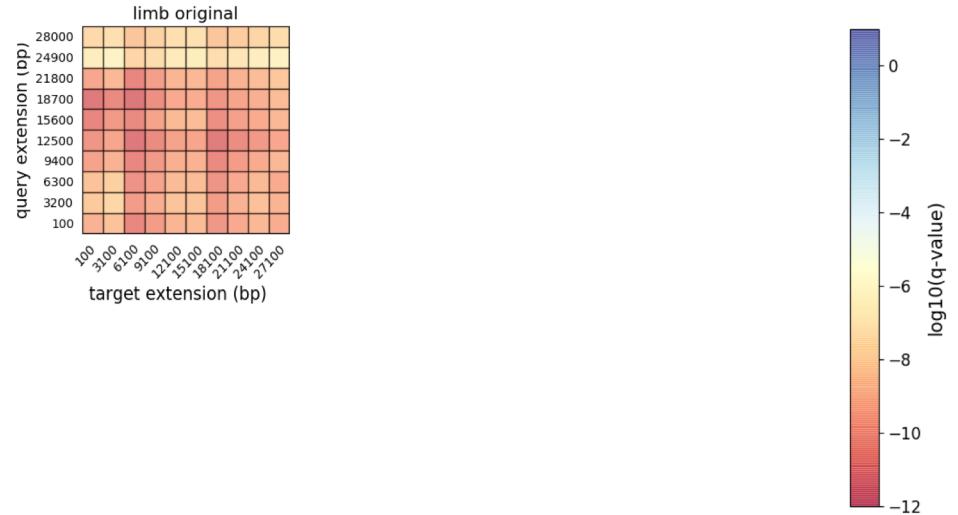
Grid search of extension parameters



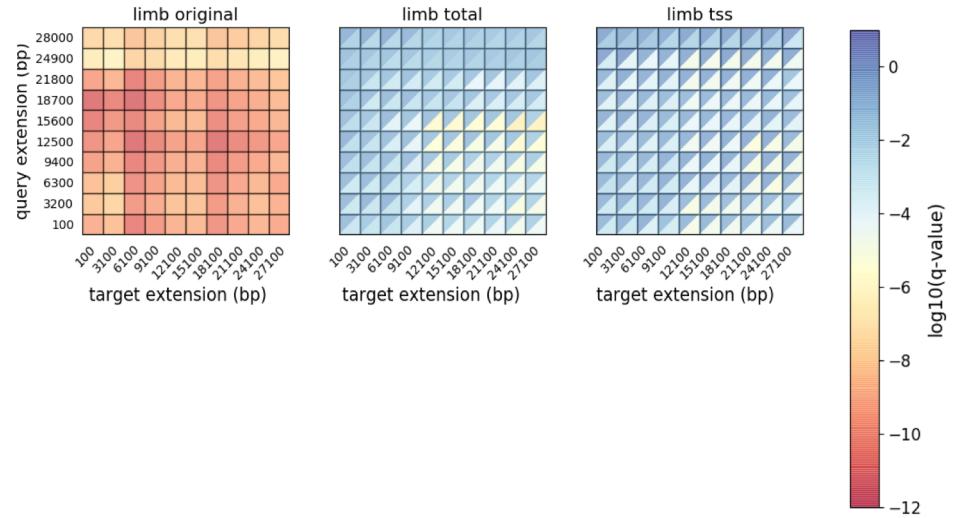
Shuffled controls

- Expectation: BEHST outputs more significant enrichment from original data than random data
- Applied BEHST to 7 sets of VISTA enhancers
- Compared with two shuffled negative controls:
 1. **Total shuffle**: randomly shuffle the enhancers across the whole genome
 2. **TSS shuffle**: shuffle in a way that preserved distance to the nearest transcription start site (TSS)

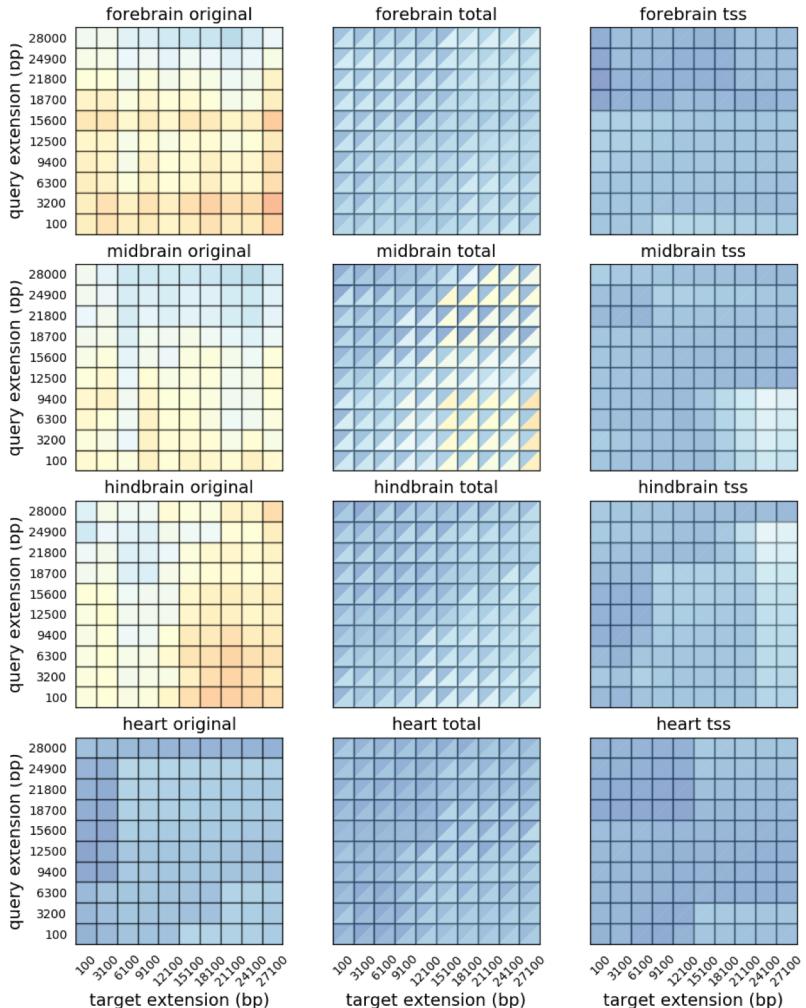
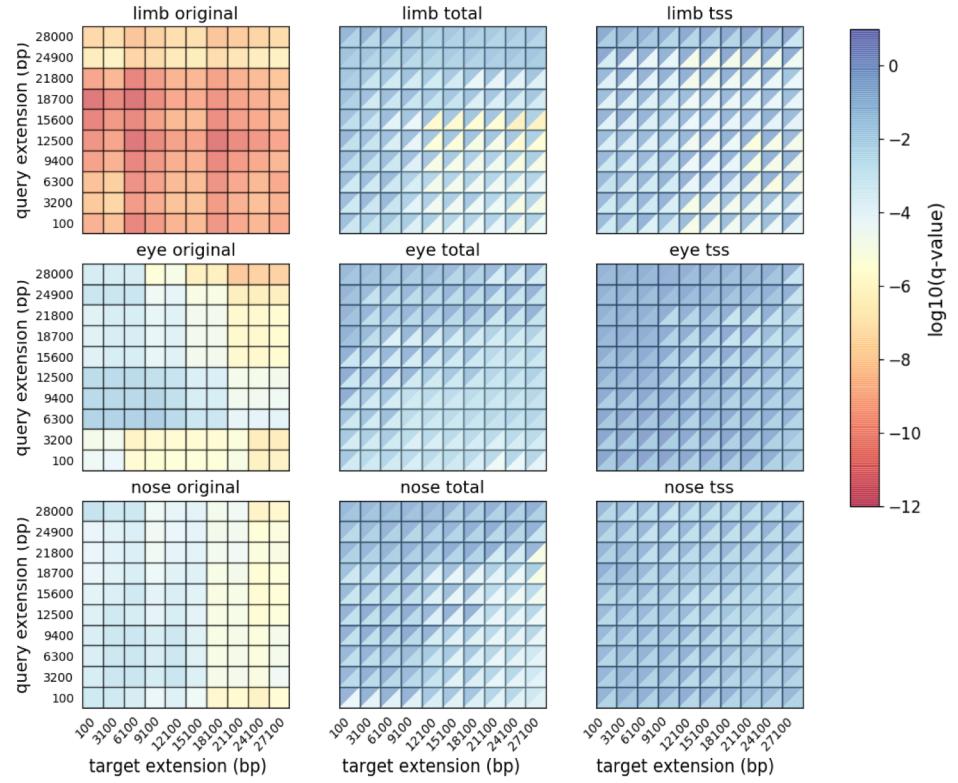
Comparing to shuffled controls



Comparing to shuffled controls



Comparing to shuffled controls



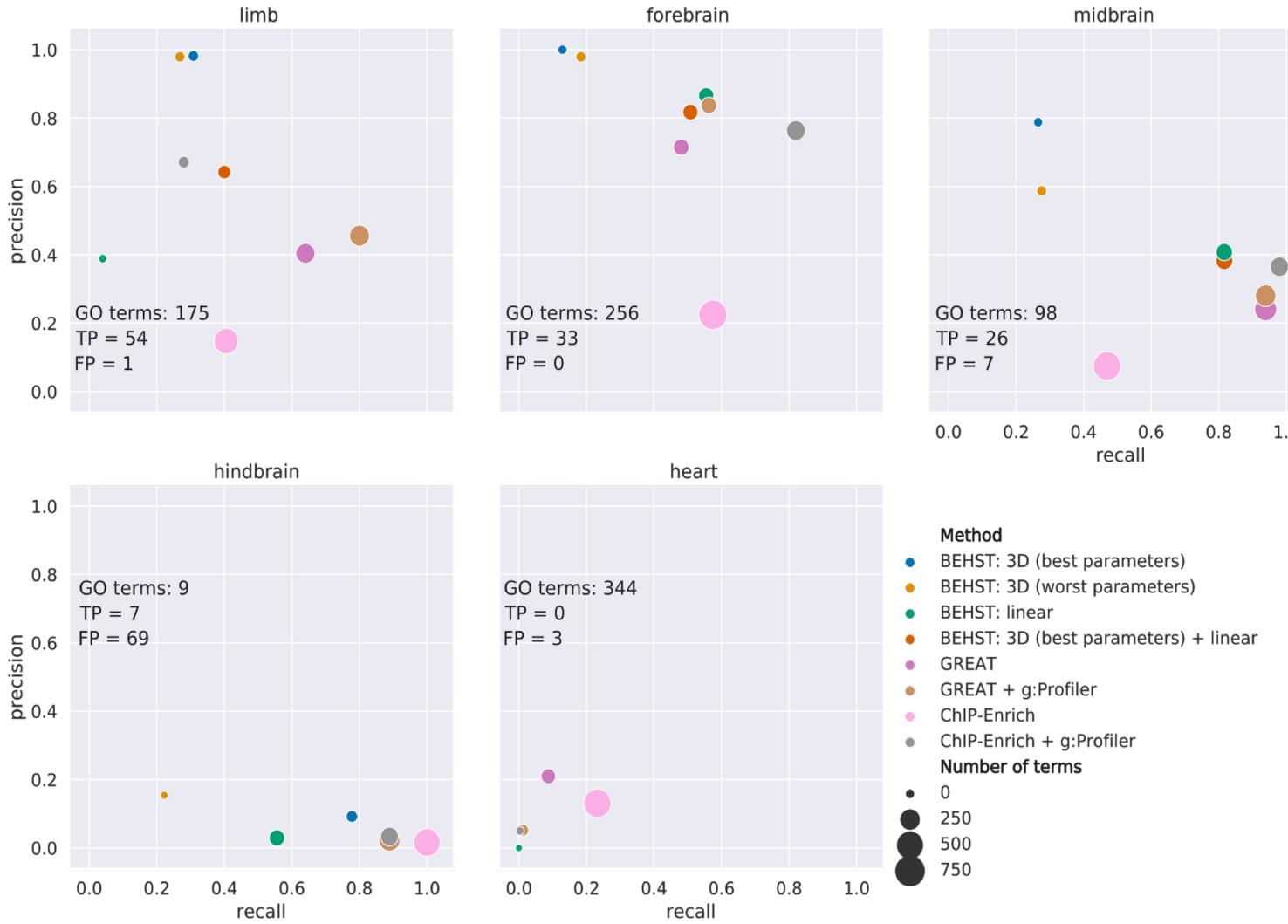
Comparison between BEHST and other tools

- Problem with old method:
manually, biased, ad-hoc interpretation

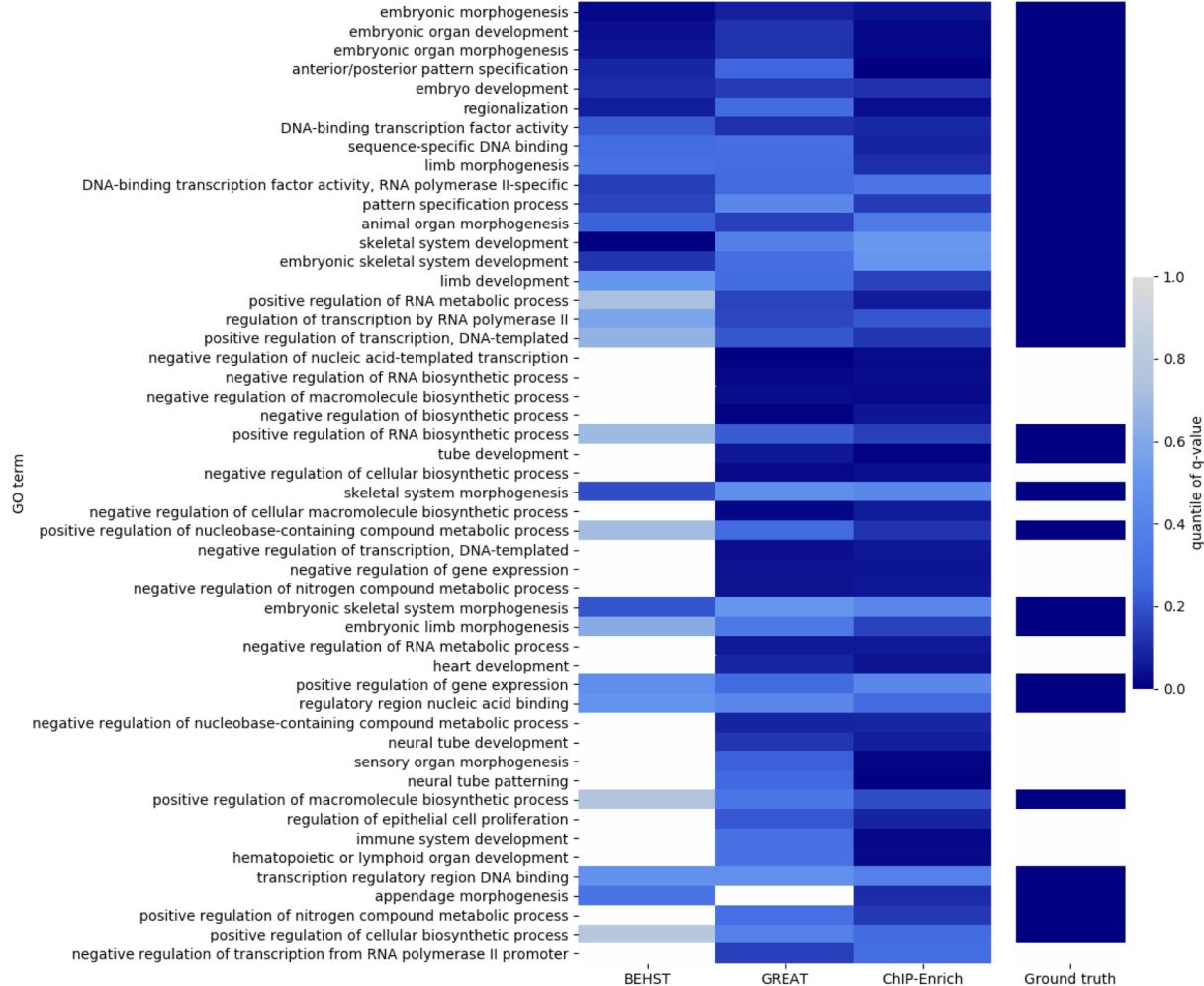
	sub-ontology	term ID	EF/UF	term name
8.21 × 10 ⁻¹⁴	MF	GO:0003700		sequence-specific DNA binding factor activity
1.04 × 10 ⁻¹³	MF	GO:0001071		nucleic acid binding transcription factor activity
2.47 × 10 ⁻⁰⁹	BP	GO:0072358	UF	cardiovascular system development
3.00 × 10 ⁻⁰⁹		GO:0007507	UF	heart development
1.06 × 10 ⁻⁰⁸	BP	GO:0035108	EF	limb morphogenesis
1.07 × 10 ⁻⁰⁸	BP	GO:0060173	EF	limb development
1.21 × 10 ⁻⁰⁸	BP	GO:0045892		negative regulation of transcription, DNA-dependent
1.27 × 10 ⁻⁰⁸	BP	GO:0001287		organ morphogenesis
1.33 × 10 ⁻⁰⁸	BP	GO:0001286		positive regulation of cellular macromolecule synthetic process
2.88 × 10 ⁻⁰⁸	BP	GO:0051235		negative regulation of RNA metabolic process
3.31 × 10 ⁻⁰⁸	BP	GO:0035295		tube development
3.36 × 10 ⁻⁰⁸	BP	GO:0010629		negative regulation of gene expression
3.82 × 10 ⁻⁰⁸	BP	GO:0010555		negative regulation of macromolecule synthetic process
7.97 × 10 ⁻⁰⁸	BP	GO:0001288	UF	limb morphogenesis
9.81 × 10 ⁻⁰⁸	BP	GO:0001262		embryo organ morphogenesis
1.42 × 10 ⁻⁰⁷	BP	GO:0030326	EF	embryo limb morphogenesis
1.80 × 10 ⁻⁰⁷	BP	GO:0060562		epithelial morphogenesis
2.19 × 10 ⁻⁰⁷	BP	GO:0035239		tube morphogenesis
2.31 × 10 ⁻⁰⁷	MF	GO:0043565		sequence-specific DNA binding
2.32 × 10 ⁻⁰⁷		GO:0060429		epithelium development
4.26 × 10 ⁻⁰⁷		GO:0000981		sequence-specific DNA binding RNA polymerase II transcript factor activity
2.64 × 10 ⁻⁰⁷	BP	GO:0048643	EF	regulation of skeletal muscle development

Comparison between BEHST and other tools

- New comparison
 - Create a list of **ground-truth GO terms**
 - Choose tissue-specific genes from RNA-seq data
 - $\text{TPM} > 1$ and $\text{TPM} > 5 \langle \text{TPM}_{\text{other}} \rangle$
 - Run g:Profiler on these genes
 - Intersect the ground-truth GO term list with the GO terms from
 - BEHST
 - GREAT, GREAT-g:Profiler hybrid
 - ChIP-Enrich, ChIP-Enrich-g:Profiler hybrid
 - GO terms in both lists are true positive terms
 - GO terms only in output list but not ground-truth list are false positive terms



GO BP terms found by three methods

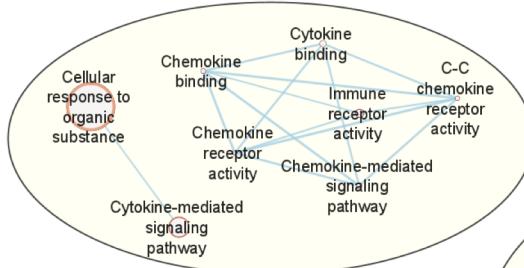


UK Biobank GWAS Data

- Get 17 anthropometric and blood-panel traits in the UK Biobank
- Select positions where p-value of beta-meta significance test $< 10^{-8}$
- Add eQ = 1000 bp to the single positions and run BEHST
- Find clusters of gene sets with Enrichment Map

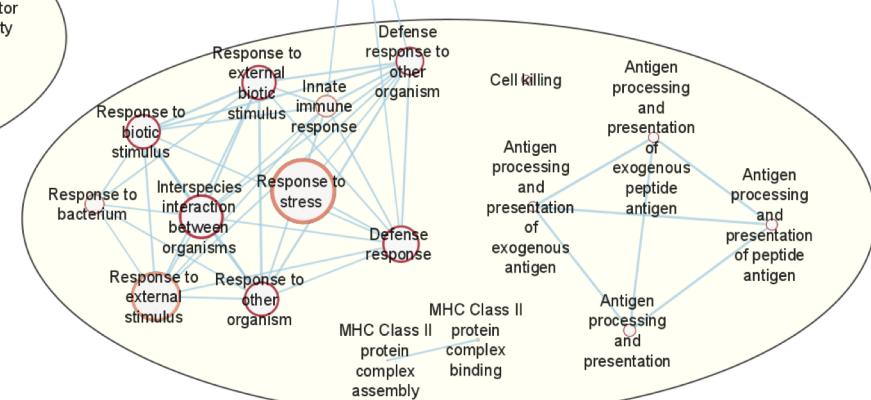
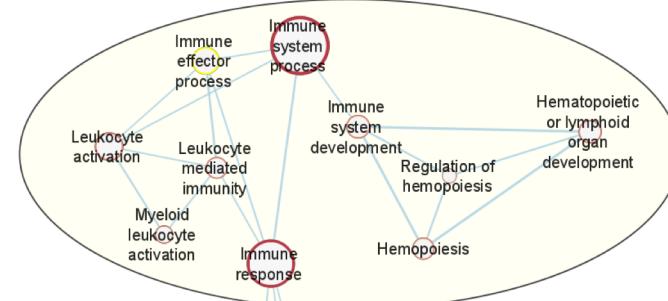
Application to UK Biobank GWAS for Basophil number

Receptor Chemokine Binding



Threshold for node and edge display:
p-value < 0.001, edge threshold < 0.5

Lymphoid Development Hemopoiesis



Immune Response

BEHST Biological Enrichment of Hidden Sequence Targets

Genomic set enrichment analysis enhanced through integration of chromatin long-range interactions

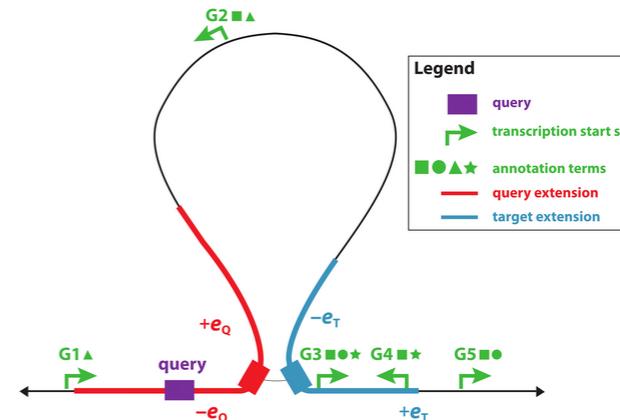
Chicco D, Bi HS, Reimand J, Hoffman MM. 2017. BEHST – Genomic set enrichment analysis enhanced through integration of chromatin long-range interactions. *In preparation.*

The free BEHST software package efficiently associates functional enriched Gene Ontology terms to input genomic regions

BEHST reads a dataset of genomic regions, and intersects them with the chromatin interactions available in the Hi-C dataset (Rao et al, Cell, 2014). Of these genomic regions, BEHST selects those that are present in the regulatory regions of genes a dataset of principal isoform annotations. We defined these cis-regulatory regions upon the position of their nearest transcription start site of the genes' principal transcripts, plus an upstream and downstream extension. Afterwards, BEHST selects the genes of the resulting partner loci found in gene regulatory regions, and inserts them into g:Profiler. BEHST, finally, produces the list of the most significant Gene Ontology terms detected by g:Profiler.

Installation

BEHST can run on any Linux and Mac computers. You can find the





Search



Upload your files

Query regions

File upload/ URL (.bed)

Optional parameters and files:

Query extension (bp)

Target extension (bp)

Gene annotation (.gtf)

Chromosomal interactions (.hiccups)

Principal transcripts (.bed)



g:GOST Gene Group Functional Profiling

g:Cocoa Compact Compare of Annotations

g:Convert Gene ID Converter

g:Sorter Expression Similarity Search

g:Orth Orthology search

g:SNPense Convert rsID

[Welcome!](#) [Contact](#) [FAQ](#) [R / APIs](#) [Beta](#) [Archive](#)

J. Reimand, T. Arak, P. Adler, L. Kolberg, S. Reisberg, H. Peterson, J. Vilo: g:Profiler -- a web server for functional interpretation of gene lists (2016 update) Nucleic Acids Research 2016; doi: 10.1093/nar/gkw199 ([PDF](#), [more](#))

[?] Organism

Homo sapiens

[?] Query (genes, proteins, probes)

ENSG00000000938 ENSG00000003393 ^
 ENSG00000003400 ENSG00000003402
 ENSG00000005844 ENSG00000006059
 ENSG00000007933 ENSG00000013583
 ENSG00000013588 ENSG00000019186
 ENSG00000020181 ENSG00000022556
 ENSG00000024422 ENSG00000025708
 ENSG00000025770 ENSG00000026103
 ENSG00000026751 ENSG00000028528

[?] or Term ID:

 [g:Profile!](#) [Clear](#)

Example or random query

g:Profiler version r1732_e89_eg36. Version info

Options

- [?] Significant only
- [?] Ordered query
- [?] No electronic GO annotations
- [?] Chromosomal regions
- [?] Hierarchical sorting
- [?] Hierarchical filtering
- Best per parent group (strong) ▾
- [?] Output type
- Graphical (PNG) ▾
- Show advanced options



[?] Gene Ontology

Inferred from experiment [IDA, IPI, IMP, IGI, IEP]



[?] Biological process



[?] Cellular component



[?] Molecular function



Inferred from experiment [IDA, IPI, IMP, IGI, IEP]



Direct assay [IDA] / Mutant phenotype [IMP]



Genetic interaction [IGI] / Physical interaction [IPI]



Traceable author [TAS] / Non-traceable author [NAS] / Inferred by curator [IC]



Expression pattern [IEP] / Sequence or structural similarity [ISS] / Genomic context [IGC]



Biological aspect of ancestor [IBA] / Rapid divergence [IRD]



Reviewed computational analysis [RCA] / Electronic annotation [IEA]



No biological data [ND] / Not annotated or not in background [NA]



Biological pathways □ KEGG □ Reactome



[?] Regulatory motifs in DNA □ TRANSFAC TFBS □ miRBase microRNAs



[?] Protein databases □ Human Protein Atlas □ CORUM protein complexes



[?] Human Phenotype Ontology (sequence homologs in other species)



[?] BioGRID protein-protein interactions

>> g:Convert

Gene ID Converter

>> g:Orth

Orthology Search

>> g:Sorter

Expression Similarity Search

>> g:Cocoa

Compact Compare of Annotations

>> Static URL

or generate compact link

source term name

term ID	n. of term	n. of query	n. of common genes	corrected p-value	ENSG00000000000	ENSG00000000001	ENSG00000000002	ENSG00000000003	ENSG00000000004	ENSG00000000005	ENSG00000000006	ENSG00000000007	ENSG00000000008	ENSG00000000009	ENSG00000000010	ENSG00000000011	ENSG00000000012	ENSG00000000013	ENSG00000000014	ENSG00000000015	ENSG00000000016	ENSG00000000017	ENSG00000000018	ENSG00000000019	ENSG00000000020	ENSG00000000021	ENSG00000000022	ENSG00000000023	ENSG00000000024	ENSG00000000025	ENSG00000000026	ENSG00000000027	ENSG00000000028	ENSG00000000029	ENSG00000000030	ENSG00000000031	ENSG00000000032	ENSG00000000033	ENSG00000000034	ENSG00000000035	ENSG00000000036	ENSG00000000037	ENSG00000000038	ENSG00000000039	ENSG00000000040	ENSG00000000041	ENSG00000000042	ENSG00000000043	ENSG00000000044	ENSG00000000045	ENSG00000000046	ENSG00000000047	ENSG00000000048	ENSG00000000049	ENSG00000000050	ENSG00000000051	ENSG00000000052	ENSG00000000053	ENSG00000000054	ENSG00000000055	ENSG00000000056	ENSG00000000057	ENSG00000000058	ENSG00000000059	ENSG00000000060	ENSG00000000061	ENSG00000000062	ENSG00000000063	ENSG00000000064	ENSG00000000065	ENSG00000000066	ENSG00000000067	ENSG00000000068	ENSG00000000069	ENSG00000000070	ENSG00000000071	ENSG00000000072	ENSG00000000073	ENSG00000000074	ENSG00000000075	ENSG00000000076	ENSG00000000077	ENSG00000000078	ENSG00000000079	ENSG00000000080	ENSG00000000081	ENSG00000000082	ENSG00000000083	ENSG00000000084	ENSG00000000085	ENSG00000000086	ENSG00000000087	ENSG00000000088	ENSG00000000089	ENSG00000000090	ENSG00000000091	ENSG00000000092	ENSG00000000093	ENSG00000000094	ENSG00000000095	ENSG00000000096	ENSG00000000097	ENSG00000000098	ENSG00000000099	ENSG00000000100	ENSG00000000101	ENSG00000000102	ENSG00000000103	ENSG00000000104	ENSG00000000105	ENSG00000000106	ENSG00000000107	ENSG00000000108	ENSG00000000109	ENSG00000000110	ENSG00000000111	ENSG00000000112	ENSG00000000113	ENSG00000000114	ENSG00000000115	ENSG00000000116	ENSG00000000117	ENSG00000000118	ENSG00000000119	ENSG00000000120	ENSG00000000121	ENSG00000000122	ENSG00000000123	ENSG00000000124	ENSG00000000125	ENSG00000000126	ENSG00000000127	ENSG00000000128	ENSG00000000129	ENSG00000000130	ENSG00000000131	ENSG00000000132	ENSG00000000133	ENSG00000000134	ENSG00000000135	ENSG00000000136	ENSG00000000137	ENSG00000000138	ENSG00000000139	ENSG00000000140	ENSG00000000141	ENSG00000000142	ENSG00000000143	ENSG00000000144	ENSG00000000145	ENSG00000000146	ENSG00000000147	ENSG00000000148	ENSG00000000149	ENSG00000000150	ENSG00000000151	ENSG00000000152	ENSG00000000153	ENSG00000000154	ENSG00000000155	ENSG00000000156	ENSG00000000157	ENSG00000000158	ENSG00000000159	ENSG00000000160	ENSG00000000161	ENSG00000000162	ENSG00000000163	ENSG00000000164	ENSG00000000165	ENSG00000000166	ENSG00000000167	ENSG00000000168	ENSG00000000169	ENSG00000000170	ENSG00000000171	ENSG00000000172	ENSG00000000173	ENSG00000000174	ENSG00000000175	ENSG00000000176	ENSG00000000177	ENSG00000000178	ENSG00000000179	ENSG00000000180	ENSG00000000181	ENSG00000000182	ENSG00000000183	ENSG00000000184	ENSG00000000185	ENSG00000000186	ENSG00000000187	ENSG00000000188	ENSG00000000189	ENSG00000000190	ENSG00000000191	ENSG00000000192	ENSG00000000193	ENSG00000000194	ENSG00000000195	ENSG00000000196	ENSG00000000197	ENSG00000000198	ENSG00000000199	ENSG00000000200	ENSG00000000201	ENSG00000000202	ENSG00000000203	ENSG00000000204	ENSG00000000205	ENSG00000000206	ENSG00000000207	ENSG00000000208	ENSG00000000209	ENSG00000000210	ENSG00000000211	ENSG00000000212	ENSG00000000213	ENSG00000000214	ENSG00000000215	ENSG00000000216	ENSG00000000217	ENSG00000000218	ENSG00000000219	ENSG00000000220	ENSG00000000221	ENSG00000000222	ENSG00000000223	ENSG00000000224	ENSG00000000225	ENSG00000000226	ENSG00000000227	ENSG00000000228	ENSG00000000229	ENSG00000000230	ENSG00000000231	ENSG00000000232	ENSG00000000233	ENSG00000000234	ENSG00000000235	ENSG00000000236	ENSG00000000237	ENSG00000000238	ENSG00000000239	ENSG00000000240	ENSG00000000241	ENSG00000000242	ENSG00000000243	ENSG00000000244	ENSG00000000245	ENSG00000000246	ENSG00000000247	ENSG00000000248	ENSG00000000249	ENSG00000000250	ENSG00000000251	ENSG00000000252	ENSG00000000253	ENSG00000000254	ENSG00000000255	ENSG00000000256	ENSG00000000257	ENSG00000000258	ENSG00000000259	ENSG00000000260	ENSG00000000261	ENSG00000000262	ENSG00000000263	ENSG00000000264	ENSG00000000265	ENSG00000000266	ENSG00000000267	ENSG00000000268	ENSG00000000269	ENSG00000000270	ENSG00000000271	ENSG00000000272	ENSG00000000273	ENSG00000000274	ENSG00000000275	ENSG00000000276	ENSG00000000277	ENSG00000000278	ENSG00000000279	ENSG00000000280	ENSG00000000281	ENSG00000000282	ENSG00000000283	ENSG00000000284	ENSG00000000285	ENSG00000000286	ENSG00000000287	ENSG00000000288	ENSG00000000289	ENSG00000000290	ENSG00000000291	ENSG00000000292	ENSG00000000293	ENSG00000000294	ENSG00000000295	ENSG00000000296	ENSG00000000297	ENSG00000000298	ENSG00000000299	ENSG00000000300	ENSG00000000301	ENSG00000000302	ENSG00000000303	ENSG00000000304	ENSG00000000305	ENSG00000000306	ENSG00000000307	ENSG00000000308	ENSG00000000309	ENSG00000000310	ENSG00000000311	ENSG00000000312	ENSG00000000313	ENSG00000000314	ENSG00000000315	ENSG00000000316	ENSG00000000317	ENSG00000000318	ENSG00000000319	ENSG00000000320	ENSG00000000321	ENSG00000000322	ENSG00000000323	ENSG00000000324	ENSG00000000325	ENSG00000000326	ENSG00000000327	ENSG00000000328	ENSG00000000329	ENSG00000000330	ENSG00000000331	ENSG00000000332	ENSG00000000333	ENSG00000000334	ENSG00000000335	ENSG00000000336	ENSG00000000337	ENSG00000000338	ENSG00000000339	ENSG00000000340	ENSG00000000341	ENSG00000000342	ENSG00000000343	ENSG00000000344	ENSG00000000345	ENSG00000000346	ENSG00000000347	ENSG00000000348	ENSG00000000349	ENSG00000000350	ENSG00000000351	ENSG00000000352	ENSG00000000353	ENSG00000000354	ENSG00000000355	ENSG00000000356	ENSG00000000357	ENSG00000000358	ENSG00000000359	ENSG00000000360	ENSG00000000361	ENSG00000000362	ENSG00000000363	ENSG00000000364	ENSG00000000365	ENSG00000000366	ENSG00000000367	ENSG00000000368	ENSG00000000369	ENSG00000000370	ENSG00000000371	ENSG00000000372	ENSG00000000373	ENSG00000000374	ENSG00000000375	ENSG00000000376	ENSG00000000377	ENSG00000000378	ENSG00000000379	ENSG00000000380	ENSG00000000381	ENSG00000000382	ENSG00000000383	ENSG00000000384	ENSG00000000385	ENSG00000000386	ENSG00000000387	ENSG00000000388	ENSG00000000389	ENSG00000000390	ENSG00000000391	ENSG00000000392	ENSG00000000393	ENSG00000000394	ENSG00000000395	ENSG00000000396	ENSG00000000397	ENSG00000000398	ENSG00000000399	ENSG00000000400	ENSG00000000401	ENSG00000000402	ENSG00000000403	ENSG00000000404	ENSG00000000405	ENSG00000000406	ENSG00000000407	ENSG00000000408	ENSG00000000409	ENSG00000000410	ENSG00000000411	ENSG00000000412	ENSG00000000413	ENSG00000000414	ENSG00000000415	ENSG00000000416	ENSG00000000417	ENSG00000000418	ENSG00000000419	ENSG00000000420	ENSG00000000421	ENSG00000000422	ENSG00000000423	ENSG00000000424	ENSG00000000425	ENSG00000000426	ENSG00000000427	ENSG00000000428	ENSG00000000429	ENSG00000000430	ENSG00000000431	ENSG00000000432	ENSG00000000433	ENSG00000000434	ENSG00000000435	ENSG00000000436	ENSG00000000437	ENSG00000000438	ENSG00000000439	ENSG00000000440	ENSG00000000441	ENSG00000000442	ENSG00000000443	ENSG00000000444	ENSG00000000445	ENSG00000000446	ENSG00000000447	ENSG00000000448	ENSG00000000449	ENSG00000000450	ENSG00000000451	ENSG00000000452	ENSG00000000453	ENSG00000000454	ENSG00000000455	ENSG00000000456	ENSG00000000457	ENSG00000000458	ENSG00000000459	ENSG00000000460	ENSG00000000461	ENSG00000000462	ENSG00000000463	ENSG00000000464	ENSG00000000465	ENSG00000000466	ENSG00000000467	ENSG00000000468	ENSG00000000469	ENSG00000000470	ENSG00000000471	ENSG00000000472	ENSG00000000473	ENSG00000000474	ENSG00000000475	ENSG00000000476	ENSG00000000477	ENSG00000000478	ENSG00000000479	ENSG00000000480	ENSG00000000481	ENSG00000000482	ENSG00000000483	ENSG00000000484	ENSG00000000485	ENSG00000000486	ENSG00000000487	ENSG00000000488	ENSG00000000489	ENSG00000000490	ENSG00000000491	ENSG00000000492	ENSG00000000493	ENSG00000000494	ENSG00000000495	ENSG00000000496	ENSG00000000497	ENSG00000000498	ENSG00000000499	ENSG00000000500	ENSG00000000501	ENSG00000000502	ENSG00000000503	ENSG00000000504	ENSG00000000505	ENSG00000000506	ENSG00000000507	ENSG00000000508



Search

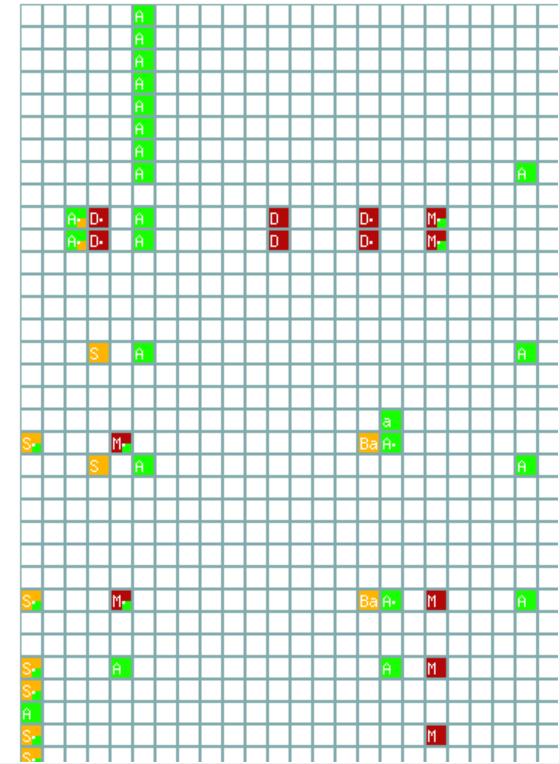


source term name

term ID	n. of term genes	n. of query genes	n. of common genes	corrected p-value
---------	------------------	-------------------	--------------------	-------------------

ENSG000000035664
ENSG000000039887
ENSG000000039850
ENSG000000037241
ENSG000000025708
ENSG000000024422
ENSG000000028528
ENSG000000026751
ENSG000000022556
ENSG000000026103
ENSG000000025770
ENSG000000019186
ENSG000000135988
ENSG00000013883
ENSG000000007933
ENSG000000020181
ENSG0000000340
ENSG00000003844
ENSG00000003400
ENSG0000000393
ENSG00000000938

BP	cornification	G0:0070268	107	621	39	1.17e-25
BP	keratinization	G0:0031424	195	621	42	9.97e-18
BP	keratinocyte differentiation	G0:0030216	254	621	45	1.41e-15
BP	epidermis development	G0:0008544	327	621	51	1.88e-15
BP	epidermal cell differentiation	G0:0009913	276	621	46	7.04e-15
BP	skin development	G0:0043588	301	621	47	4.41e-14
BP	epithelial cell differentiation	G0:0030855	507	621	61	2.95e-13
BP	epithelium development	G0:0060429	773	621	76	3.82e-12
BP	complement activation, classical pathway	G0:0006958	112	621	24	3.45e-09
BP	cell death	G0:0008219	1490	621	108	3.84e-09
BP	programmed cell death	G0:0012501	1397	621	103	4.99e-09
BP	humoral immune response mediated by circulating immunoglobulin	G0:0002455	115	621	24	6.33e-09
BP	complement activation	G0:0006956	162	621	28	1.25e-08
BP	immunoglobulin mediated immune response	G0:0016064	146	621	26	3.58e-08
BP	B cell mediated immunity	G0:0019724	148	621	26	4.93e-08
BP	tissue development	G0:0009888	1208	621	90	8.96e-08
BP	protein activation cascade	G0:0072376	184	621	28	2.85e-07
BP	regulation of humoral immune response	G0:0002920	112	621	20	7.61e-06
BP	lymphocyte mediated immunity	G0:0002449	234	621	29	1.79e-05
BP	immune response	G0:0006955	1768	621	109	4.07e-05
BP	animal organ development	G0:0048513	1812	621	111	4.13e-05
BP	regulation of complement activation	G0:0030449	103	621	18	6.59e-05
BP	regulation of protein activation cascade	G0:2000257	103	621	18	6.59e-05
BP	humoral immune response	G0:0006959	295	621	32	7.87e-05
BP	adaptive immune response based on somatic recombination of immune receptors bui ...	G0:0002460	227	621	27	1.41e-04
BP	phagocytosis, recognition	G0:0006910	58	621	13	2.68e-04
BP	immune system process	G0:0002376	2329	621	131	2.77e-04
BP	adaptive immune response	G0:0002250	265	621	29	2.86e-04
BP	regulation of acute inflammatory response	G0:0002673	127	621	19	3.61e-04
BP	regulation of immune system process	G0:0002682	1179	621	78	3.71e-04
BP	positive regulation of immune response	G0:0050778	670	621	52	4.70e-04
BP	activation of immune response	G0:0002253	578	621	47	4.84e-04
BP	positive regulation of immune system process	G0:0002684	873	621	62	7.21e-04
BP	phagocytosis	G0:0006909	263	621	28	8.35e-04



Concept and methodology in the preprint:

<https://doi.org/fm2z>

New evaluation procedure, GWAS applications:

Revised preprint coming soon!

Acknowledgments

The Hoffman Lab



Samantha Wilson

Eric Roberts

Mickaël Mendez

Annie Lu

Leo Li

Natalia Mukhina

Linh Huynh

Coby Viner

Jeffrey Niu

Aparna Gopalakrishnan

Esther Yu

Davide Chicco

Sarah Bi

Jüri Reimand

Hae Kyung Im

Wail Ba-Alawi

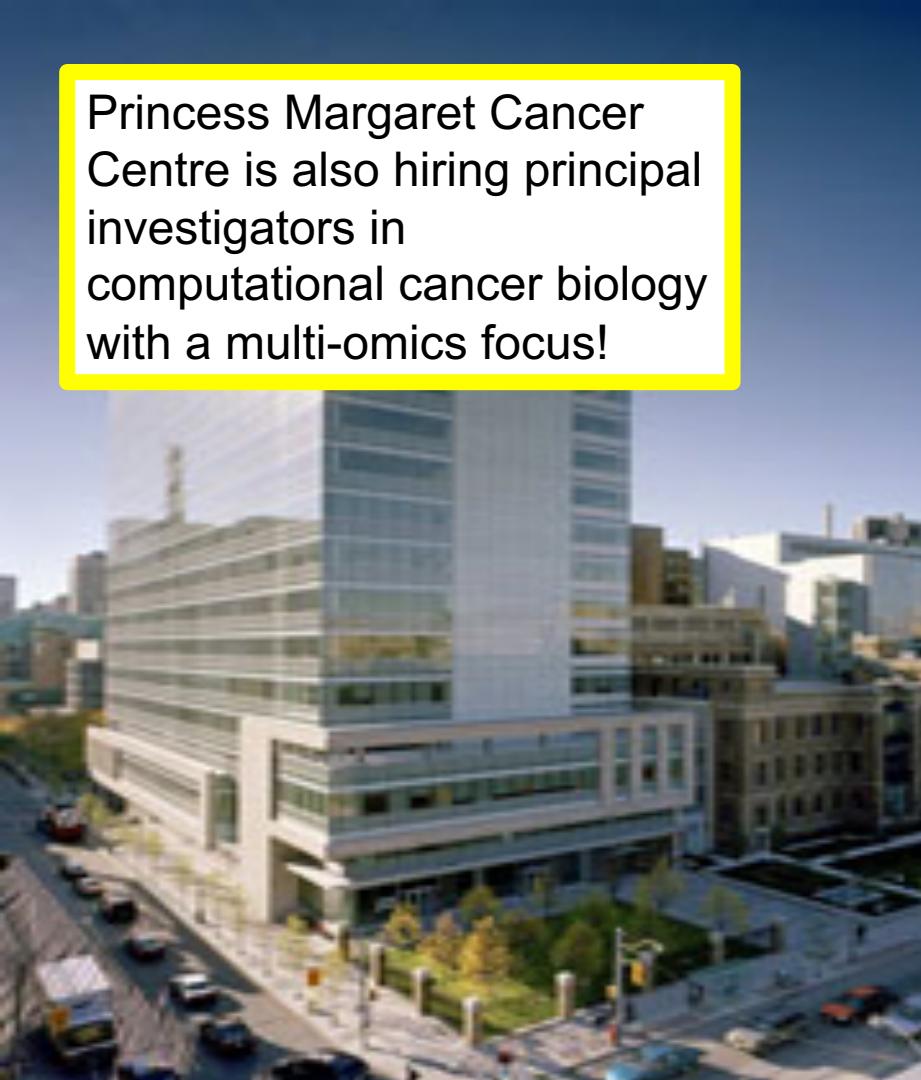
Anna Narday

Zhibin Lu

Carl Virtanen

Funding

Canadian Institutes of Health Research; Princess Margaret Cancer Foundation; Natural Sciences and Engineering Research Council of Canada; Ontario Institute for Cancer Research; Ontario Ministry of Economic Development, Job Creation and Trade; Medicine by Design; McLaughlin Centre



Princess Margaret Cancer Centre is also hiring principal investigators in computational cancer biology with a multi-omics focus!

Postdoctoral, MSc, PhD positions available in my research lab at the

Princess Margaret Cancer Centre

**Dept of Medical Biophysics
Dept of Computer Science
University of Toronto**

Please approach me for details.

Michael Hoffman
<https://hoffmanlab.org/>

michael.hoffman@utoronto.ca
@michaelhoffman