

The principal element of our analytical approach is to define specific sets of genomic regulatory loci  
WITHOUT ANY PRIOR KNOWLEDGE  
of what genes (if any) they may (or may not) control.

This is in striking contrast with the approaches that being utilized in the prior art: to identify genes that are differentially regulated between different states and/or conditions, thus introducing the confounders and biases associated with the multiple hypothesis testing, overfitting, and random co-occurrence due to the very large excess of analyzed features (genes) over the relatively small numbers of samples (100K range or more of analyzed features versus just a few hundred analyzed samples at best). Attempts to address these issues using statistical approaches (corrections for multiple hypothesis testing) and experimental design improvements (multiple replications, validation in independent data sets and using independent analytical techniques) did not adequately resolve these problems.

Importantly, our analyses revealed that this approach identifies not only genes altered expression of which implicated in the broad spectrum of human physiological and pathological conditions. It also identifies genes defined by the independent studies as genetic loci comprising mutation signatures associated with development and progression of multiple common human disorders, including cancer, neurodevelopmental, neuropsychiatric, and neurodegenerative disorders, as well autoimmune and immuno-inflammatory diseases.

Supporting the concept that many major human disorders are driven by aberrant functions of primate-specific genomic regulatory networks with prominent unique-to-human (human-specific) components, this approach identifies genetic loci implicated in inter-individual genetic mosaicism (somatic mosaicism) of cells, tissues and organs in the human body.

# Analytical Pipelines

Define panels of genomic regulatory loci

Human-specific genomic regulatory sequences (HSGRS)

Human stem cell-associated retroviral sequences (SCARS)

Identify genes regulated by HSGRS

Identify genes regulated by SCARS

Heat-Map-Guided (HMG) visualization of gene and protein expression profiles  
*using databases of distinct types of human cells, tissues, anatomical sites, organs, physiological states, pathological conditions, as well as regulatory, chemical, and environmental perturbations*

Comparative gene set enrichment (CGSE) analyses of gene and protein expression profiles  
*using databases of distinct types of human cells, tissues, anatomical sites, organs, physiological states, pathological conditions, as well as regulatory, chemical, and environmental perturbations*

Asses the statistical significance of observations

Formulate the concepts regarding the potential roles and impacts of genomic regulatory networks governed by HSGRS and SCARS in human development, physiology, and pathology

# Analytical Pipelines

Define panels of genomic regulatory loci

Functional enhancers of naïve and primed hESC

Human stem cell-associated retroviral sequences (SCARS)

Identify genes regulated by HSGRS

Identify genes regulated by SCARS

Heat-Map-Guided (HMG) visualization of gene and protein expression profiles

*using databases of distinct types of human cells, tissues, anatomical sites, organs, physiological states, pathological conditions, as well as regulatory, chemical, and environmental perturbations*

Comparative gene set enrichment (CGSE) analyses of gene and protein expression profiles

*using databases of distinct types of human cells, tissues, anatomical sites, organs, physiological states, pathological conditions, as well as regulatory, chemical, and environmental perturbations*

Asses the statistical significance of observations

Formulate the concepts regarding the potential roles and impacts of genomic regulatory networks governed by HSGRS and SCARS in human development, physiology, and pathology

# Analytical Pipelines

Define panels of genomic regulatory loci

Functional enhancers of naïve and primed hESC

Human-specific genomic regulatory sequences (HSGRS)

Identify genes regulated by HSGRS

Identify genes regulated by SCARS

Heat-Map-Guided (HMG) visualization of gene and protein expression profiles

*using databases of distinct types of human cells, tissues, anatomical sites, organs, physiological states, pathological conditions, as well as regulatory, chemical, and environmental perturbations*

Comparative gene set enrichment (CGSE) analyses of gene and protein expression profiles

*using databases of distinct types of human cells, tissues, anatomical sites, organs, physiological states, pathological conditions, as well as regulatory, chemical, and environmental perturbations*

Asses the statistical significance of observations

Formulate the concepts regarding the potential roles and impacts of genomic regulatory networks governed by HSGRS and SCARS in human development, physiology, and pathology

~59K human-specific genomic  
regulatory sequences

13,824 genes are putative regulatory targets of HSGRS

**Structurally, functionally, and evolutionary distinct families of human-specific regulatory sequences (HSRS) and associated putative regulatory target genes defined by the GREAT algorithm.**

<b>Classification category/Reference database</b>	<b>Number of records (hg19)</b>	<b>Associated genes</b>
<b>Fixed human-specific insertions.</b>	<b>11878</b>	<b>7979</b>
<b>Human-specific TE loci expressed in human dorsolateral prefrontal cortex</b>	<b>4637</b>	<b>4051</b>
<b>Set of duplicated regions in GRCh38 space</b>	<b>7599</b>	<b>6618</b>
<b>Fixed human-specific deletions</b>	<b>5883</b>	<b>5489</b>
<b>Human-specific STR expansions</b>	<b>4875</b>	<b>4844</b>
<b>hsTFBS</b>	<b>3803</b>	<b>1087</b>
<b>ace-DHS</b>	<b>3538</b>	<b>3445</b>
<b>FHSRR</b>	<b>4249</b>	<b>2810</b>
<b>Human-specific STR contractions</b>	<b>1279</b>	<b>973</b>
<b>hESC_FHSRR_DHS</b>	<b>1932</b>	<b>1458</b>
<b>DHS_FHSRR (non-hESC)</b>	<b>2118</b>	<b>552</b>
<b>HARs</b>	<b>2745</b>	<b>2281</b>
<b>haDHS</b>	<b>524</b>	<b>747</b>
<b>Human-biased CNCC enhances</b>	<b>1000</b>	<b>1439</b>
<b>Chimp-biased CNCC enhances</b>	<b>1000</b>	<b>1445</b>
<b>H3K4me3 peaks with human-specific enrichment in prefrontal neurons</b>	<b>410</b>	<b>578</b>
<b>Human-specific hESC functional enhancers</b>	<b>1619</b>	<b>1214</b>
<b>All HSRS</b>	<b>59089</b>	<b>13824</b>

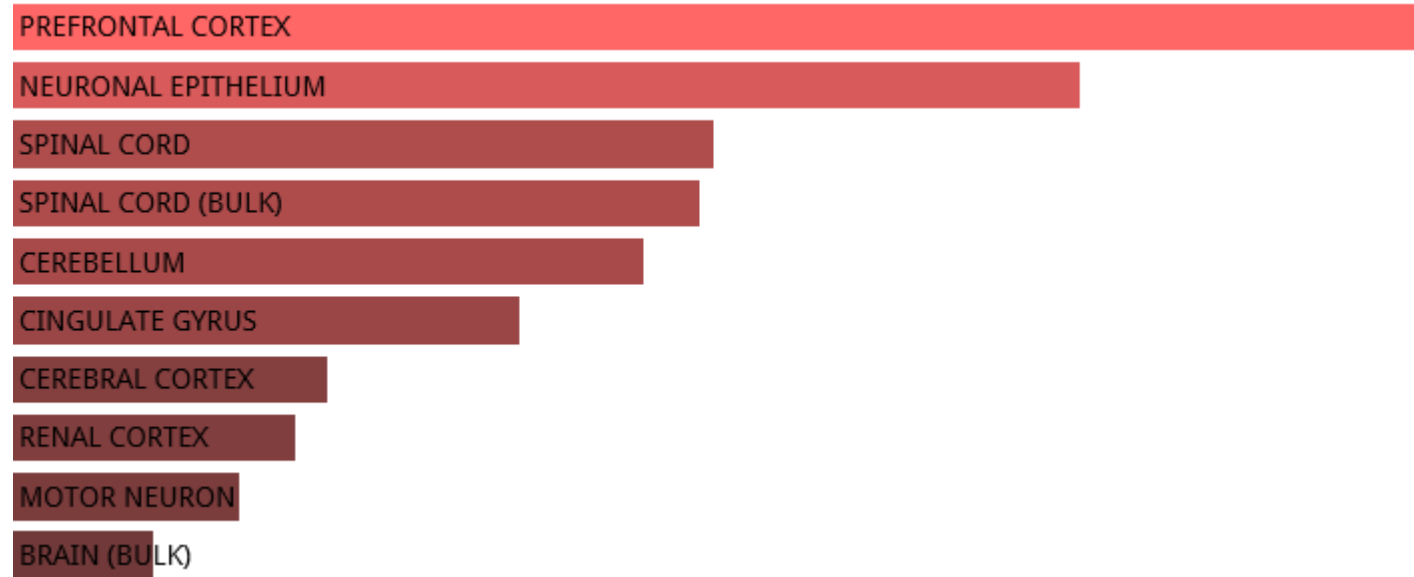
Legend: Definitions of structurally, functionally, and evolutionary distinct families of human-specific regulatory sequences (HSRS) can be found in Glinsky (2020);

**Association of human cancer survival predictor genes with human-specific genomic regulatory sequences (HSGRS)**

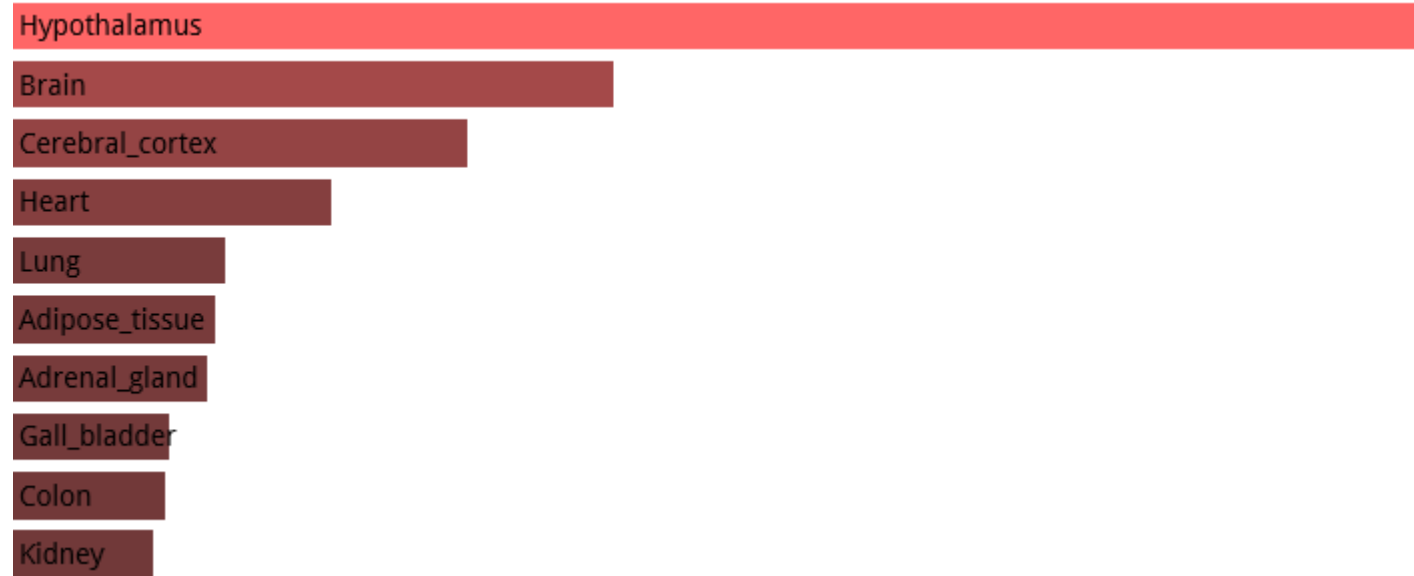
<b>Cancer type</b>	<b>All genes</b>	<b>Associated with HSGRS</b>	<b>Percent</b>
<b>Thyroid</b>	<b>347</b>	<b>269</b>	<b>77.52</b>
<b>Glioma</b>	<b>271</b>	<b>206</b>	<b>76.01</b>
<b>Melanoma</b>	<b>205</b>	<b>153</b>	<b>74.63</b>
<b>Head and neck</b>	<b>808</b>	<b>597</b>	<b>73.89</b>
<b>Colorectal</b>	<b>603</b>	<b>440</b>	<b>72.97</b>
<b>Renal</b>	<b>6070</b>	<b>4418</b>	<b>72.78</b>
<b>Ovarian</b>	<b>504</b>	<b>366</b>	<b>72.62</b>
<b>Liver</b>	<b>2892</b>	<b>2086</b>	<b>72.13</b>
<b>Lung</b>	<b>662</b>	<b>477</b>	<b>72.05</b>
<b>Breast</b>	<b>582</b>	<b>414</b>	<b>71.13</b>
<b>Urothelial</b>	<b>1101</b>	<b>783</b>	<b>71.12</b>
<b>Stomach</b>	<b>307</b>	<b>218</b>	<b>71.01</b>
<b>Prostate</b>	<b>161</b>	<b>114</b>	<b>70.81</b>
<b>Endometrial</b>	<b>1631</b>	<b>1153</b>	<b>70.69</b>
<b>Cervical</b>	<b>717</b>	<b>505</b>	<b>70.43</b>
<b>Pancreatic</b>	<b>1549</b>	<b>1075</b>	<b>69.40</b>
<b>All human cancer survival genes</b>	<b>10713</b>	<b>7738</b>	<b>72.23</b>

**11,878 fixed human-specific insertions**

ARCHS4 Tissues 7979 genes



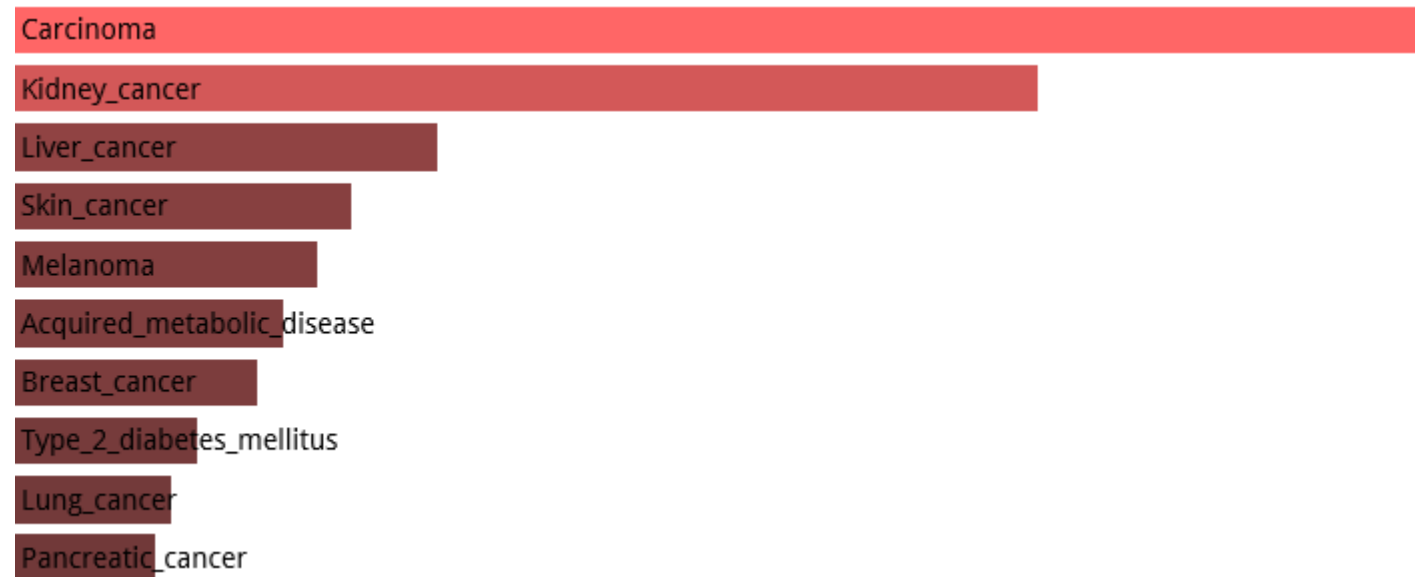
Jensen TISSUES 7979 genes



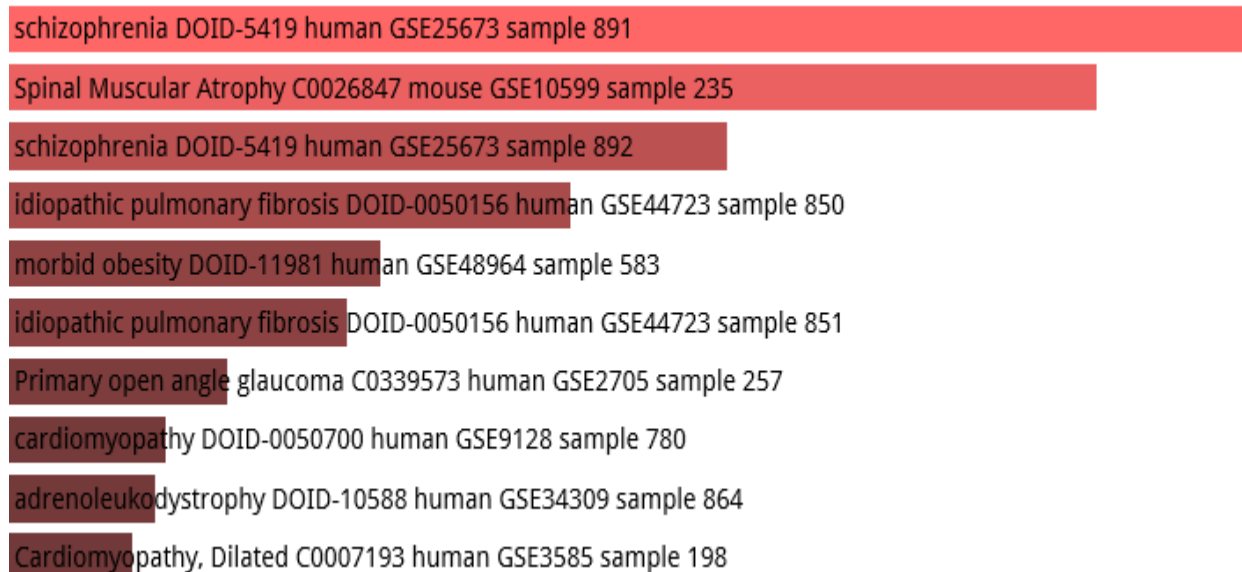


# 11,878 fixed human-specific insertions

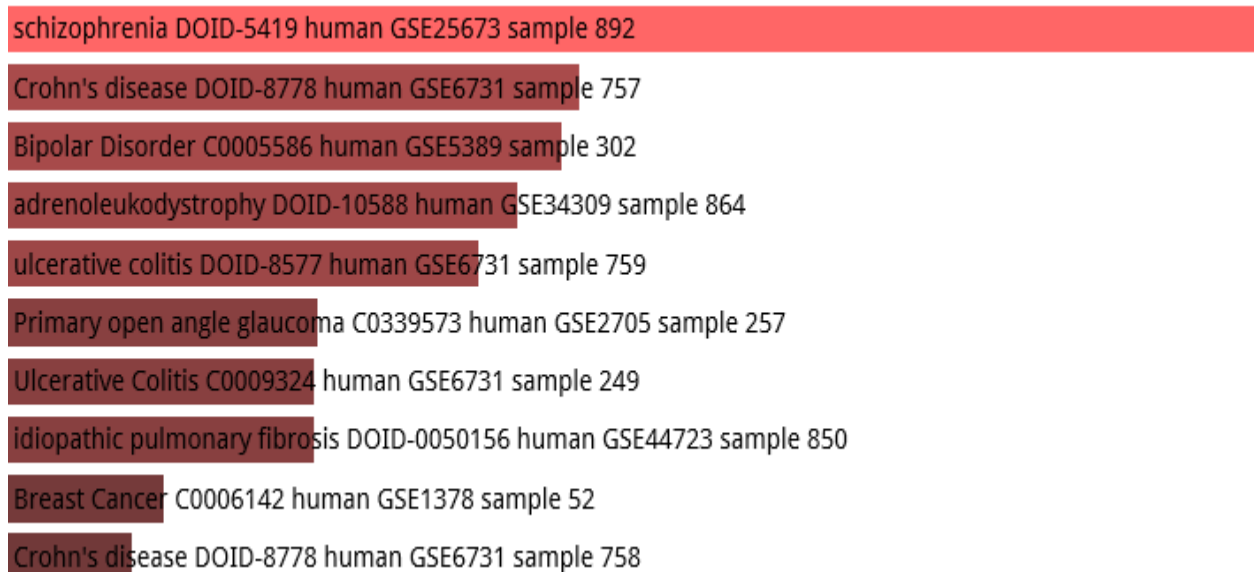
## Jensen DISEASES 7979 genes



## Disease Perturbations from GEO up



## Disease Perturbations from GEO down

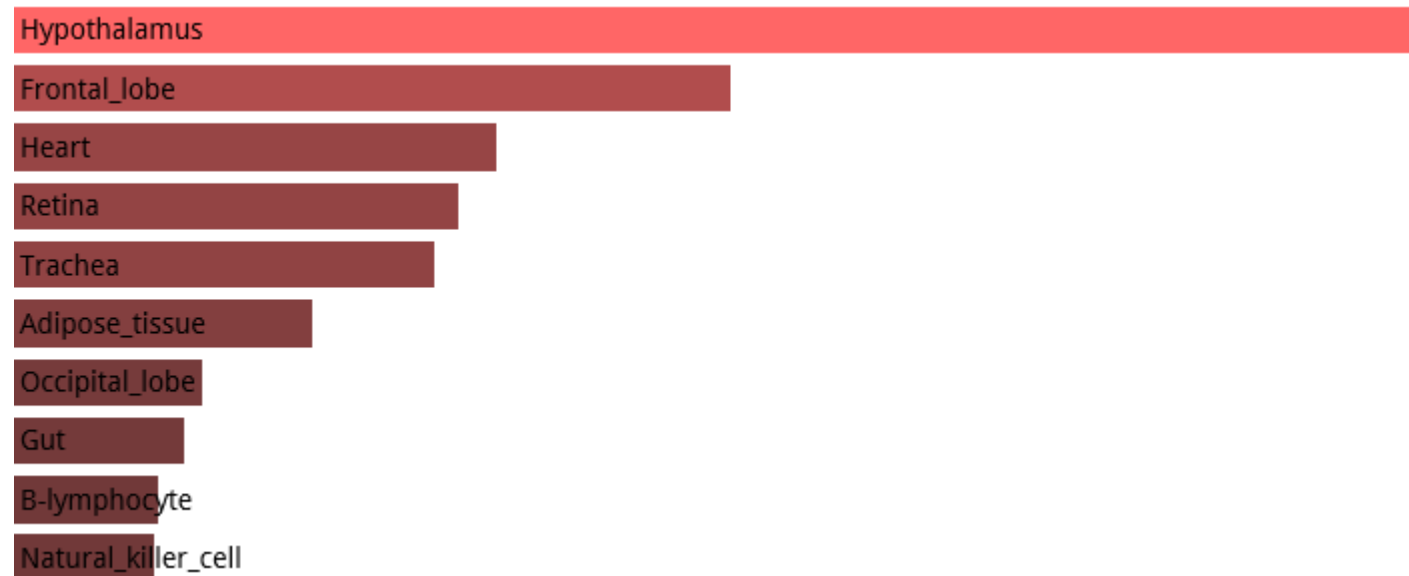


7,599 duplicated regions in the human genome

ARCHS4 Tissues 6618 genes

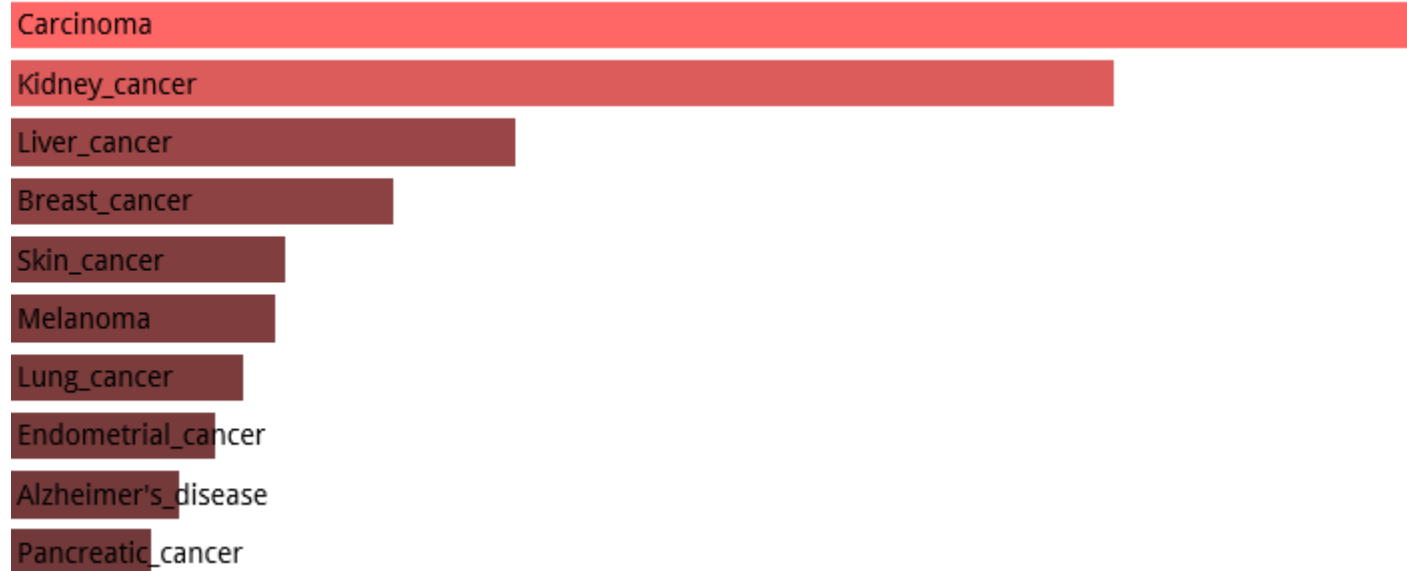


Jensen TISSUES 6618 genes

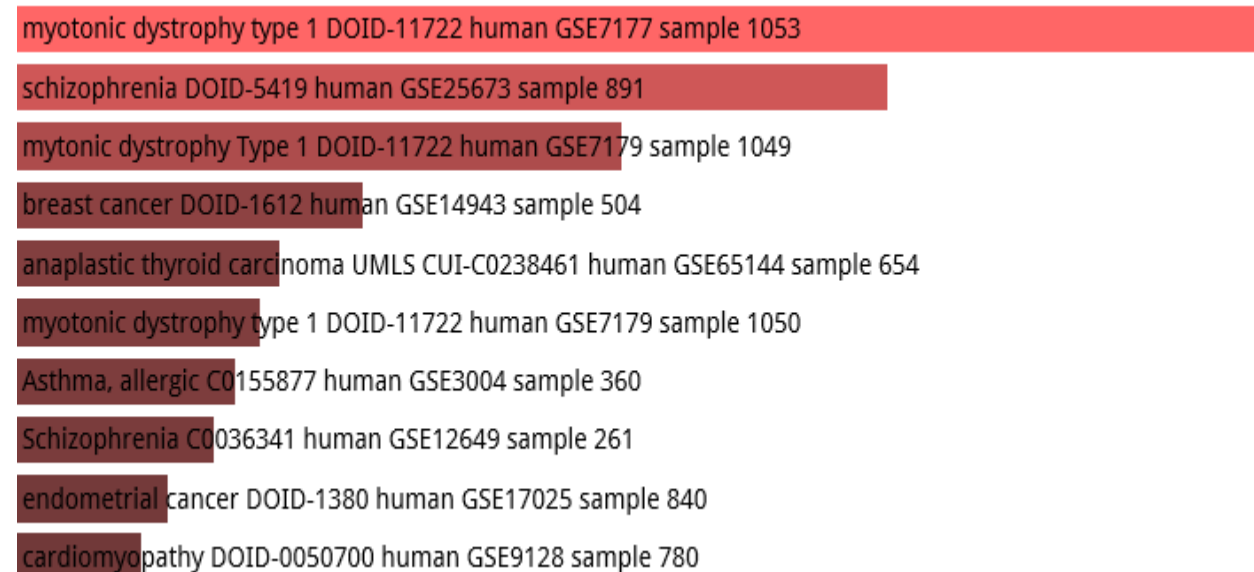


7,599 duplicated regions in the human genome

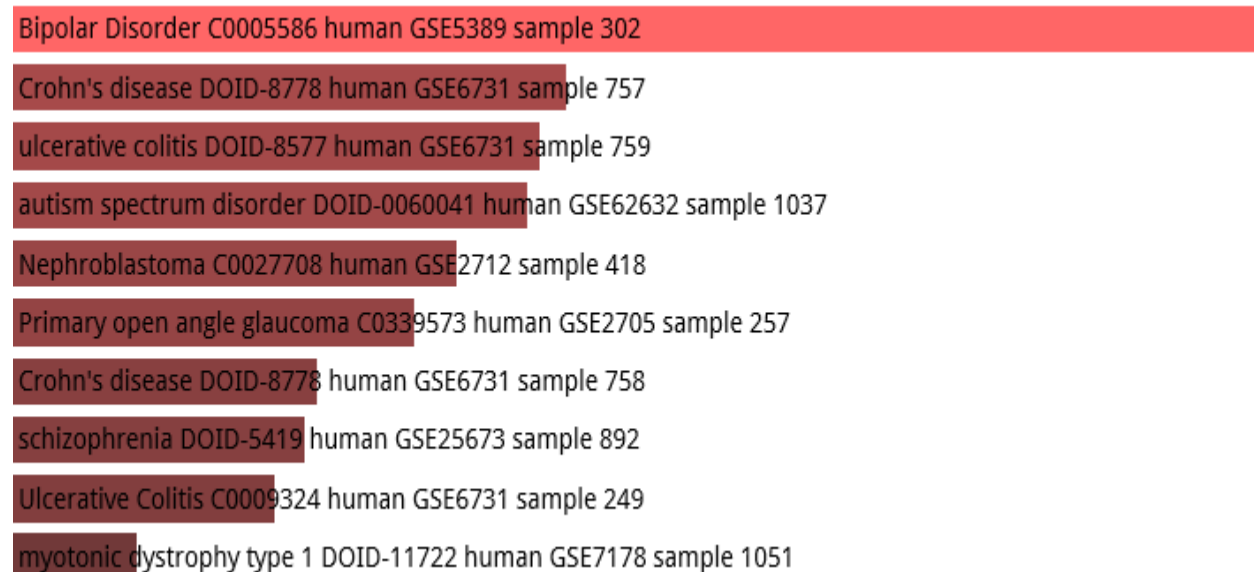
Jensen DISEASES 6618 genes



Disease Perturbations from GEO up

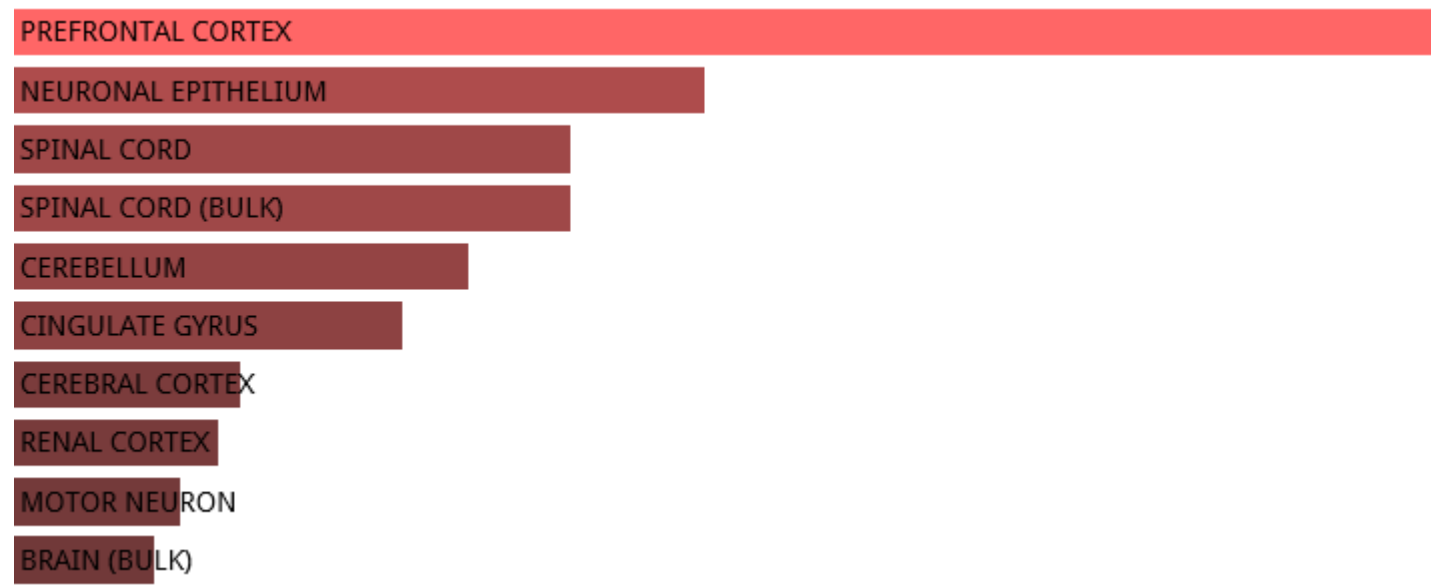


Disease Perturbations from GEO down

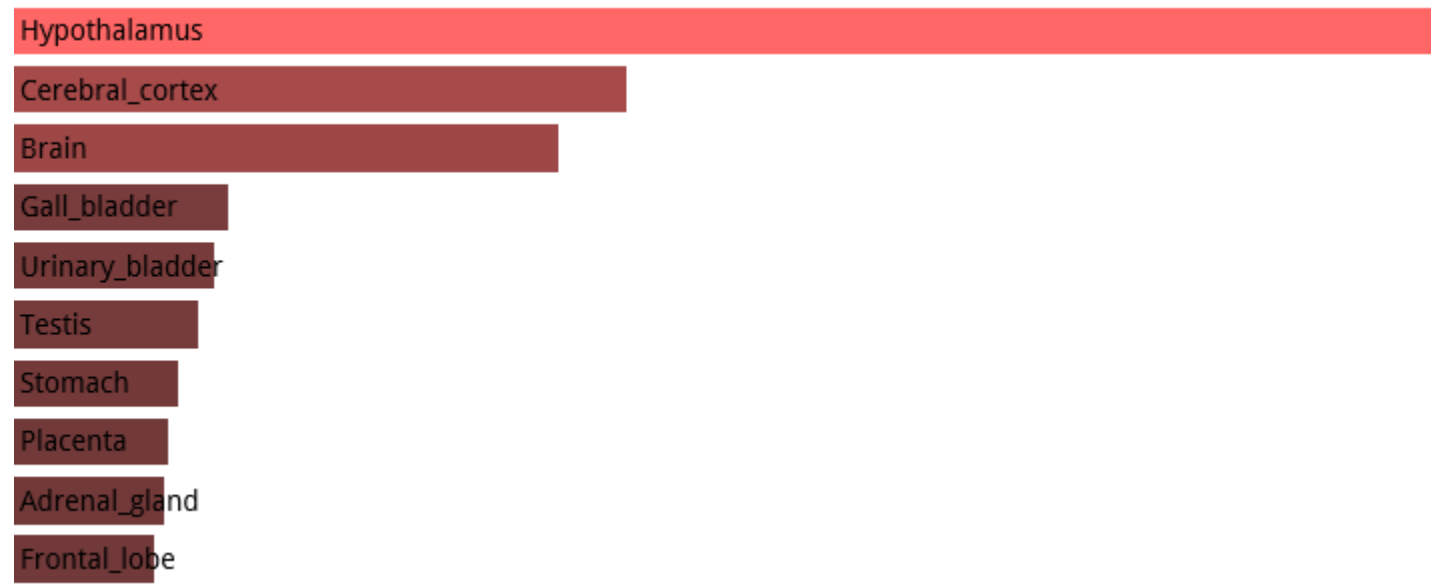


5,883 fixed human-specific deletions

ARCHS4 Tissues 5489 genes



Jensen TISSUES 5489 genes



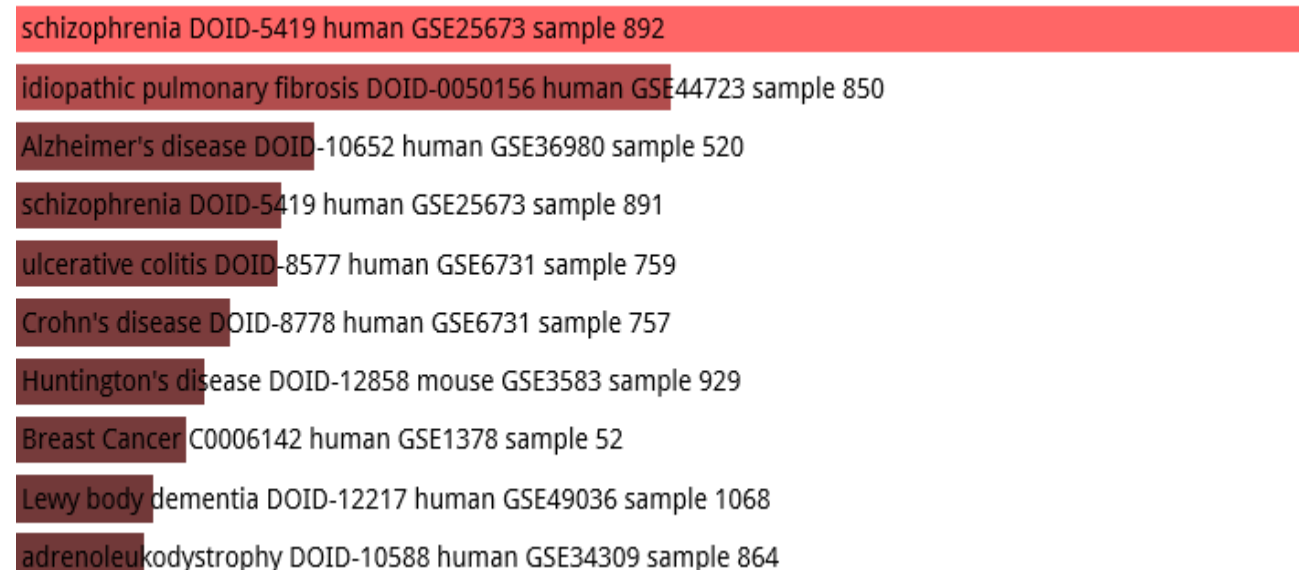
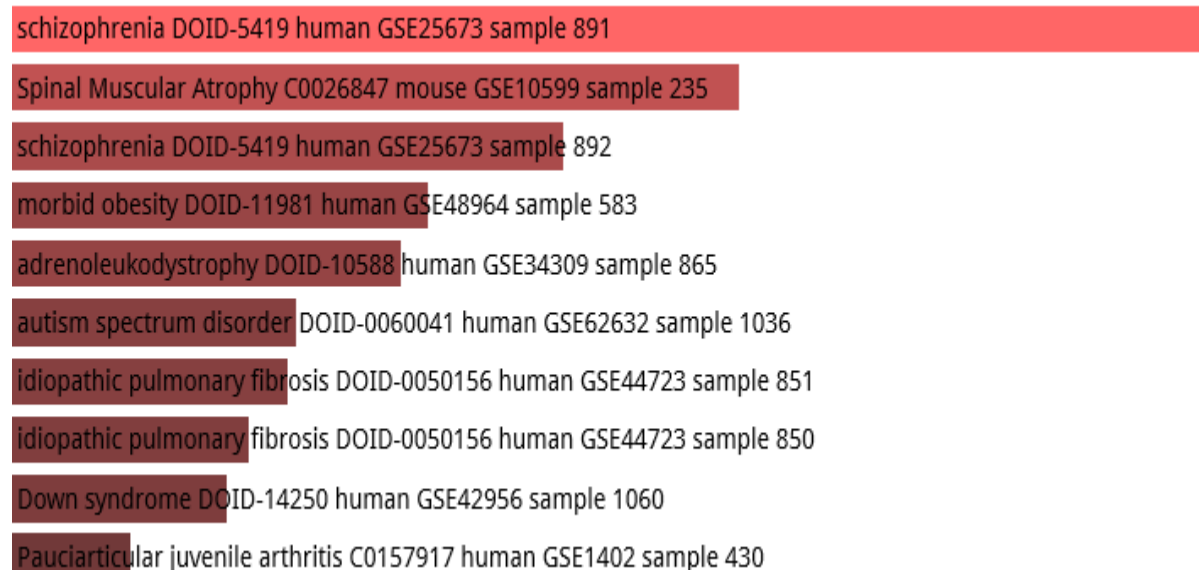
## 5,883 fixed human-specific deletions

## Jensen DISEASES 5489 genes



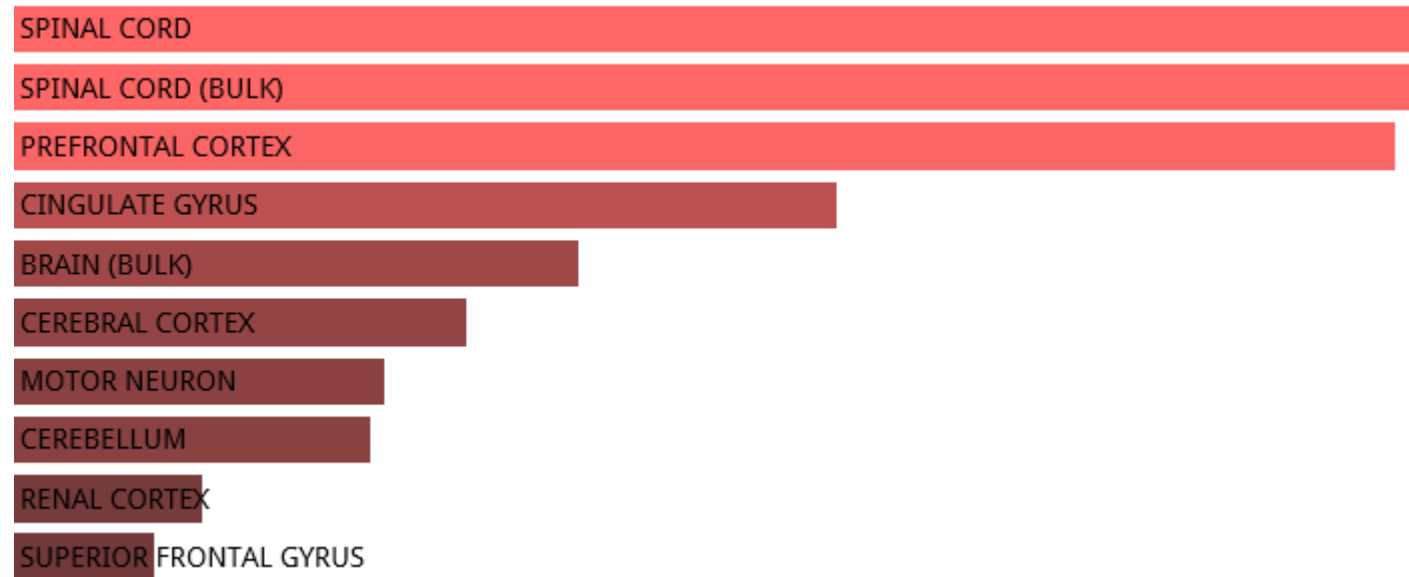
## Disease Perturbations from GEO up

## Disease Perturbations from GEO down

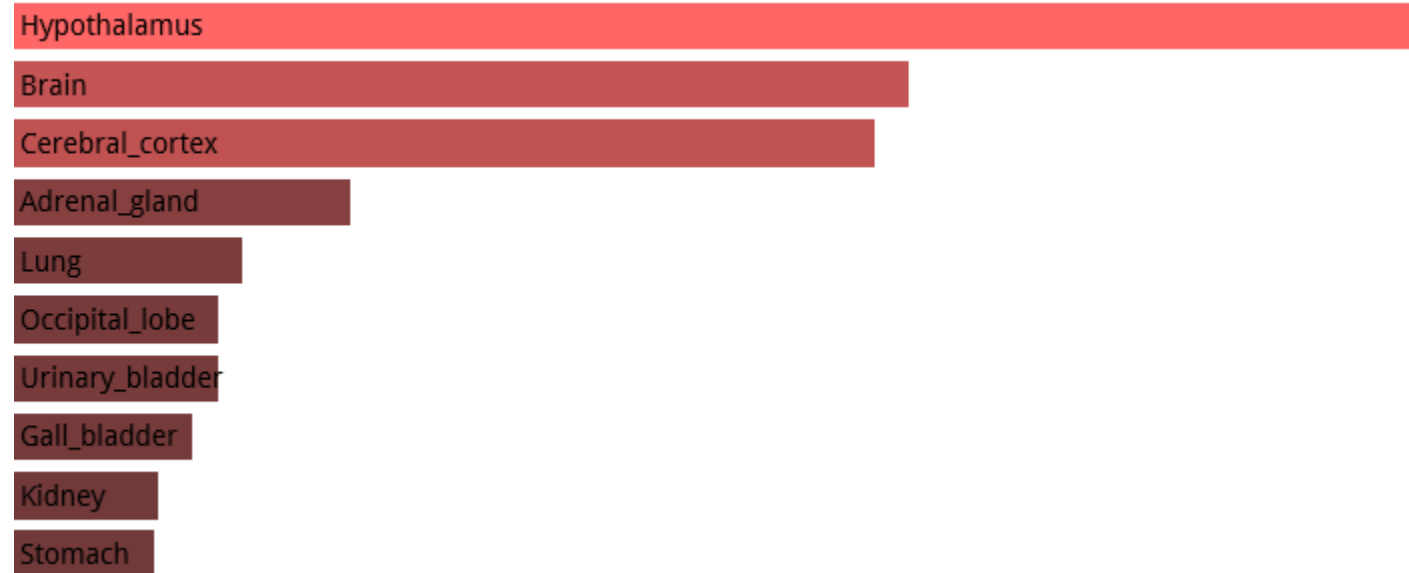


## 4,875 human-specific STR expansions

## ARCHS4 Tissues 4844 genes

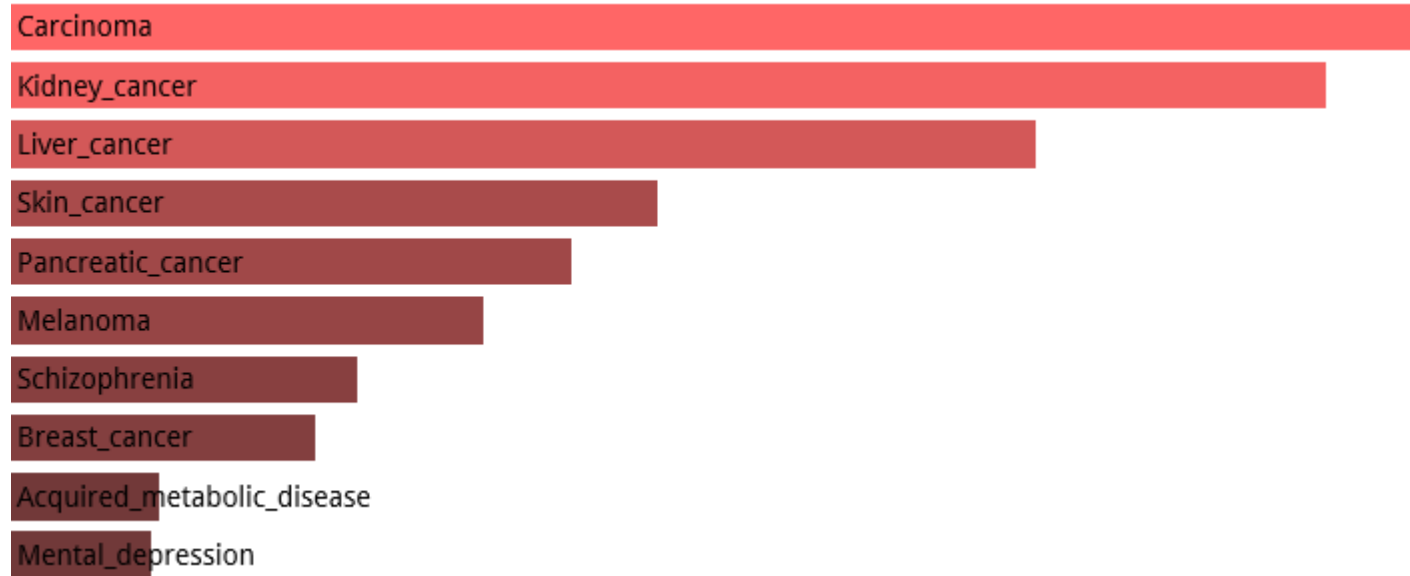


## Jensen TISSUES 4844 genes



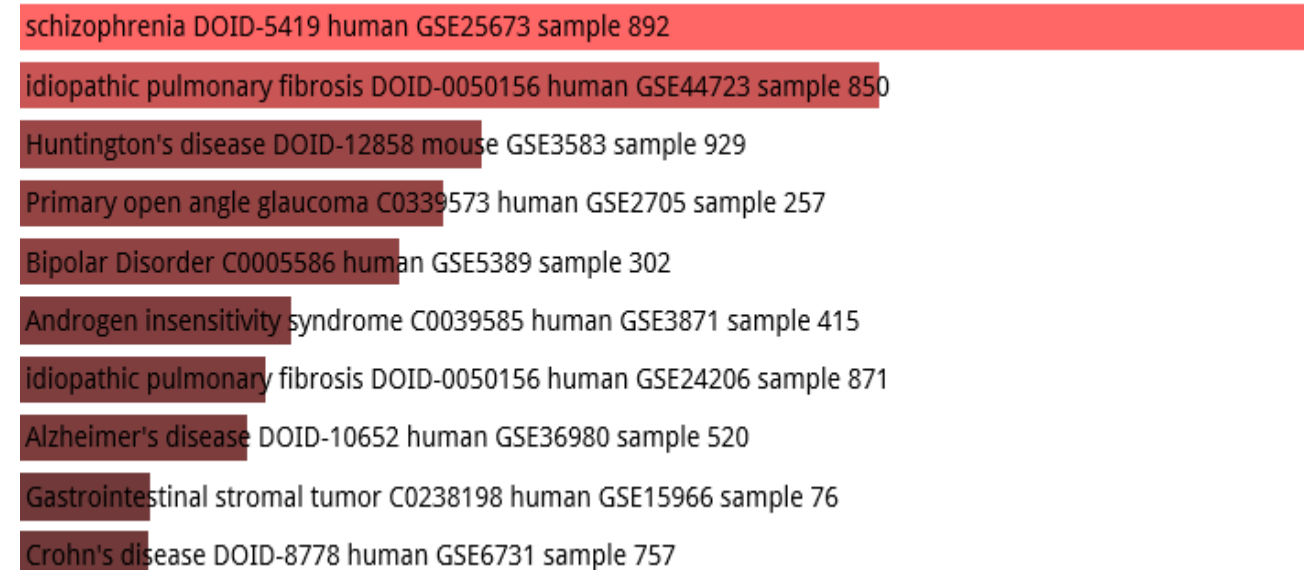
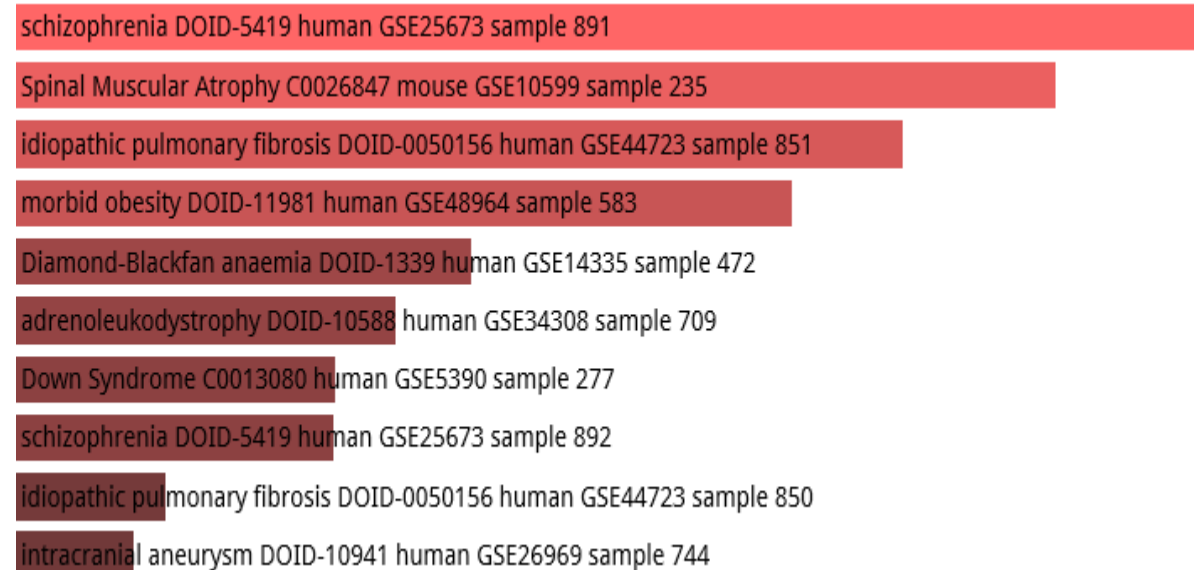
## 4,875 human-specific STR expansions

## Jensen DISEASES 4844 genes



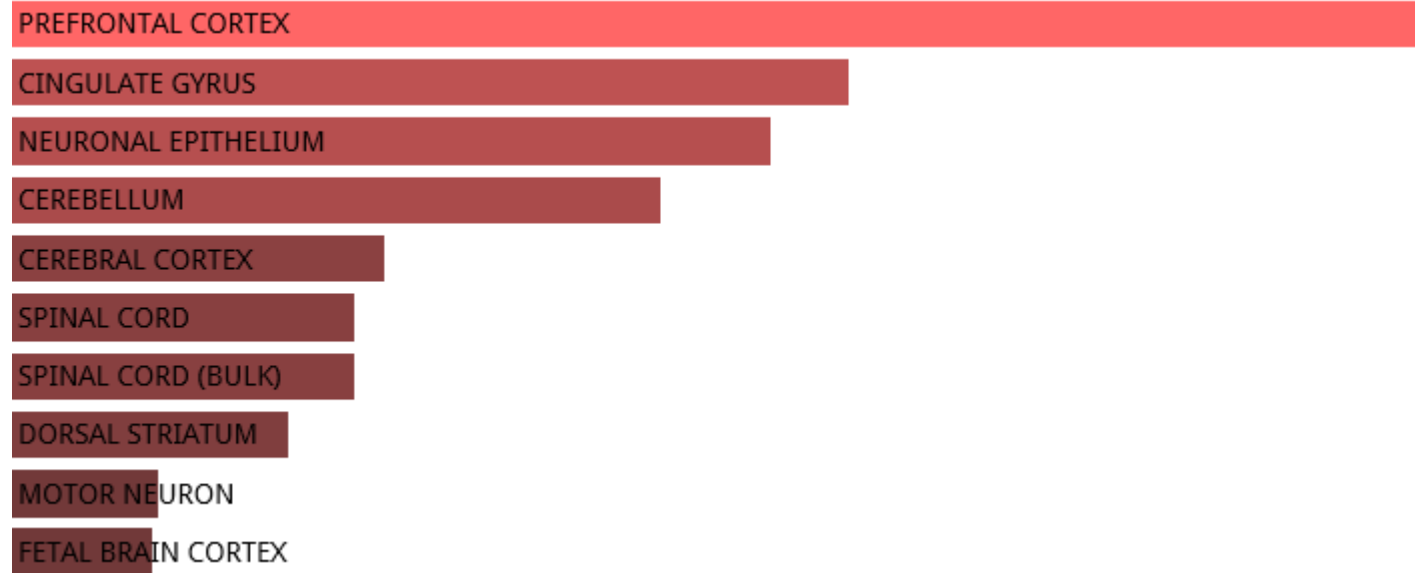
## Disease Perturbations from GEO up

## Disease Perturbations from GEO down

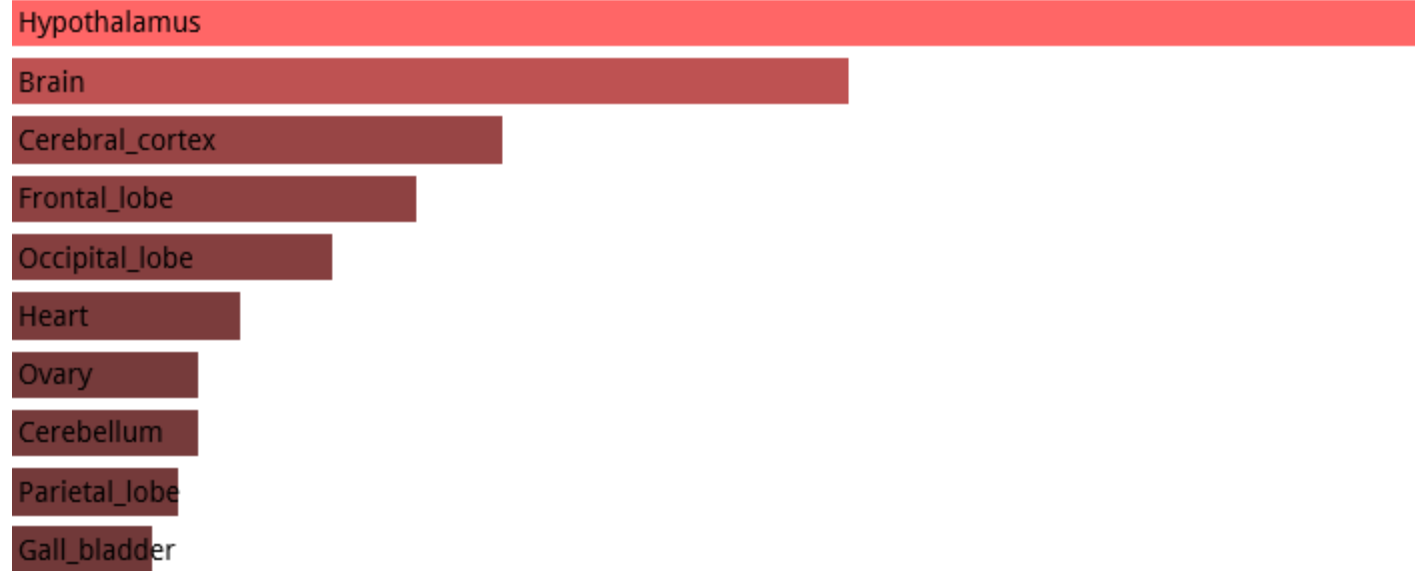


**4,637 human-specific TE-encoded loci  
expressed in human DLPFC**

ARCHS4 Tissues 4051 genes



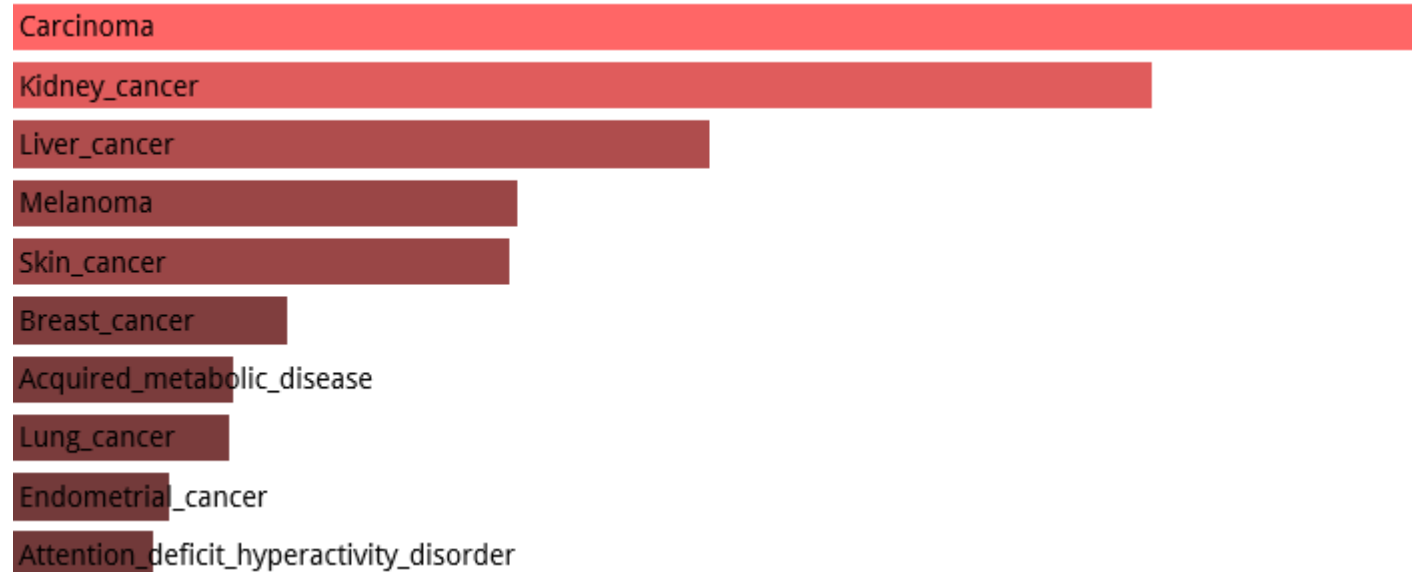
Jensen TISSUES 4051 genes





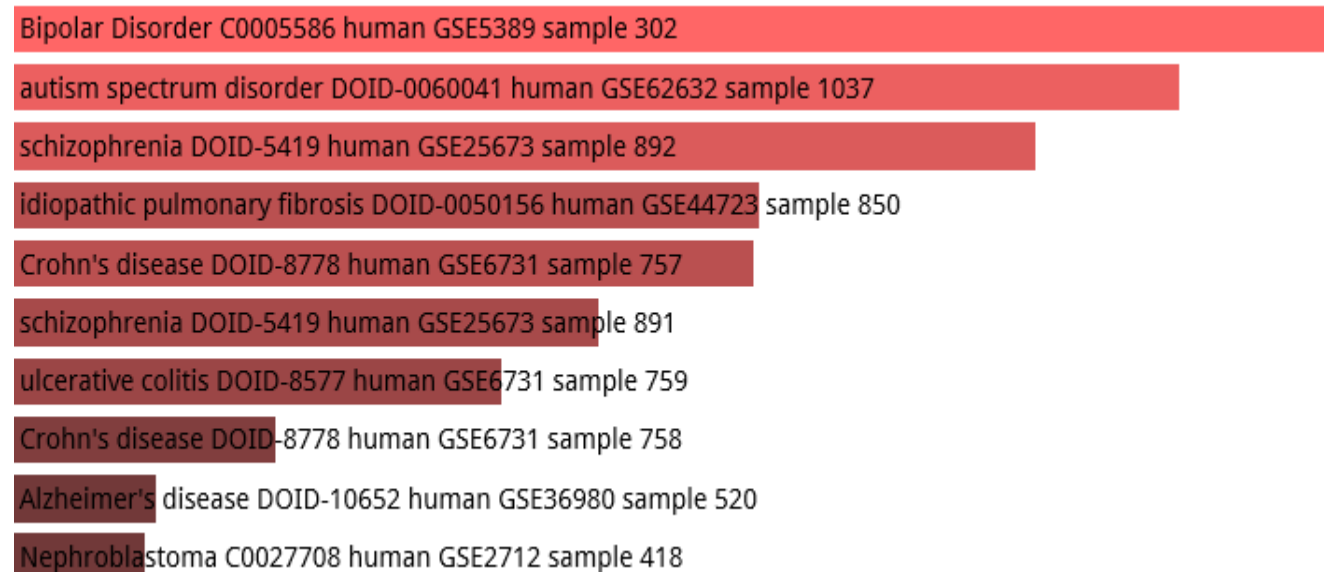
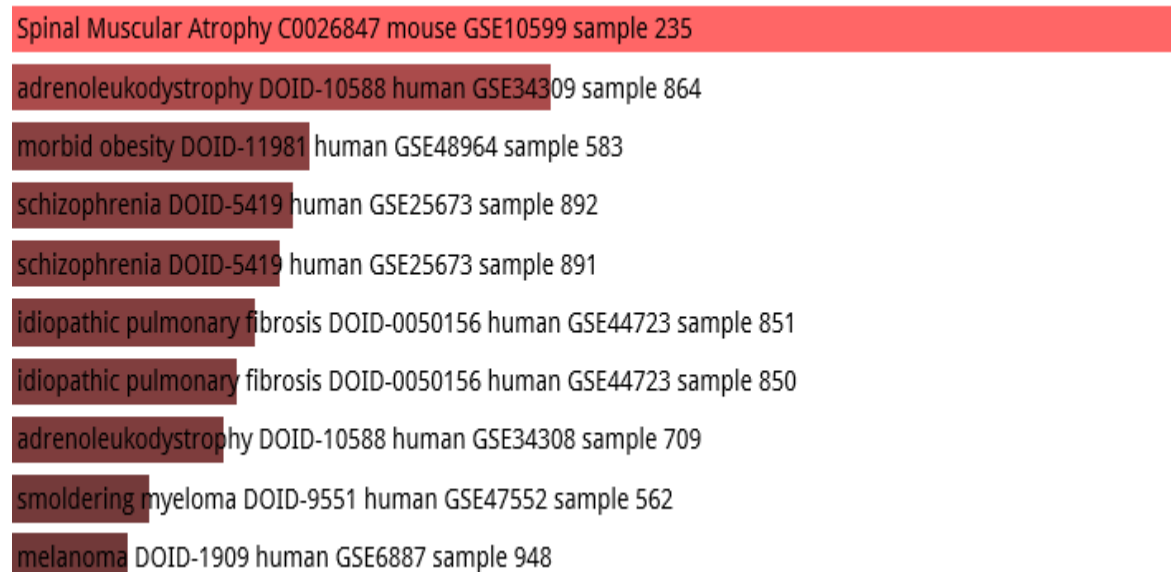
## 4,637 human-specific TE-encoded loci expressed in human DLPFC

## Jensen DISEASES 4051 genes



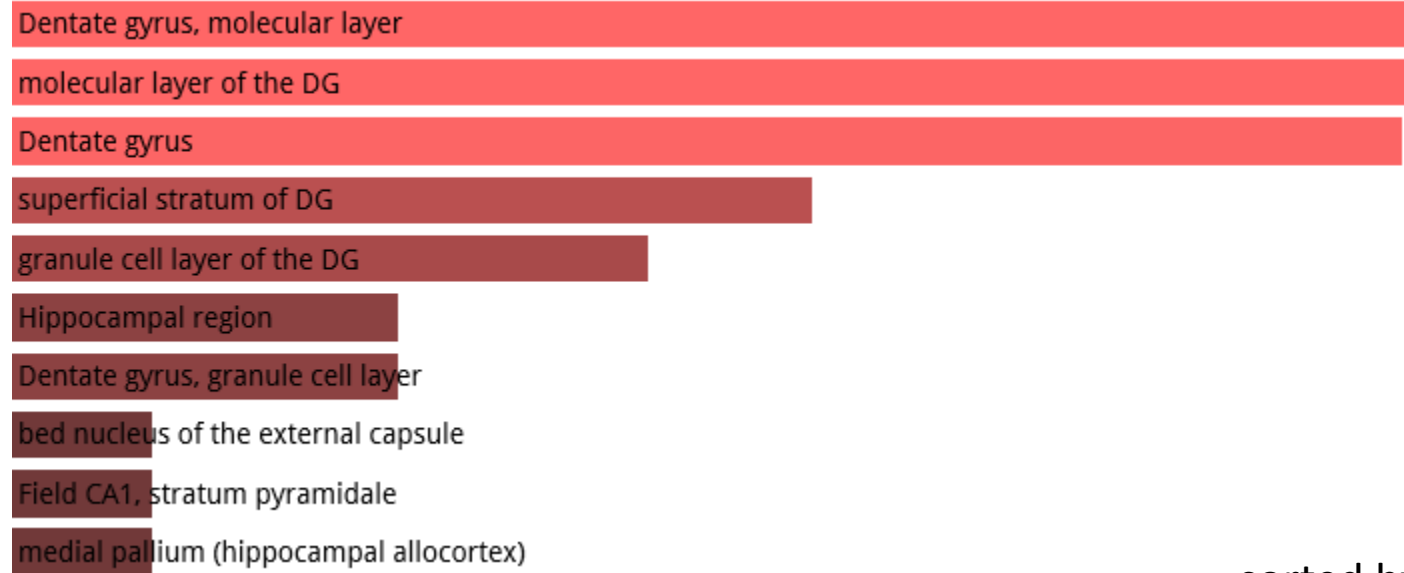
## Disease Perturbations from GEO up

## Disease Perturbations from GEO down



**4,637 human-specific TE-encoded loci  
expressed in human DLPFC**

**Allen Brain Atlas up 4051 genes**



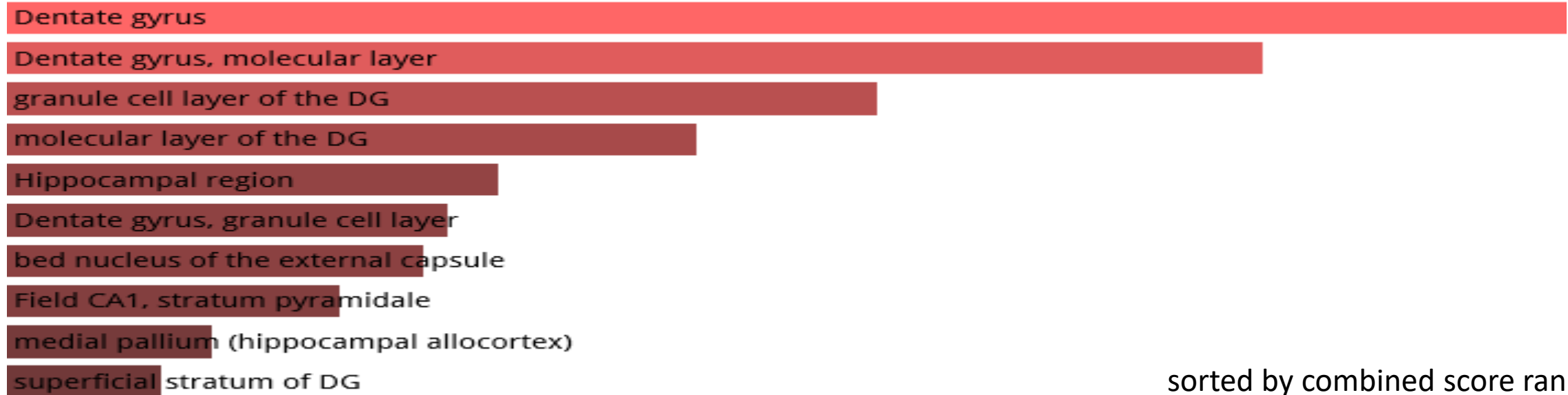
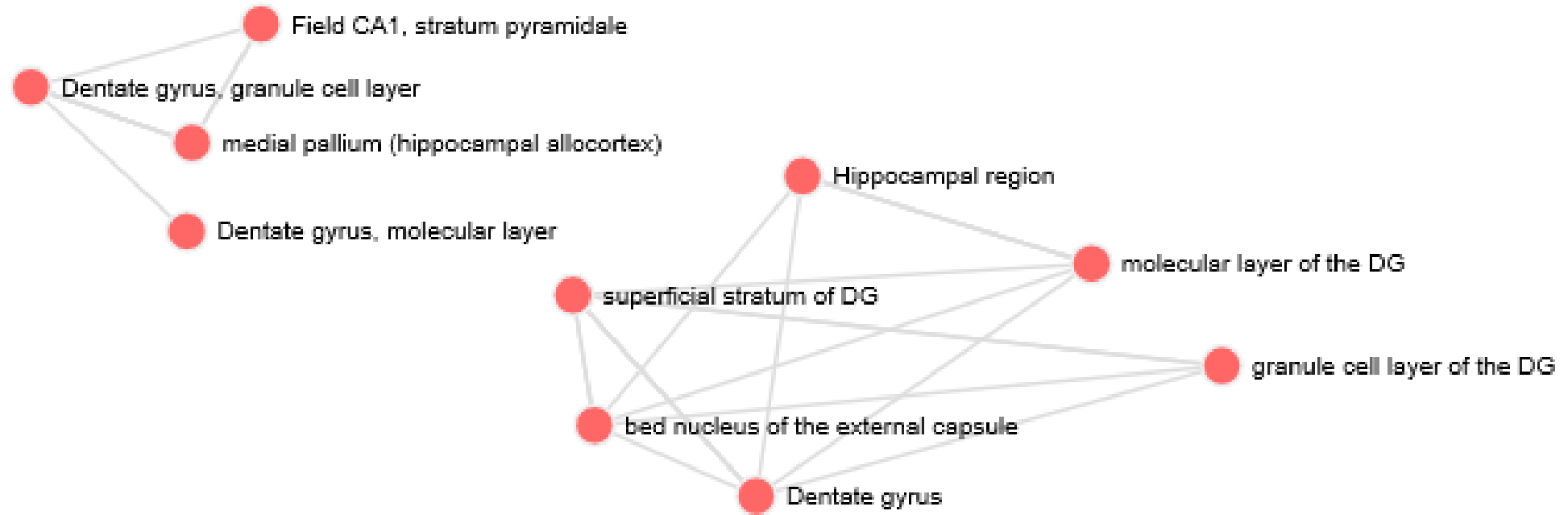
sorted by p-value ranking

### TOP 15 CATEGORIES

Term	Overlap	P-value	Adjusted P-value
Dentate gyrus	110/320	2.07E-09	1.51E-06
Dentate gyrus, molecular layer	105/301	1.92E-09	1.51E-06
molecular layer of the DG	105/301	1.92E-09	1.51E-06
superficial stratum of DG	179/601	1.14E-08	6.24E-06
granule cell layer of the DG	102/301	1.83E-08	8.03E-06
Hippocampal region	101/301	3.78E-08	1.18E-05
Dentate gyrus, granule cell layer	101/301	3.78E-08	1.18E-05
bed nucleus of the external capsule	100/301	7.67E-08	1.68E-05
Field CA1, stratum pyramidale	100/301	7.67E-08	1.68E-05
medial pallium (hippocampal allocortex)	100/301	7.67E-08	1.68E-05
mantle zone of DG	174/599	1.22E-07	2.4E-05
Subiculum, dorsal part, molecular layer	99/301	1.54E-07	2.4E-05
hilus of the DG	99/301	1.54E-07	2.4E-05
Dentate gyrus, polymorph layer	99/301	1.54E-07	2.4E-05
hippocampus (cortex Ammonis)	98/301	3.03E-07	2.88E-05

# 4,637 human-specific TE-encoded loci expressed in human DLPFC

Allen Brain Atlas up 4051 genes



sorted by combined score ranking

# 4,637 human-specific TE-encoded loci expressed in human DLPFC

Tissue Protein Expression from Human Proteome Map: 4051 genes

adult frontal cortex p-value: 0.00006954 Adjusted p-value: 0.002086

fetal brain p-value: 0.001310; Adjusted p-value: 0.01966

adult retina

b cells

adult pancreas

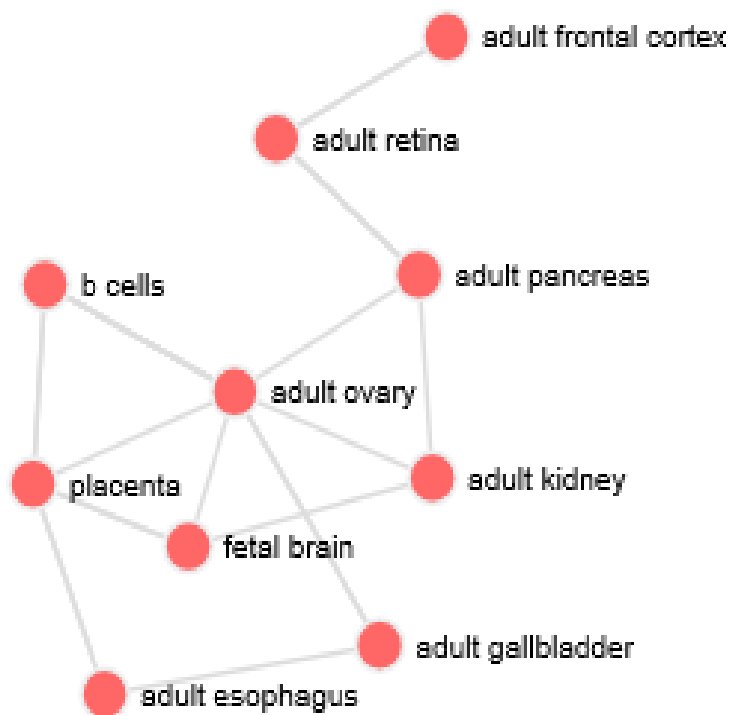
placenta

adult kidney

adult esophagus

adult gallbladder

adult ovary

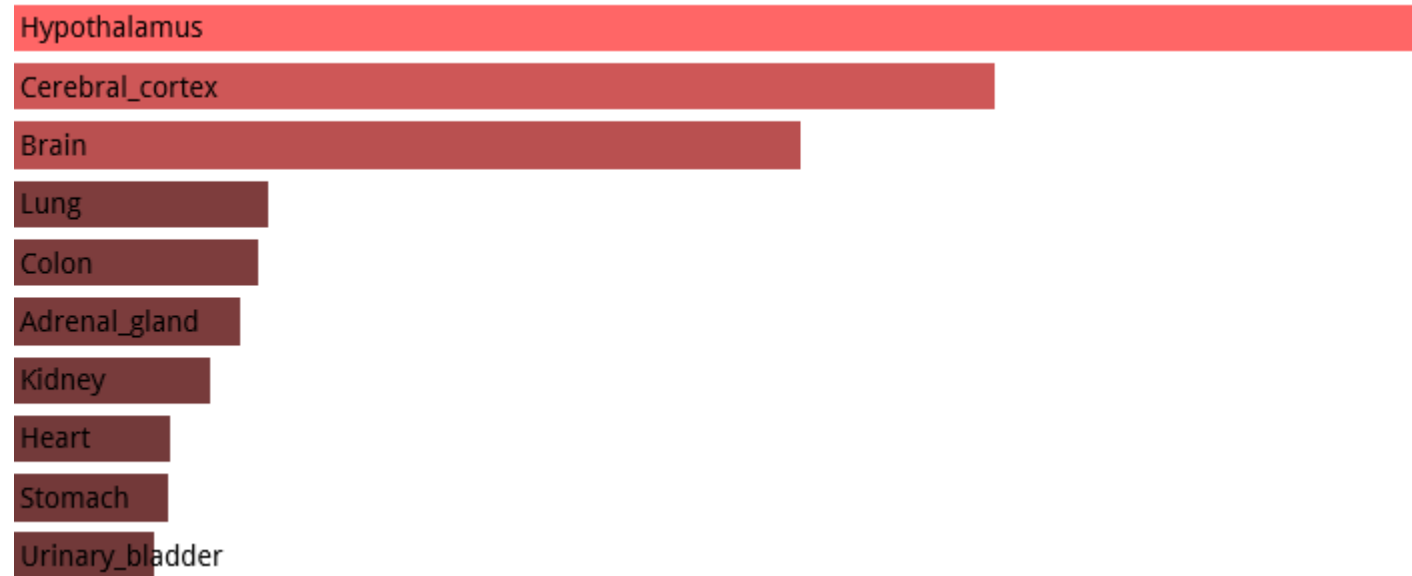


3,538 human-specific ace-DHS

ARCHS4 Tissues 3445 genes



Jensen TISSUES 3445 genes

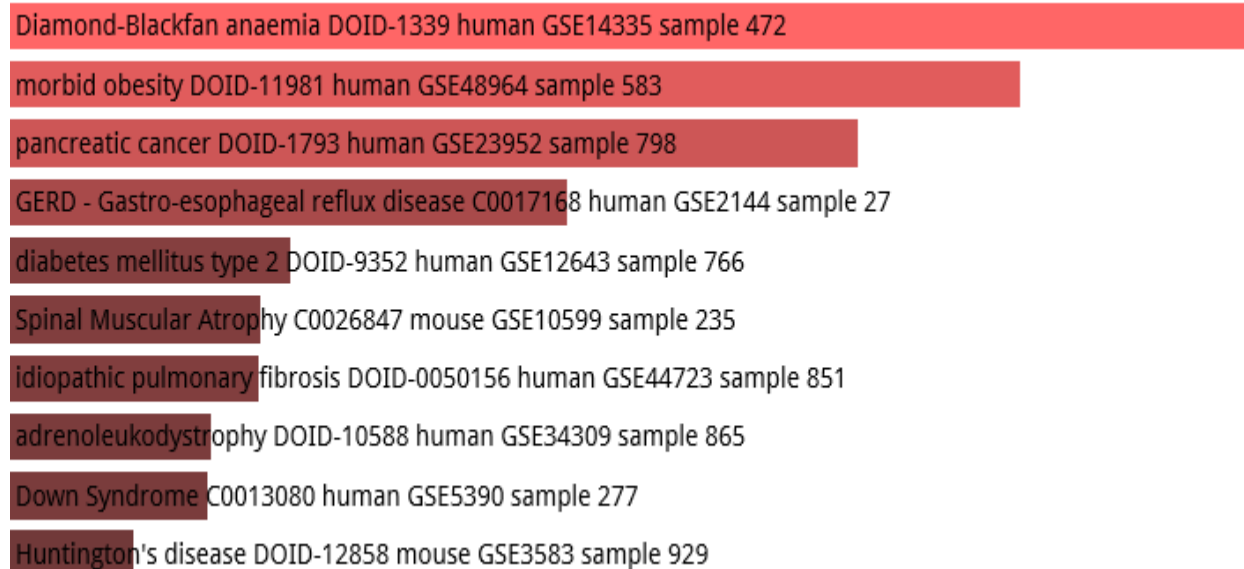


## 3,538 human-specific ace-DHS

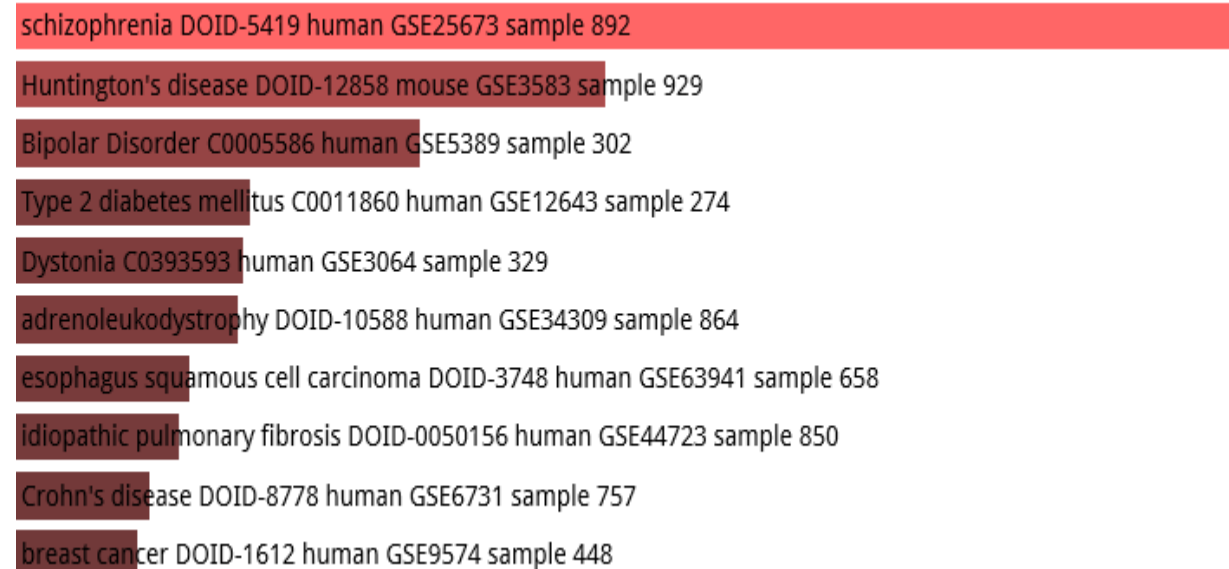
## Jensen DISEASES 3445 genes



## Disease Perturbations from GEO up

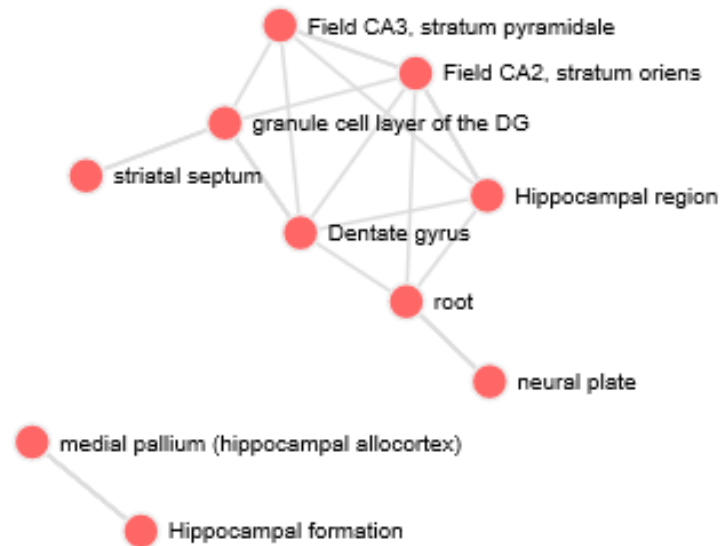
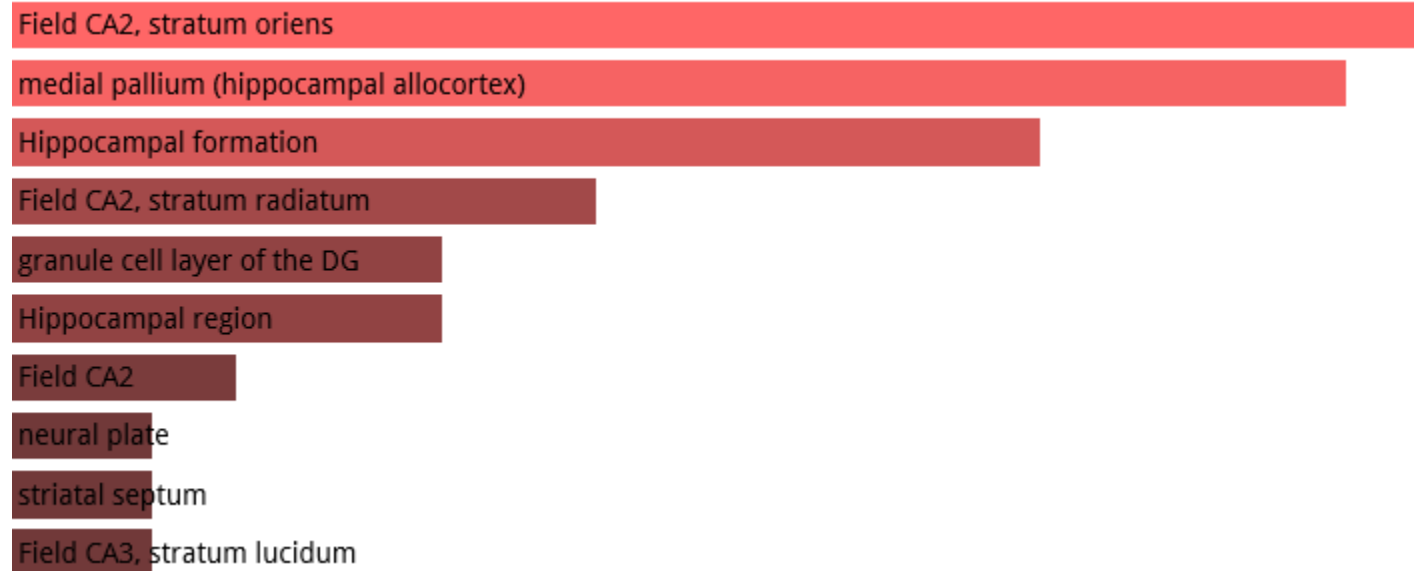


## Disease Perturbations from GEO down



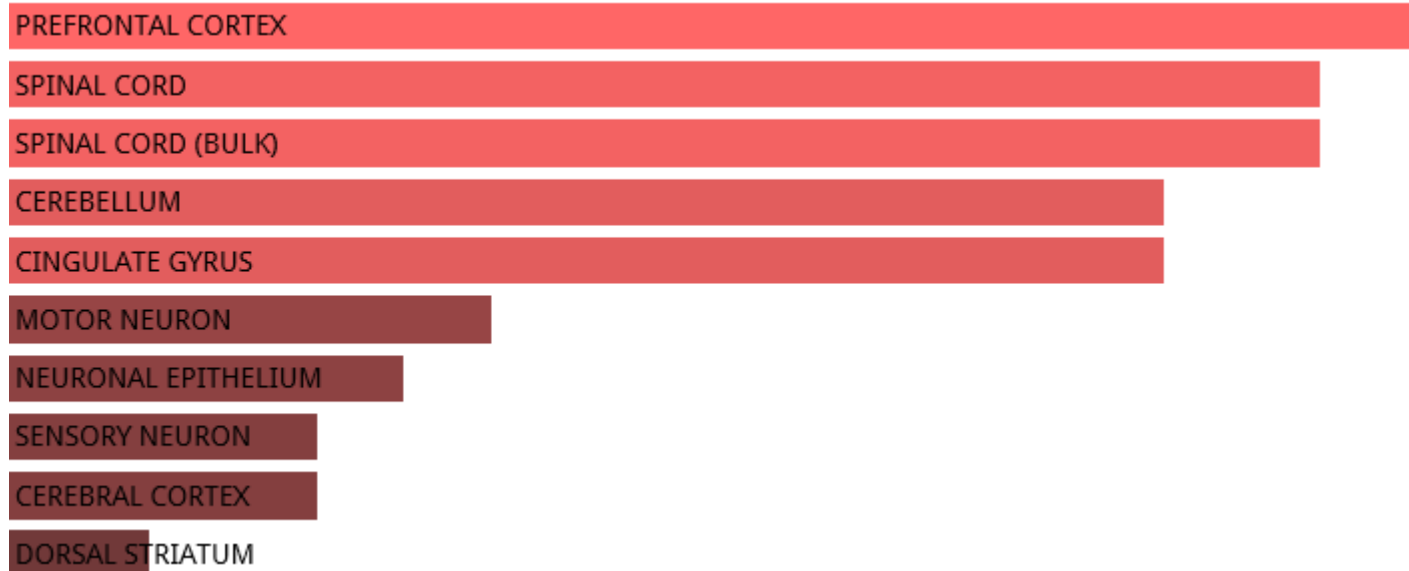
# 3,538 human-specific ace-DHS

## Allen Brain Atlas up 3445 genes

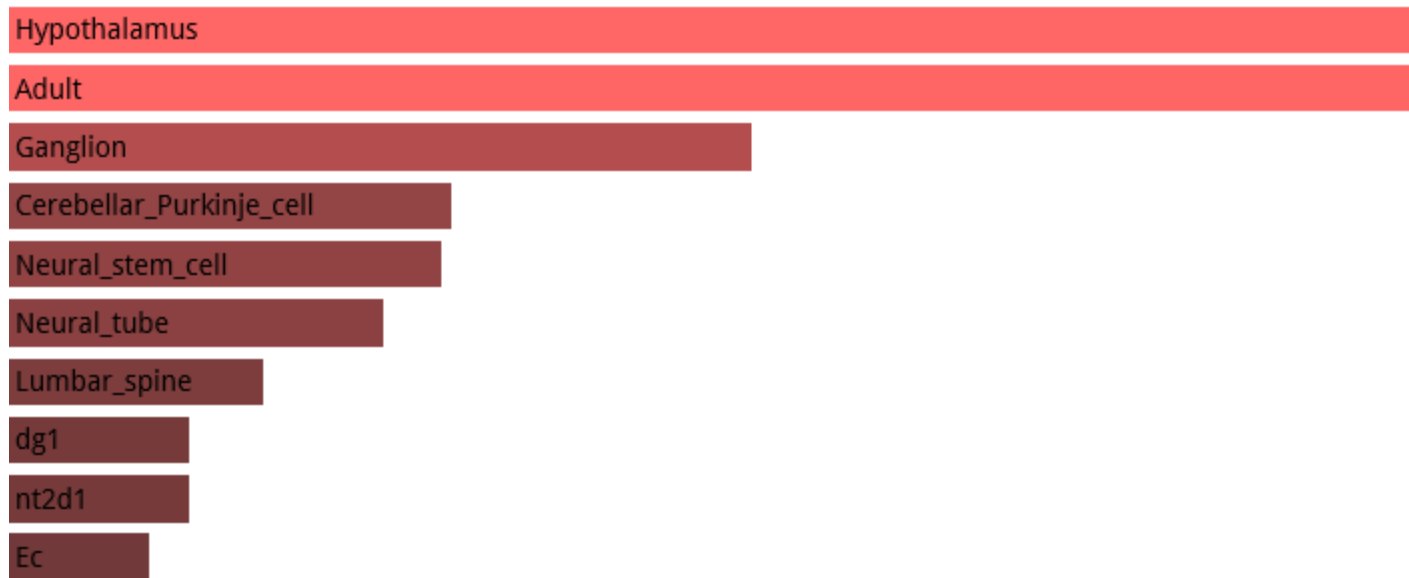


3,803 human-specific TFBS in hESC

ARCHS4 Tissues 1087 genes



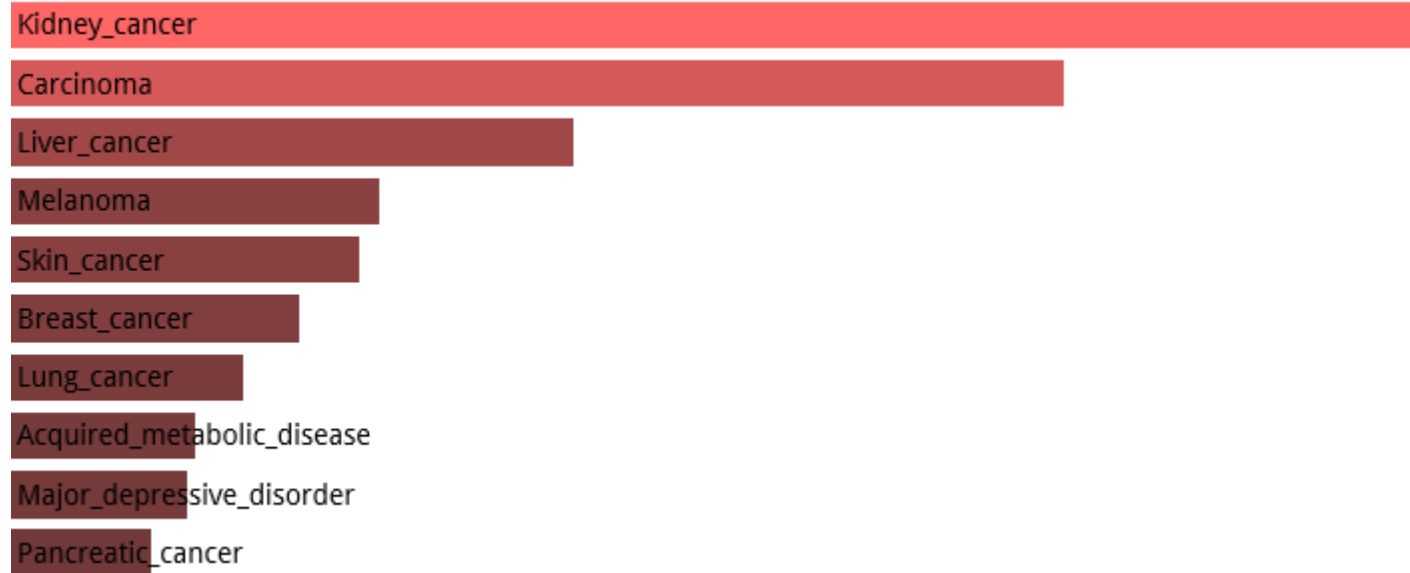
Jensen TISSUES 1087 genes





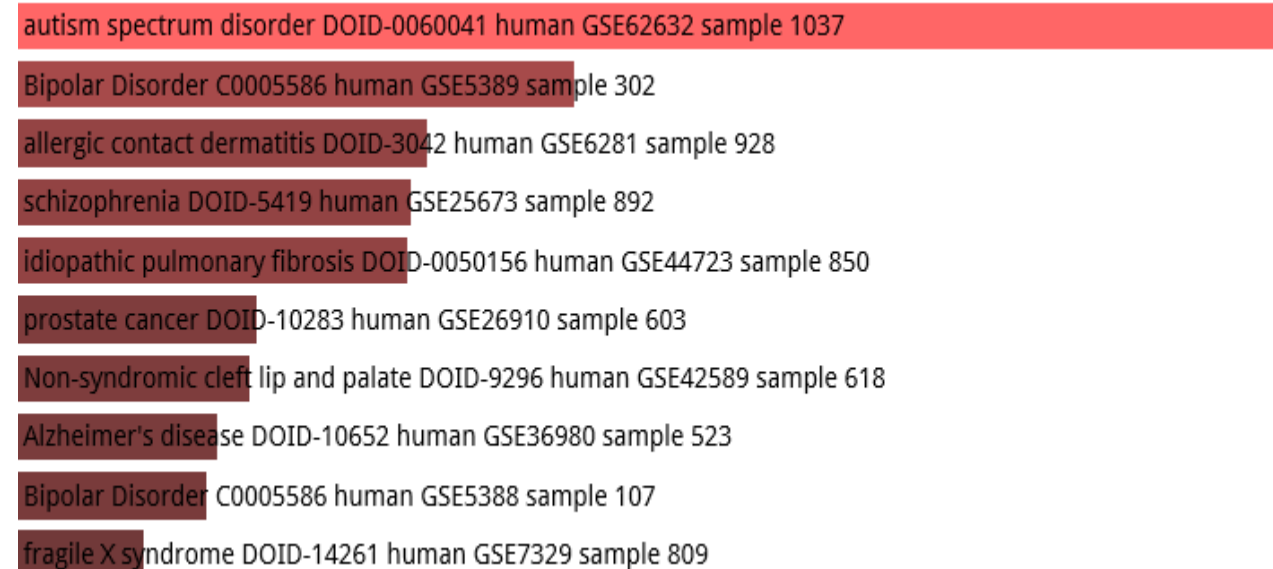
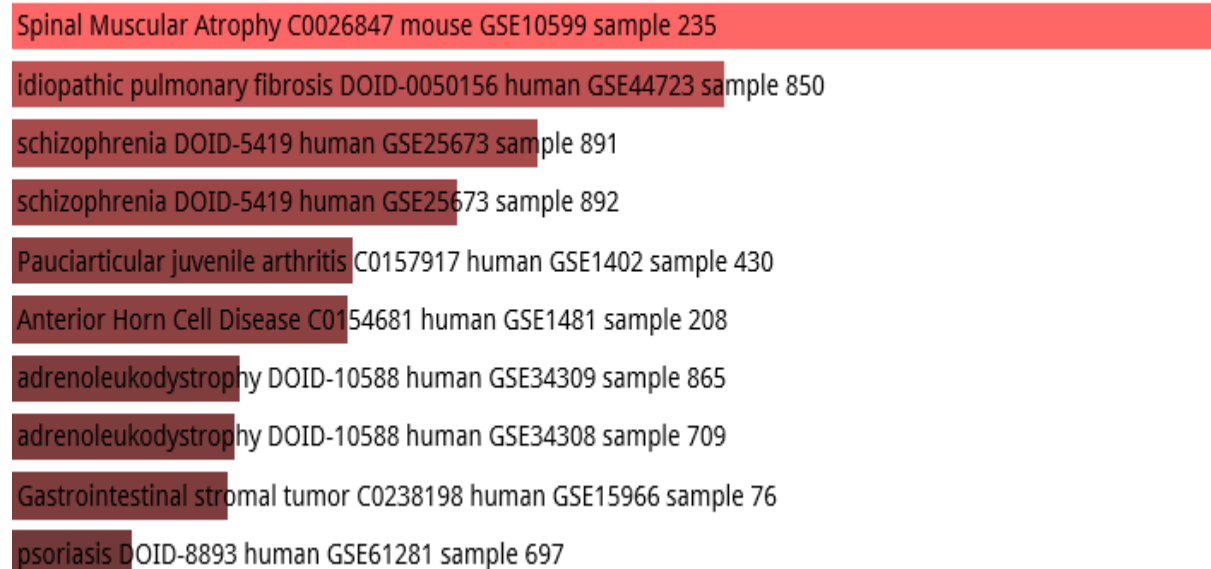
## 3,803 human-specific TFBS in hESC

## Jensen DISEASES 1087 genes



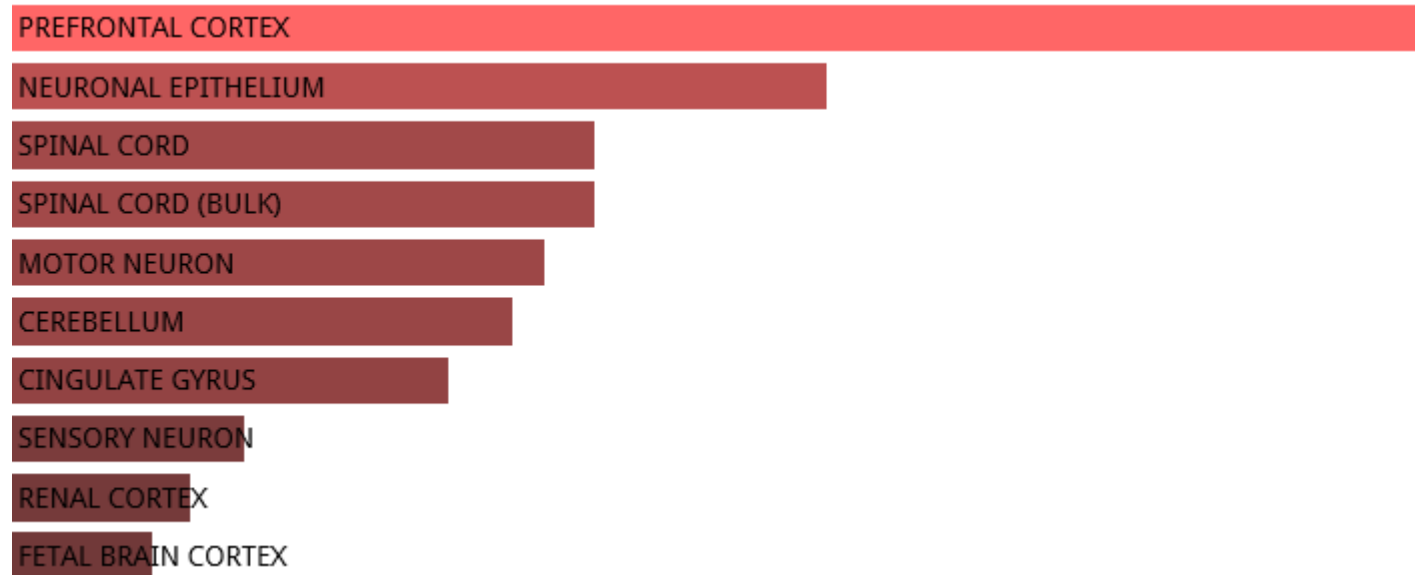
## Disease Perturbations from GEO up

## Disease Perturbations from GEO down

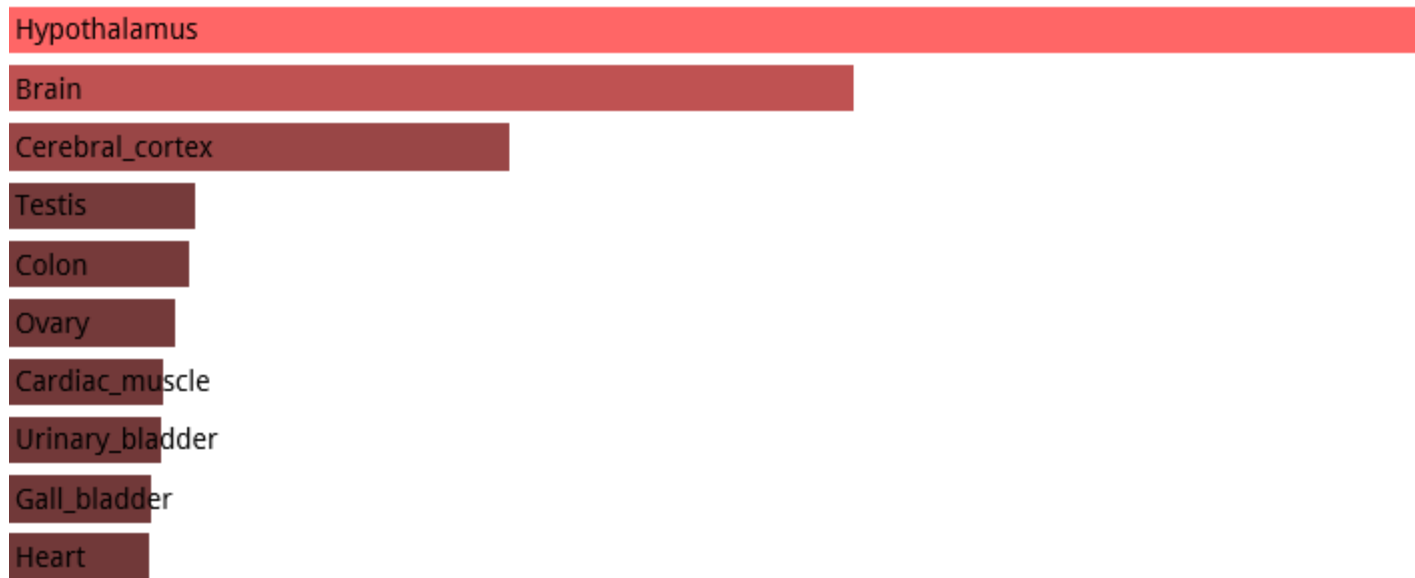


4,249 fixed human-specific regulatory loci

ARCHS4 Tissues 2810 genes

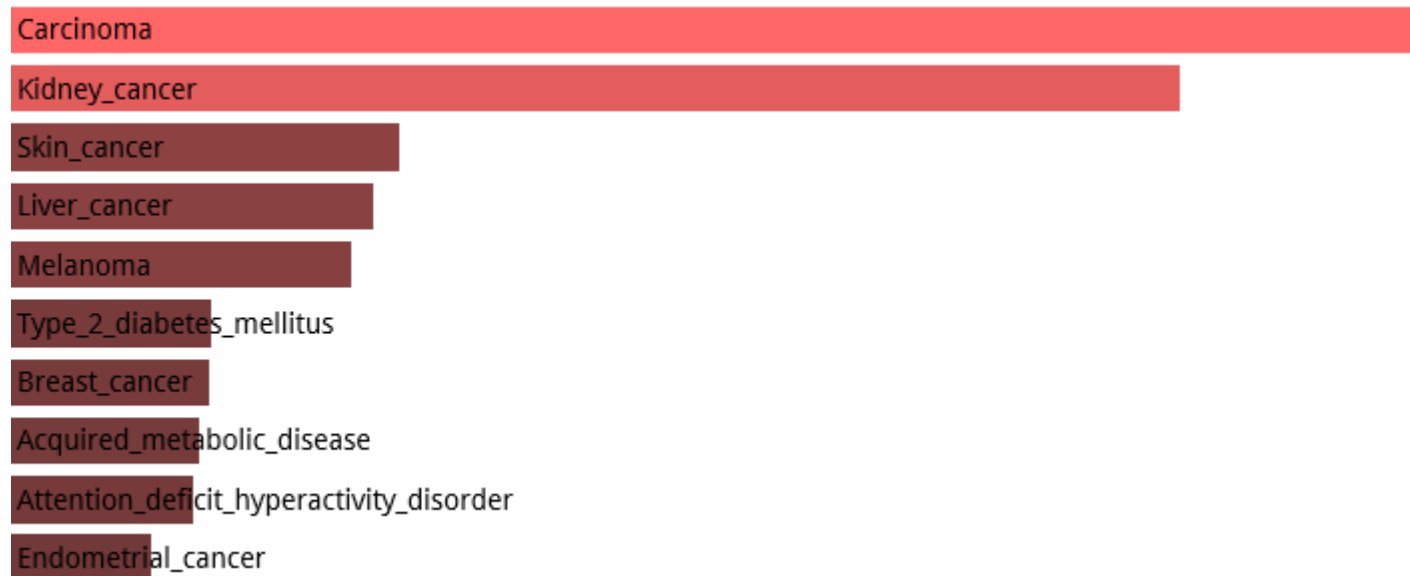


Jensen TISSUES 2810 genes

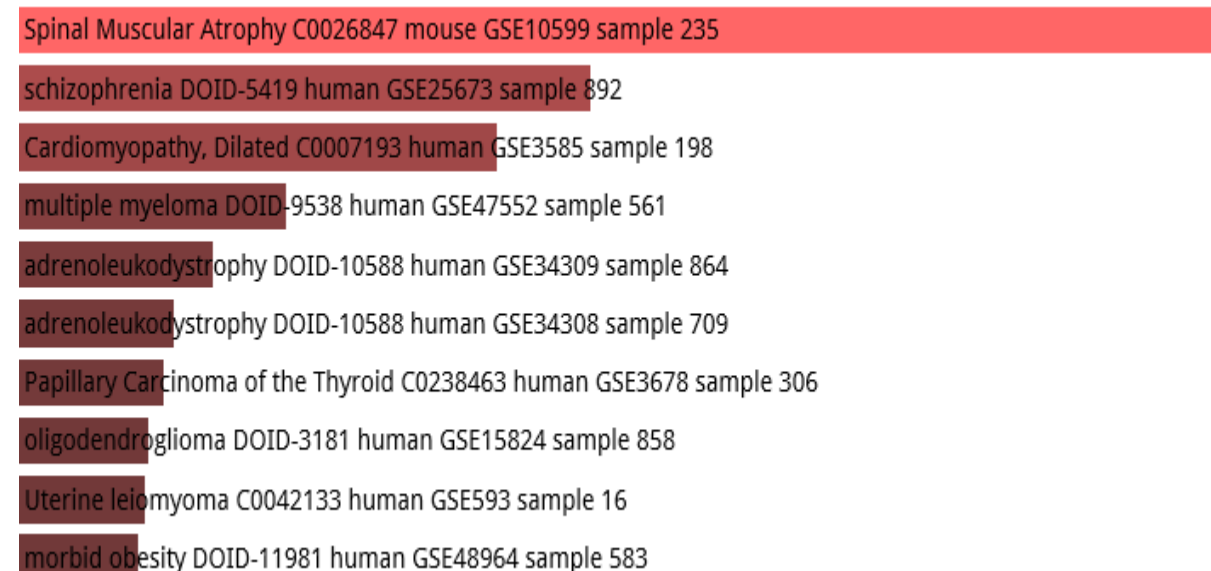


4,249 fixed human-specific regulatory loci

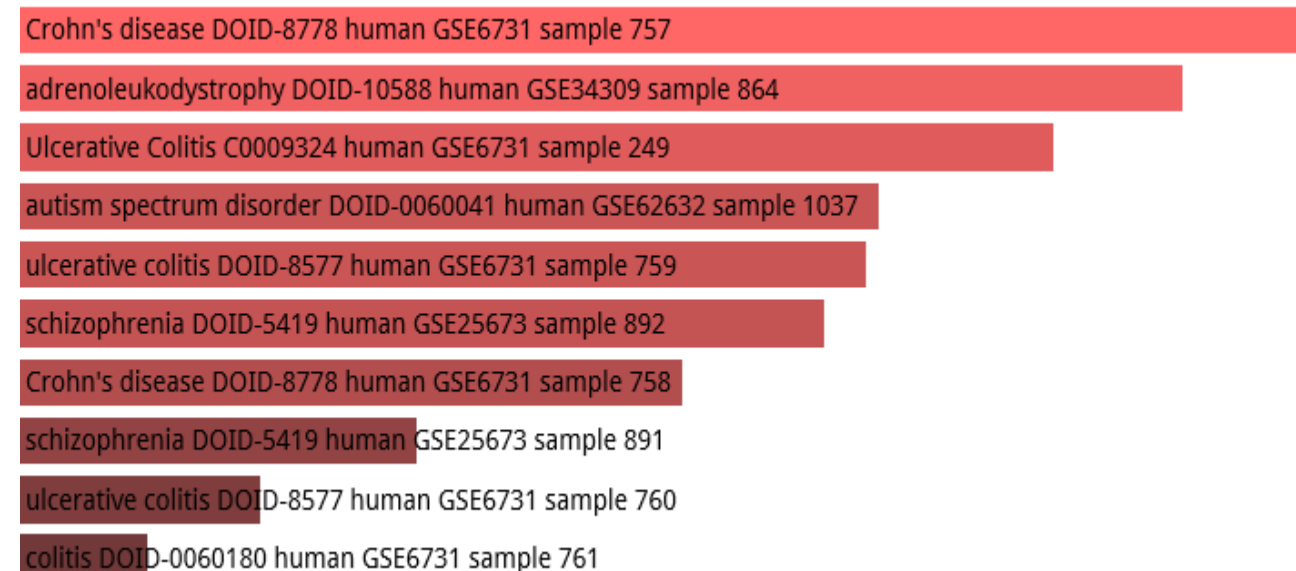
Jensen DISEASES 2810 genes



Disease Perturbations from GEO up

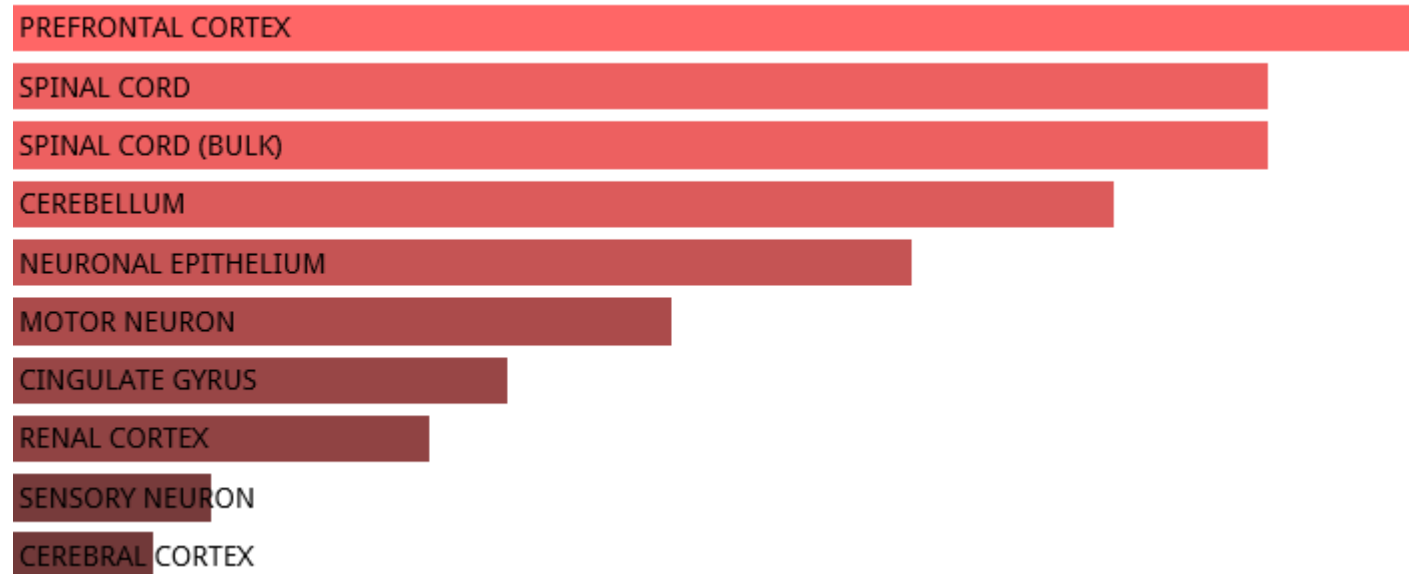


Disease Perturbations from GEO down

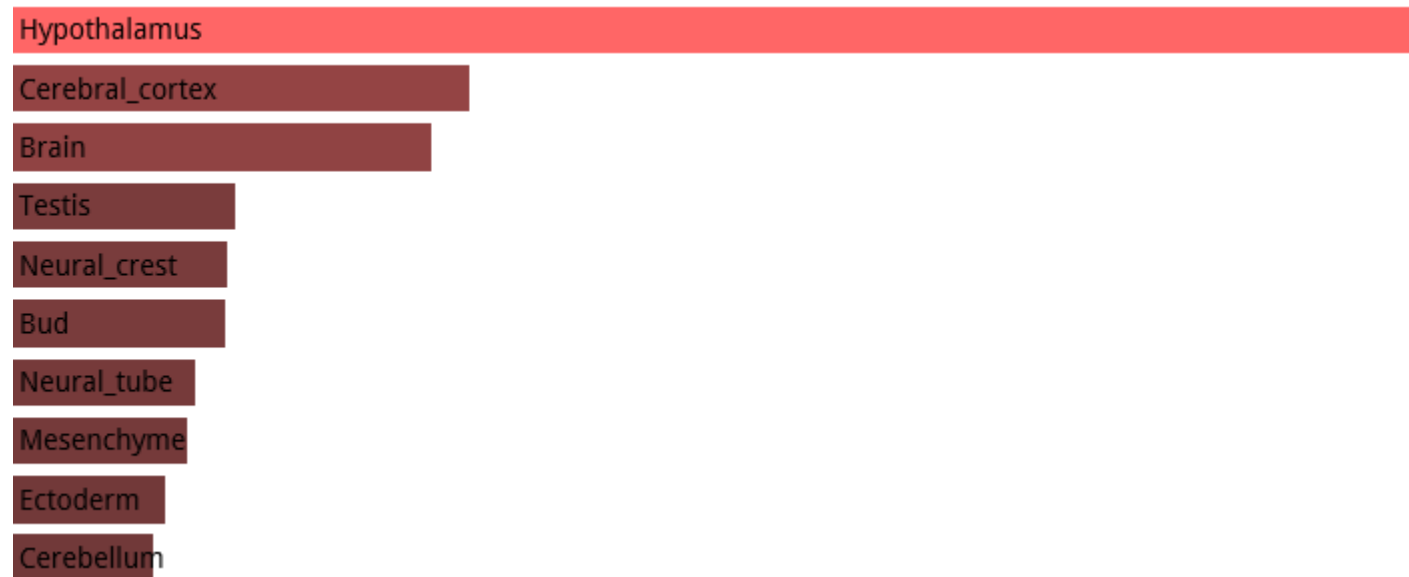


2,745 human accelerated regions

ARCHS4 Tissues 2281 genes

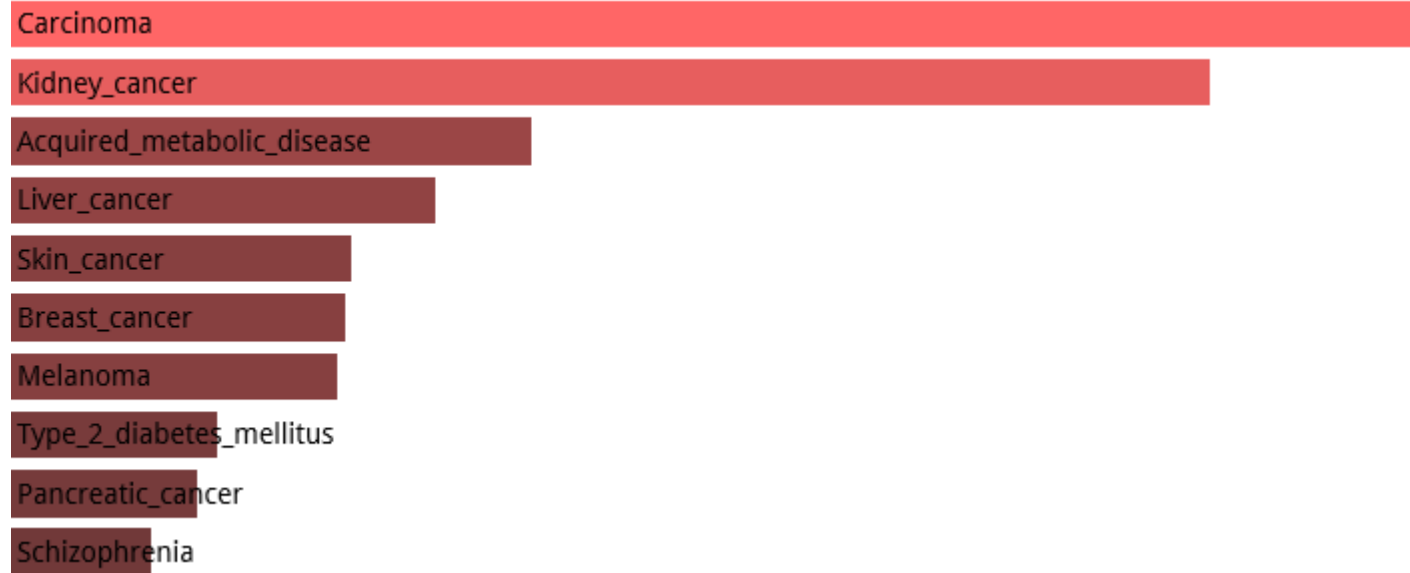


Jensen TISSUES 2281 genes

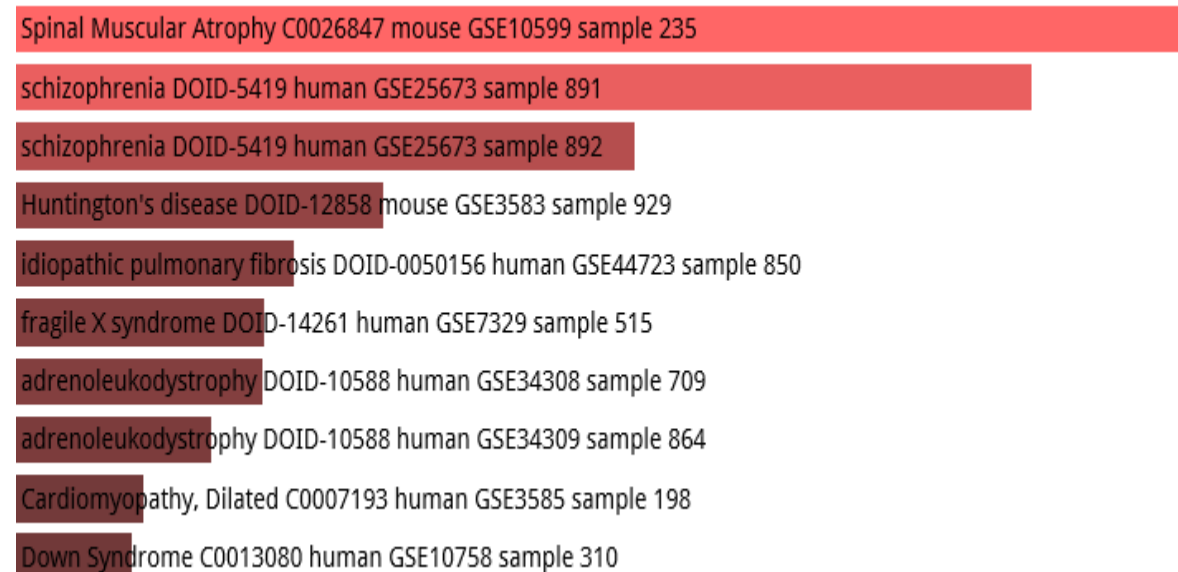


2,745 human accelerated regions

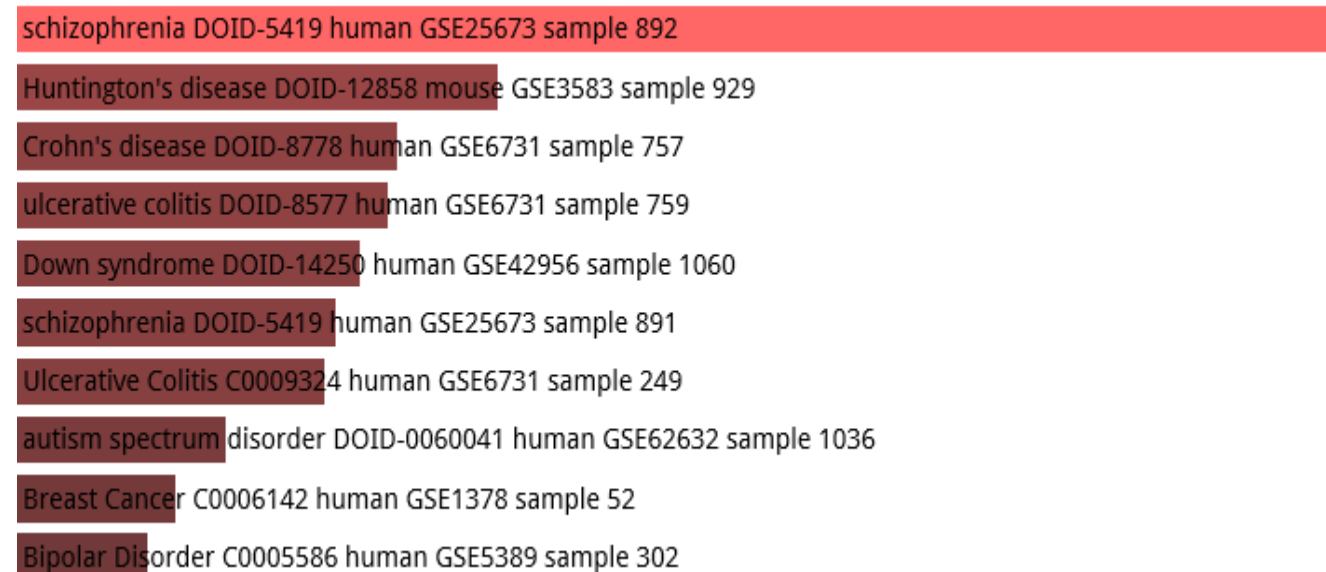
Jensen DISEASES 2281 genes



Disease Perturbations from GEO up

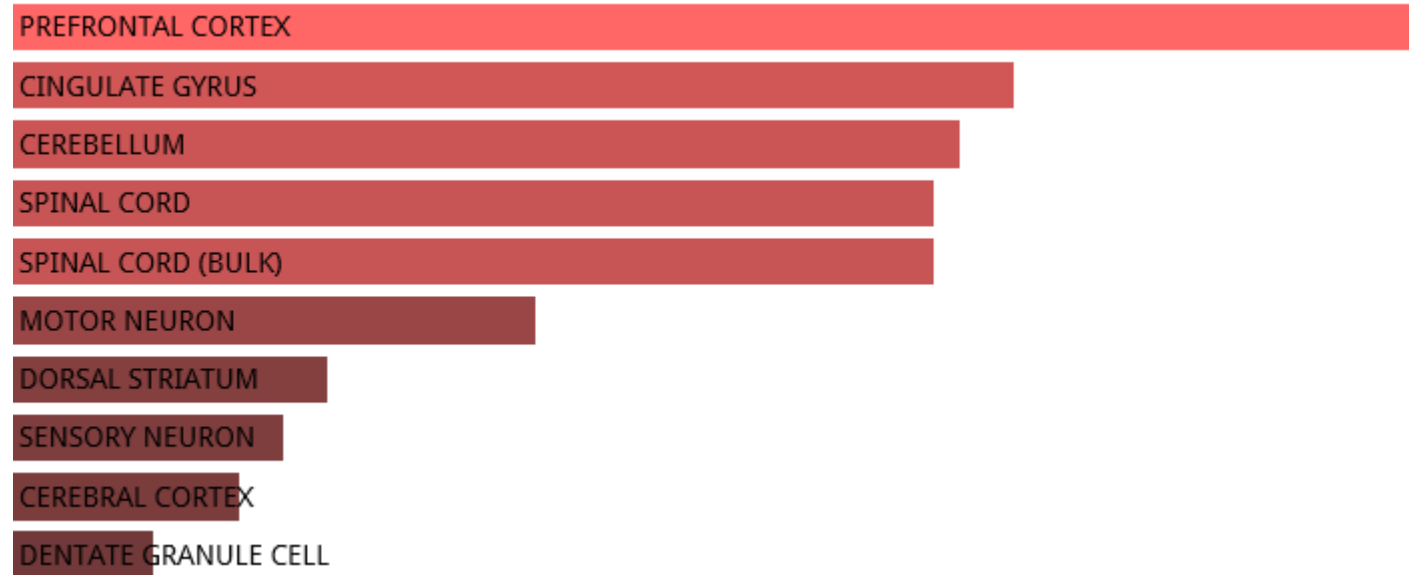


Disease Perturbations from GEO down

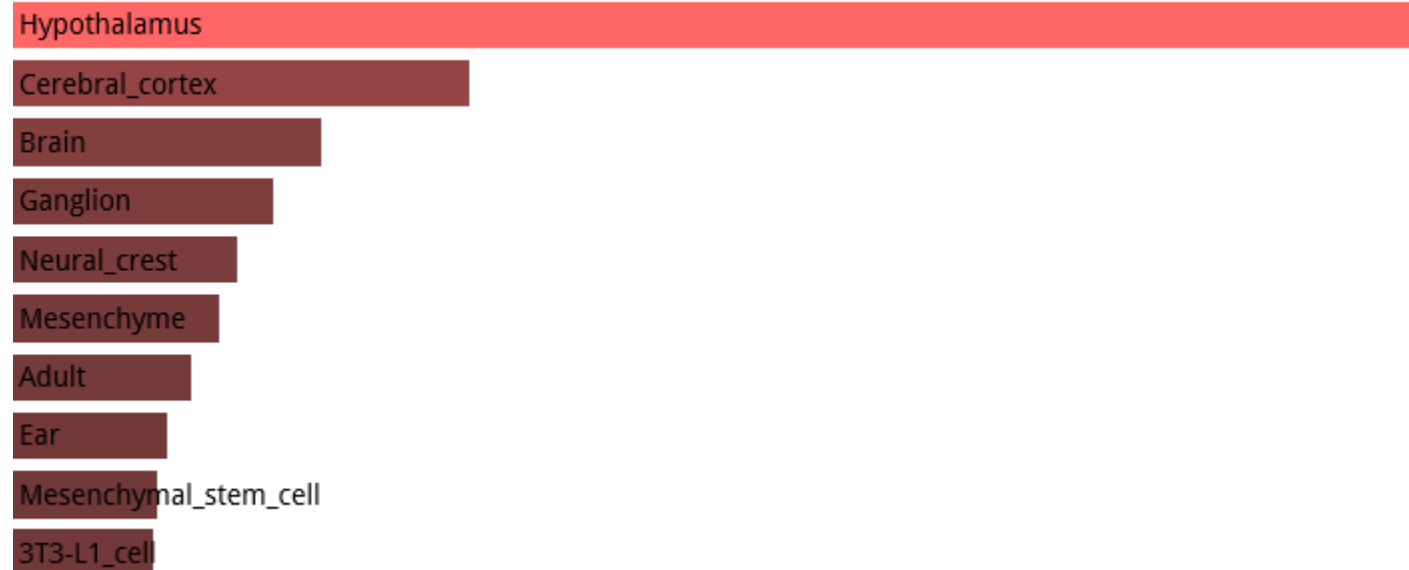


**1,932 fixed human-specific regulatory regions (hESC DHS)**

**ARCHS4 Tissues 1458 genes**

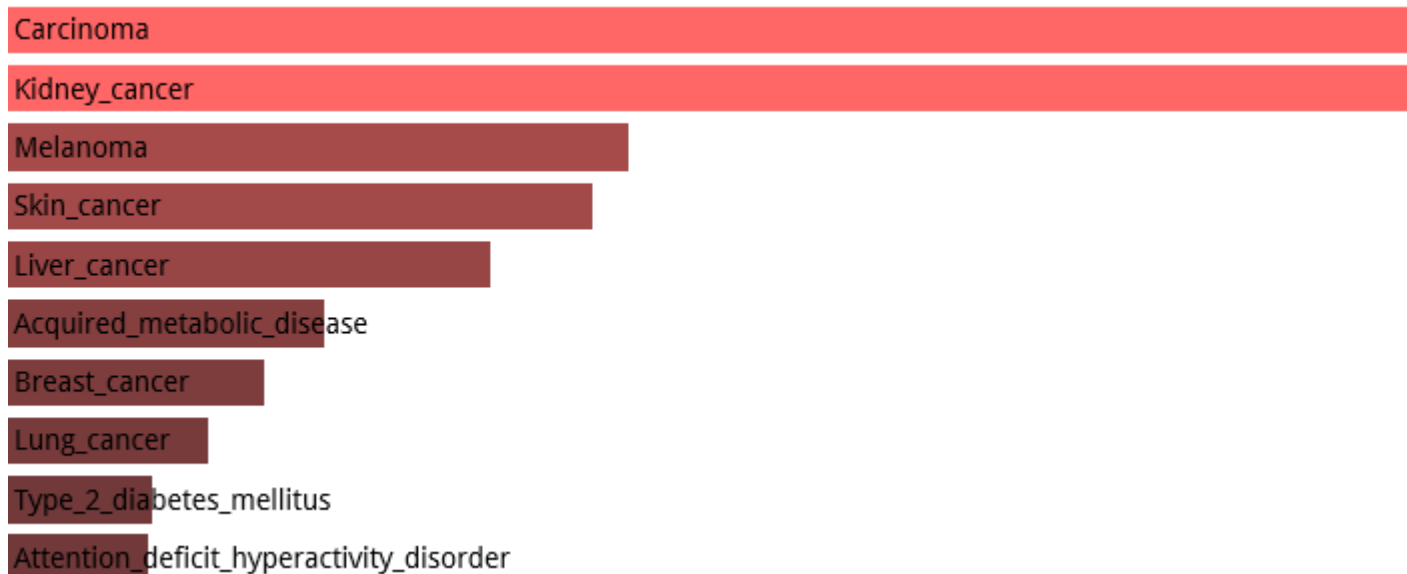


**Jensen TISSUES 1458 genes**



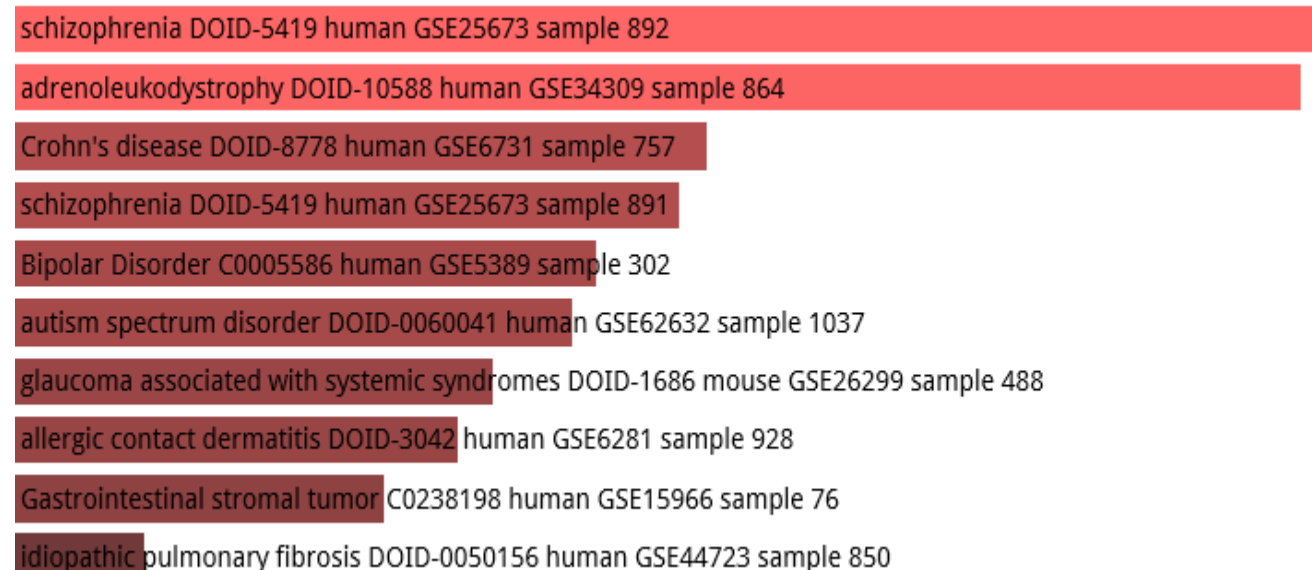
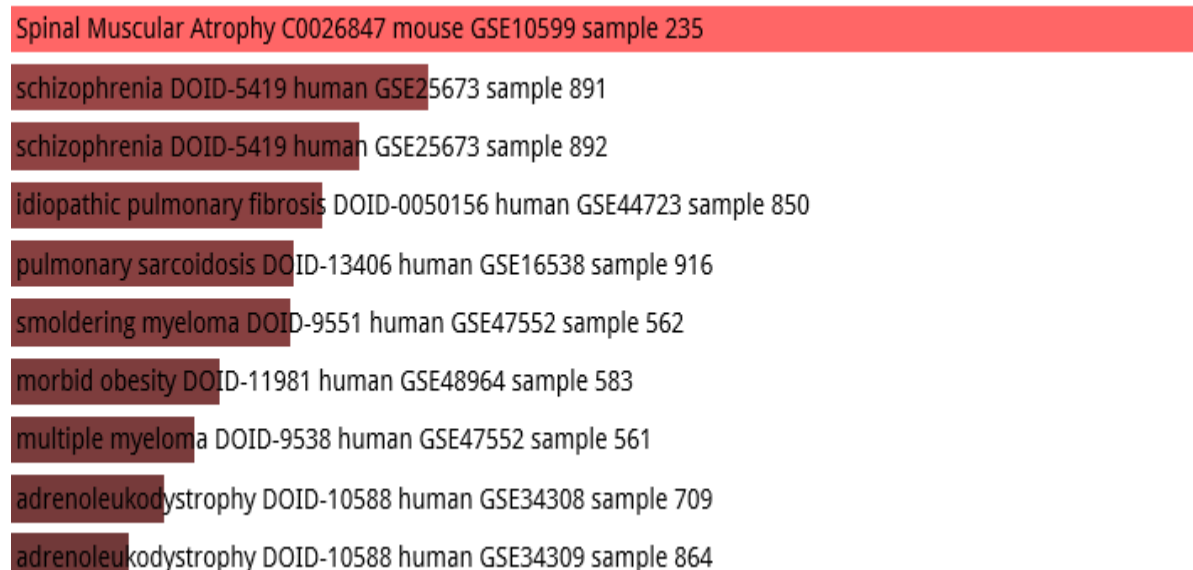
## 1,932 fixed human-specific regulatory regions (hESC DHS)

## Jensen DISEASES 1458 genes



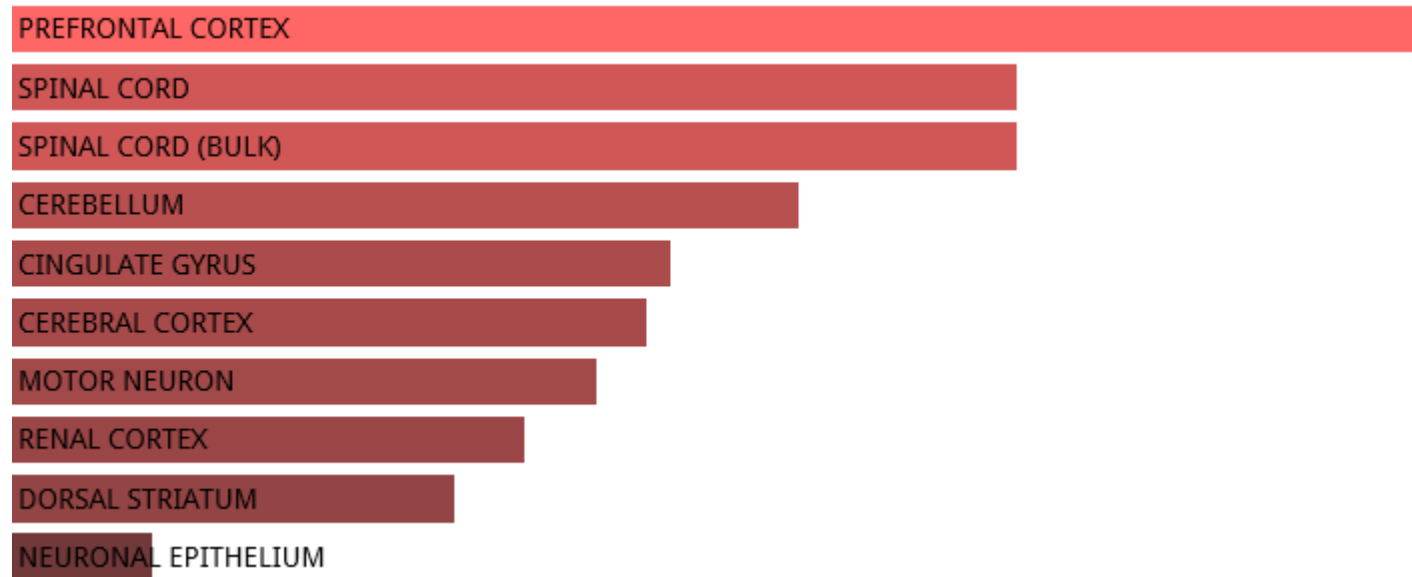
## Disease Perturbations from GEO up

## Disease Perturbations from GEO down

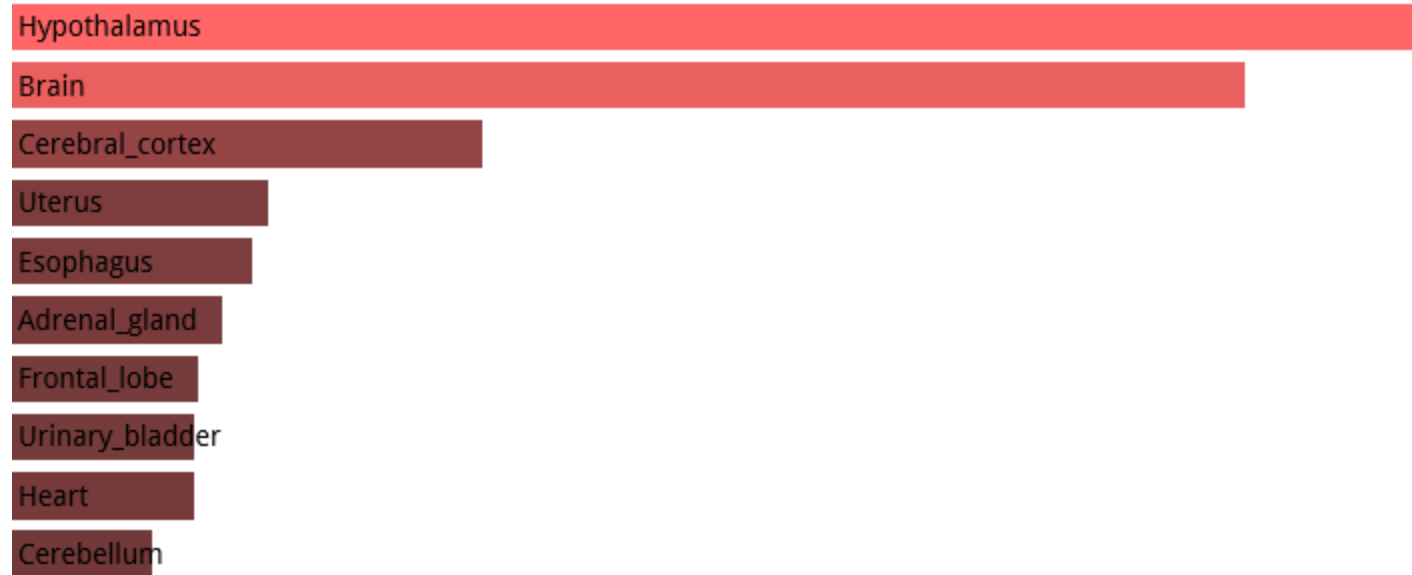


**1,000 human-biased cranial neural crest cells (CNCCs) enhancers**

**ARCHS4 Tissues 1439 genes**



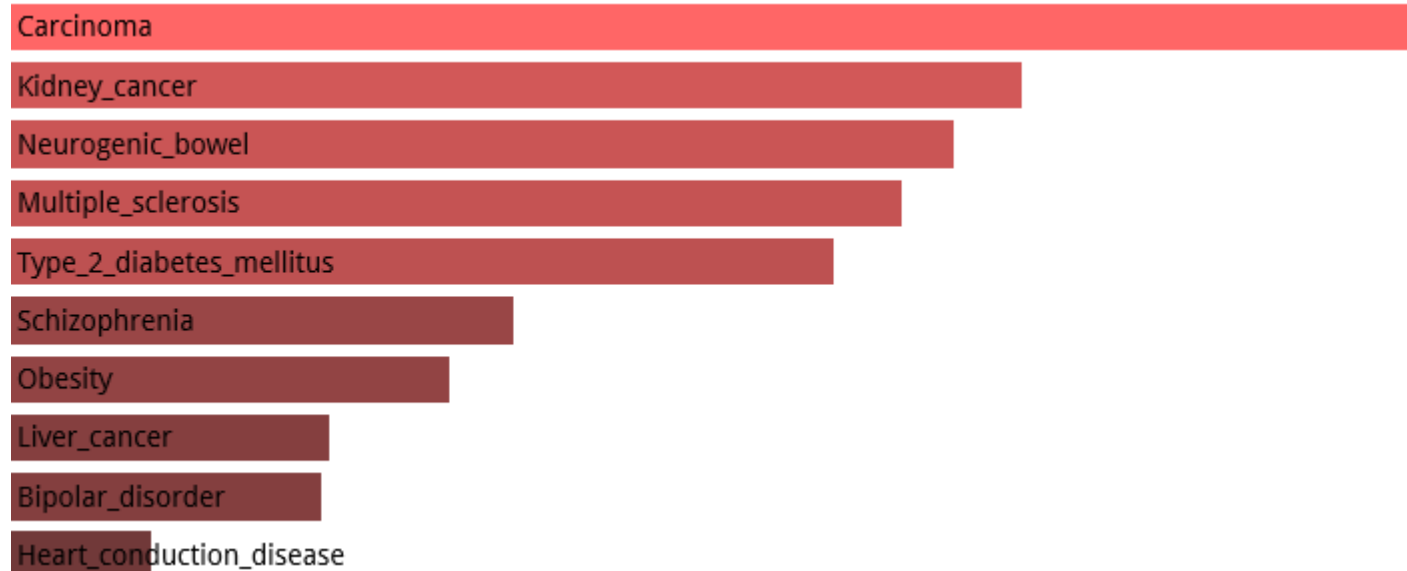
**Jensen TISSUES 1439 genes**



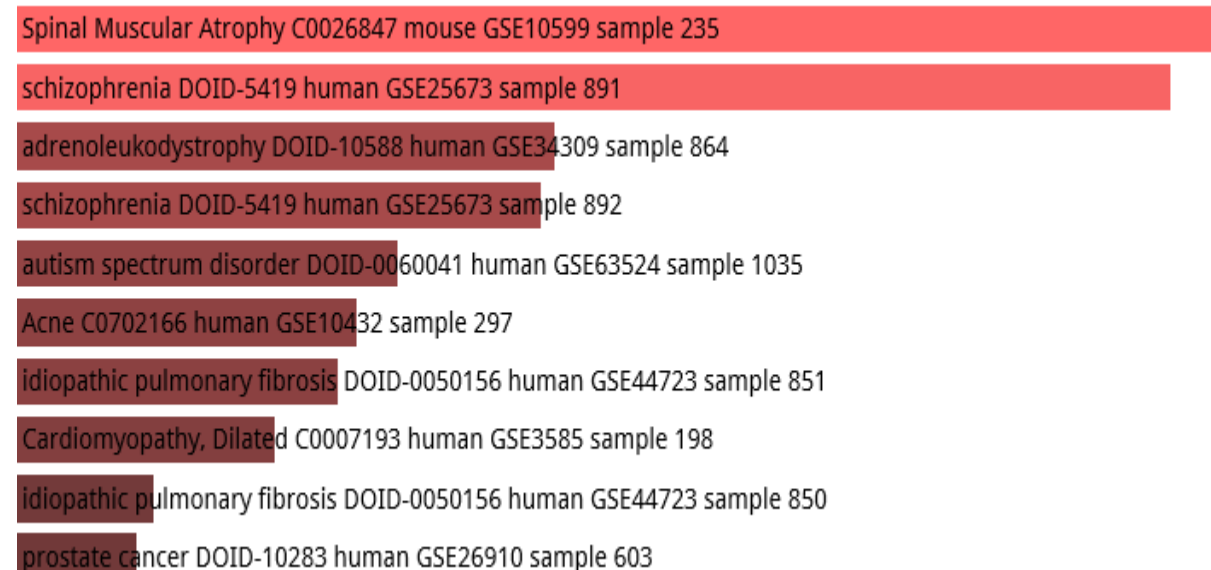


## 1,000 human-biased cranial neural crest cells (CNCCs) enhancers

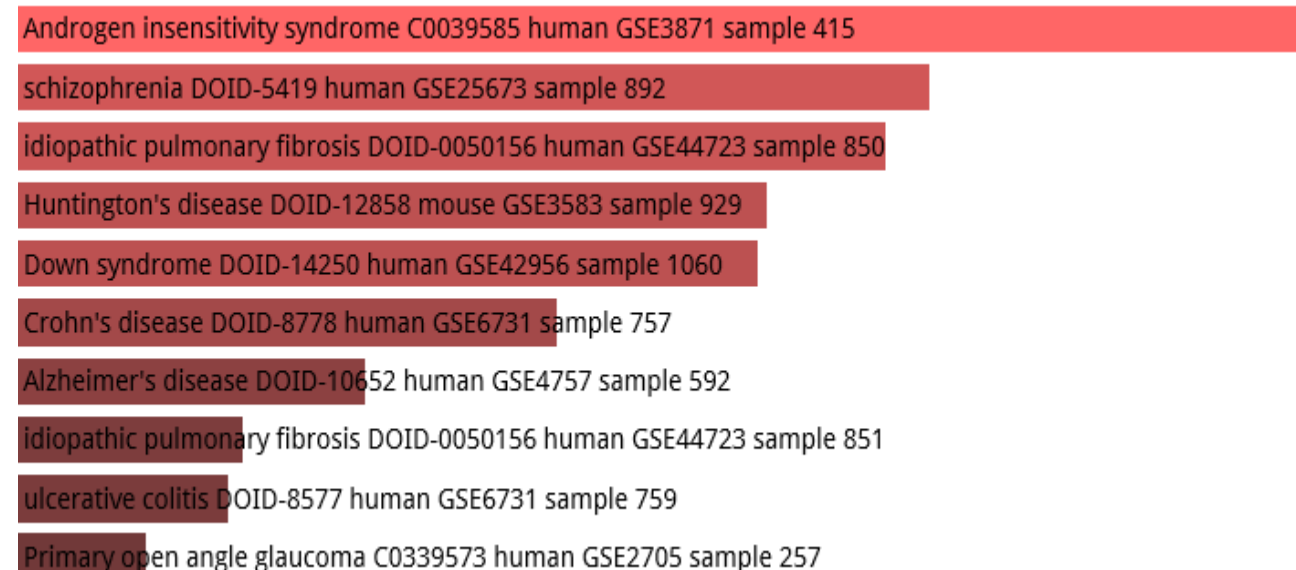
## Jensen DISEASES 1439 genes



## Disease Perturbations from GEO up

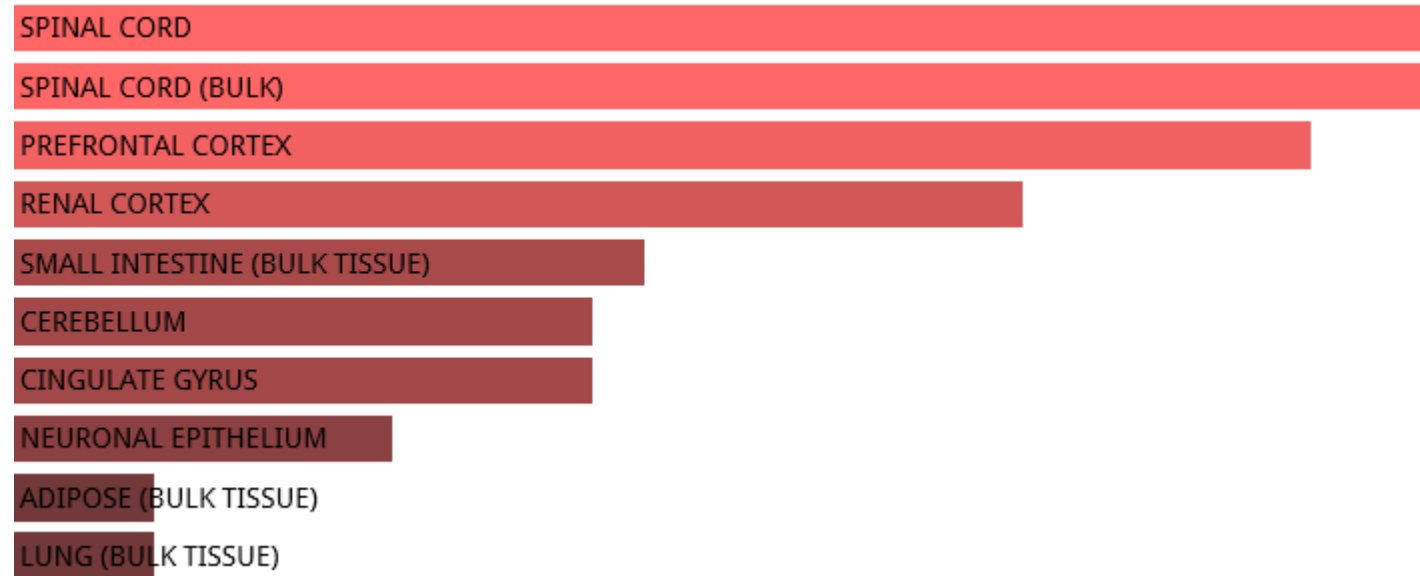


## Disease Perturbations from GEO down

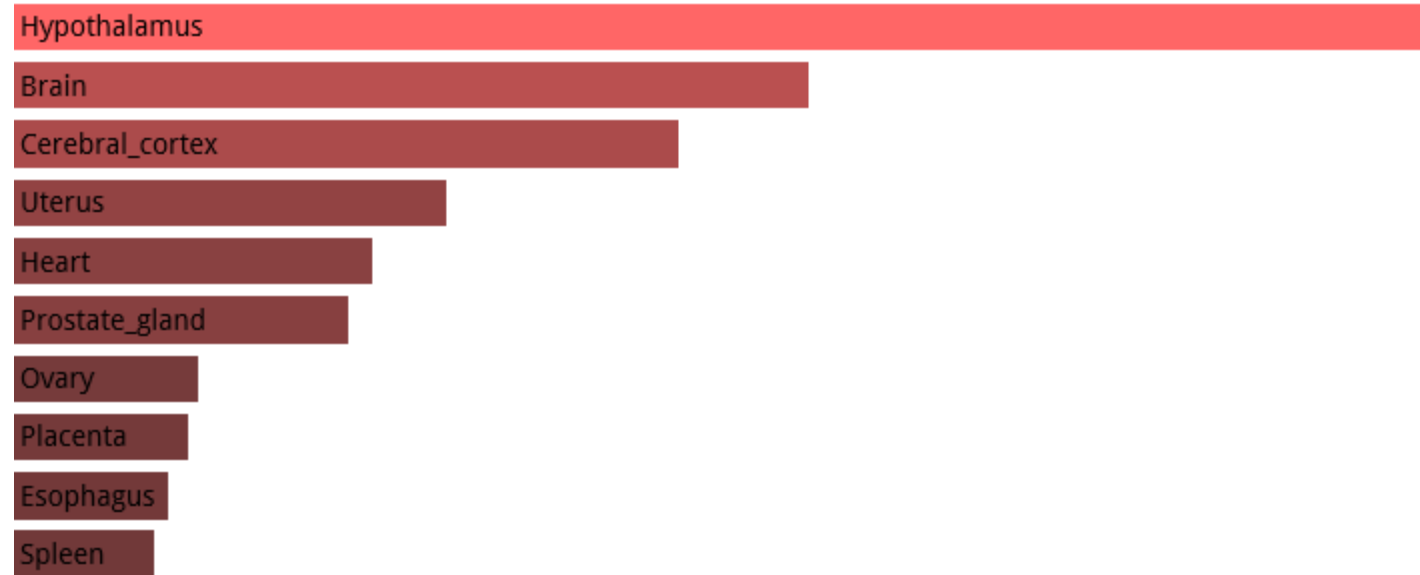


# 1,000 chimp-biased cranial neural crest cells (CNCCs) enhancers

## ARCHS4 Tissues 1445 genes

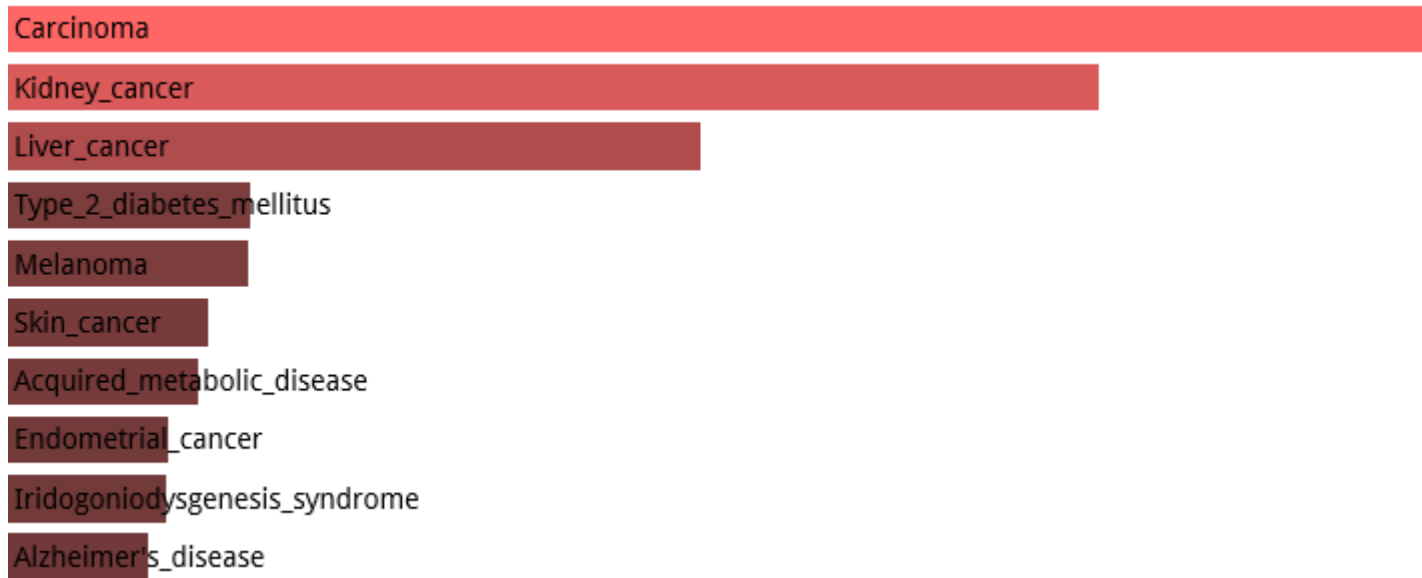


## Jensen TISSUES 1445 genes

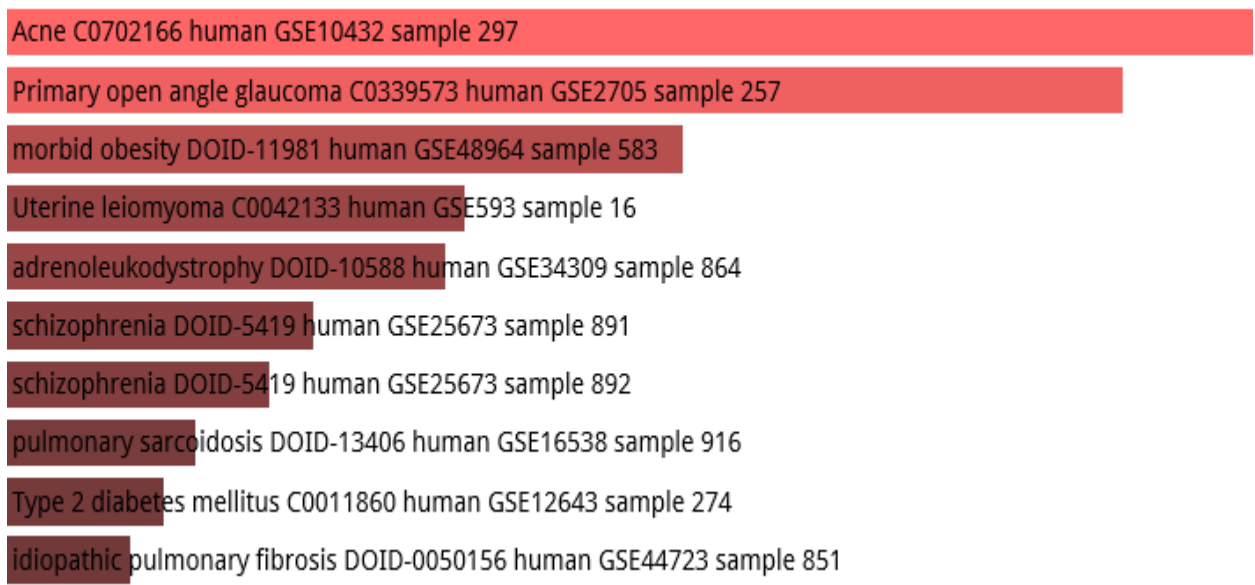


# 1,000 chimp-biased cranial neural crest cells (CNCCs) enhancers

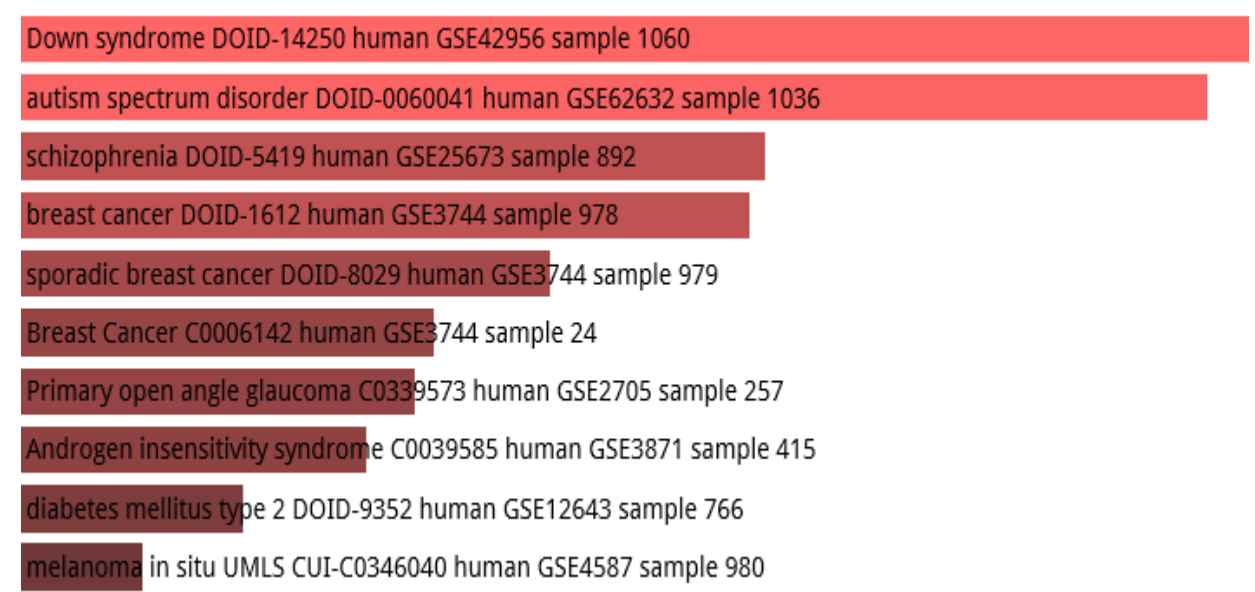
# Jensen DISEASES 1445 genes



## Disease Perturbations from GEO up

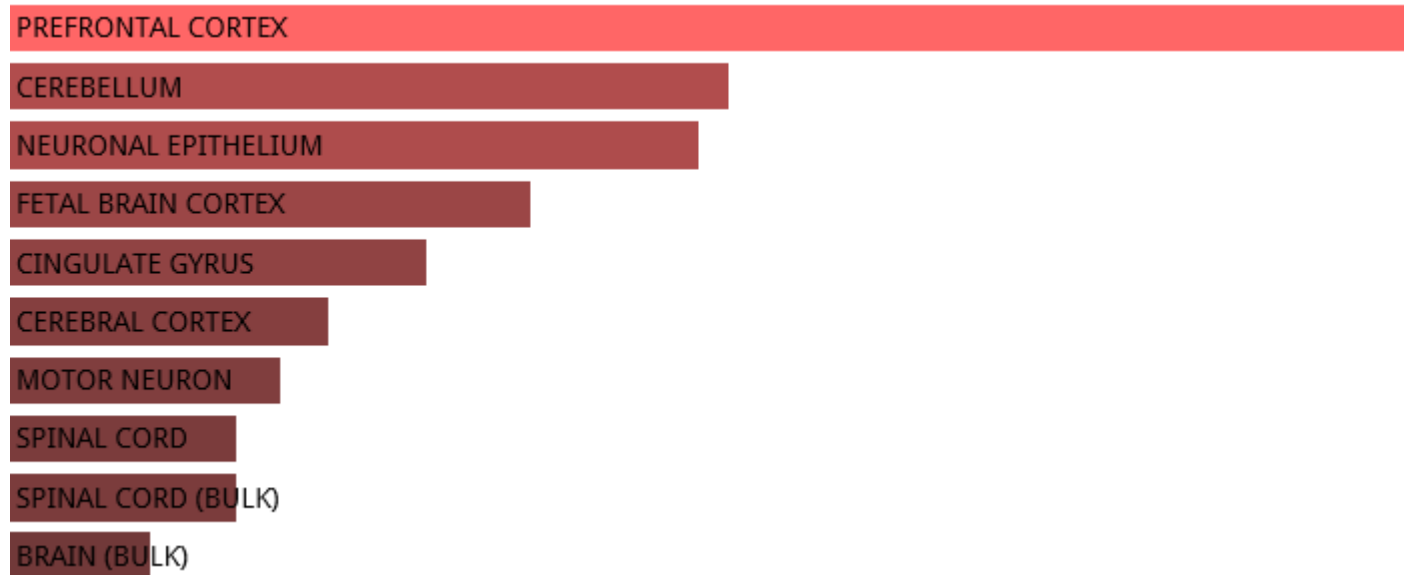


## Disease Perturbations from GEO down

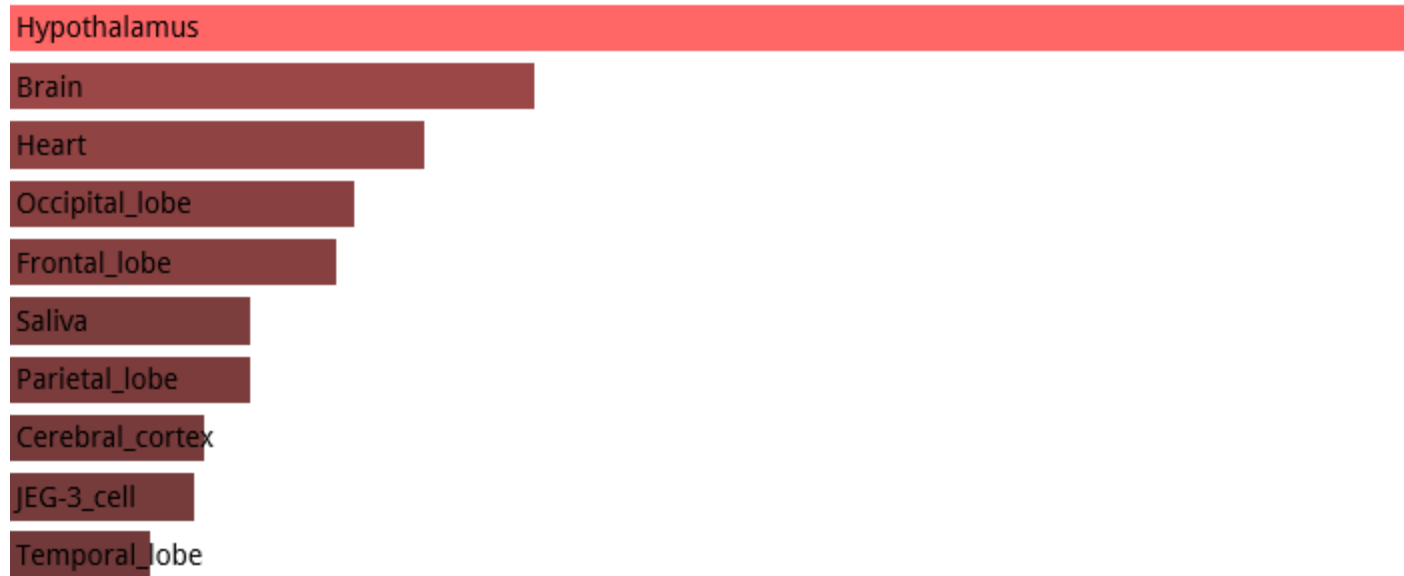


**1,619 human-specific functional  
hESC enhancers**

ARCHS4 Tissues 1214 genes

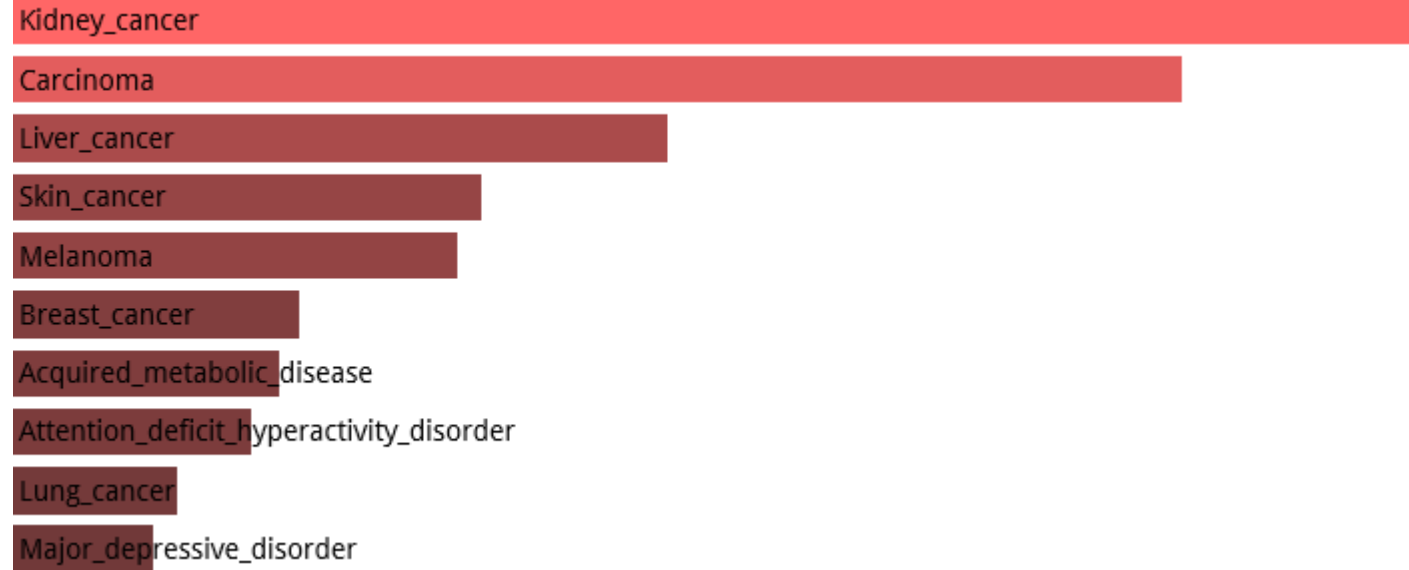


Jensen TISSUES 1214 genes

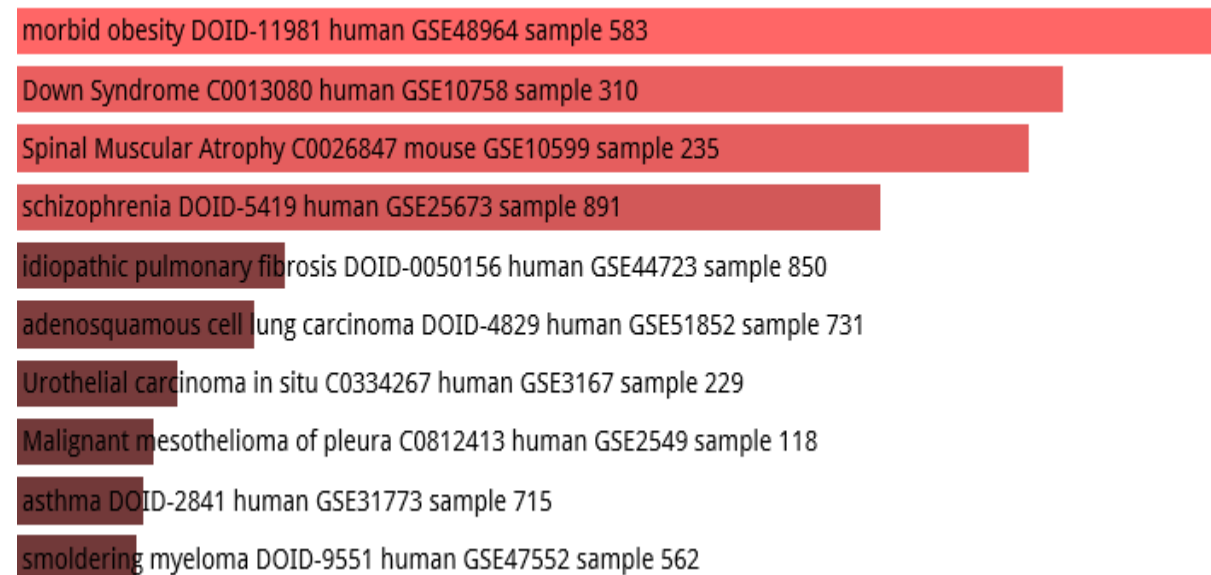


## 1,619 human-specific functional hESC enhancers

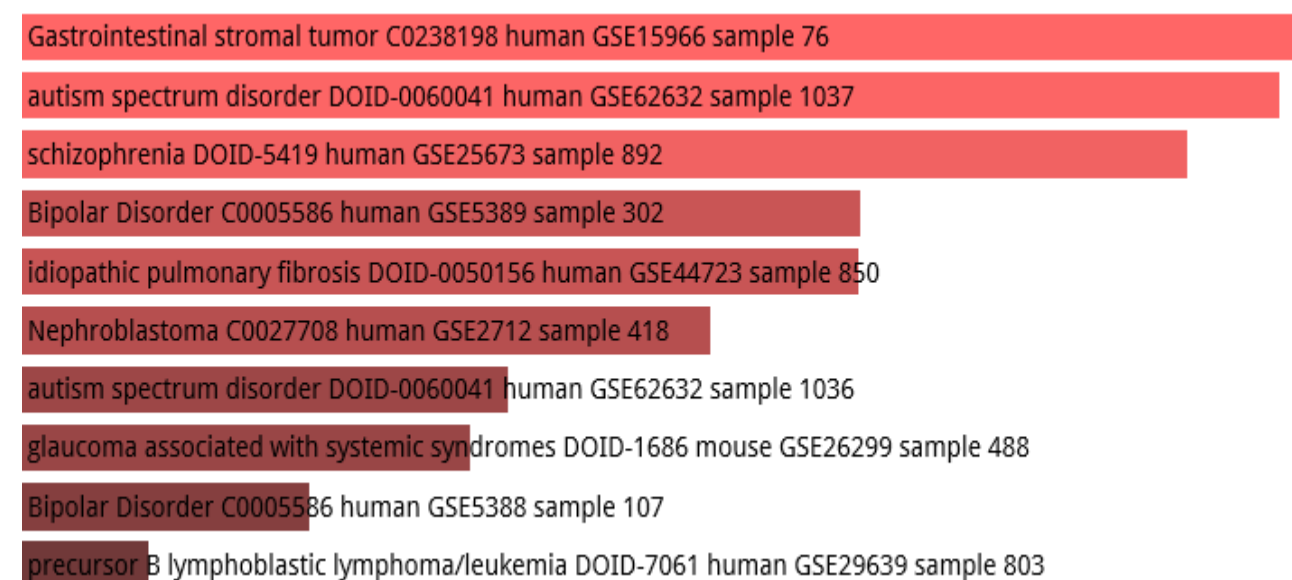
## Jensen DISEASES 1214 genes



## Disease Perturbations from GEO up 1214 genes

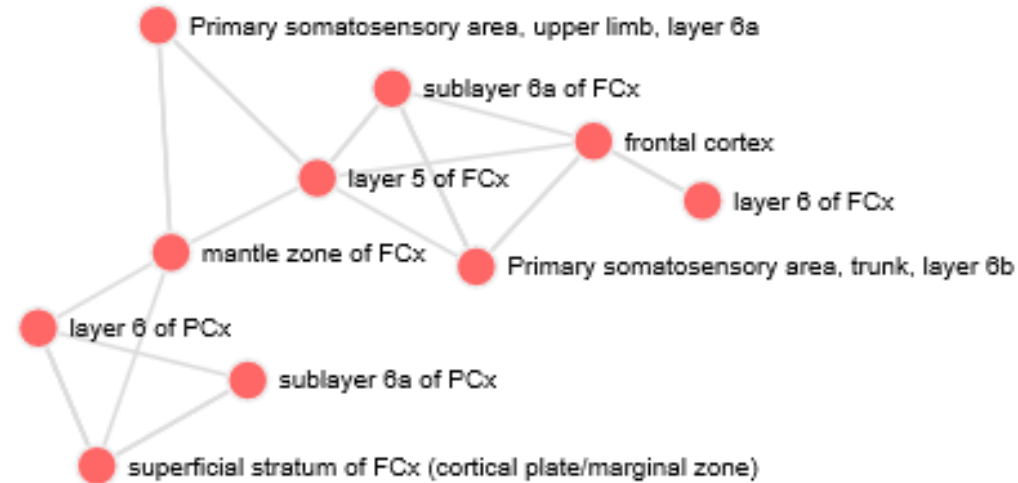
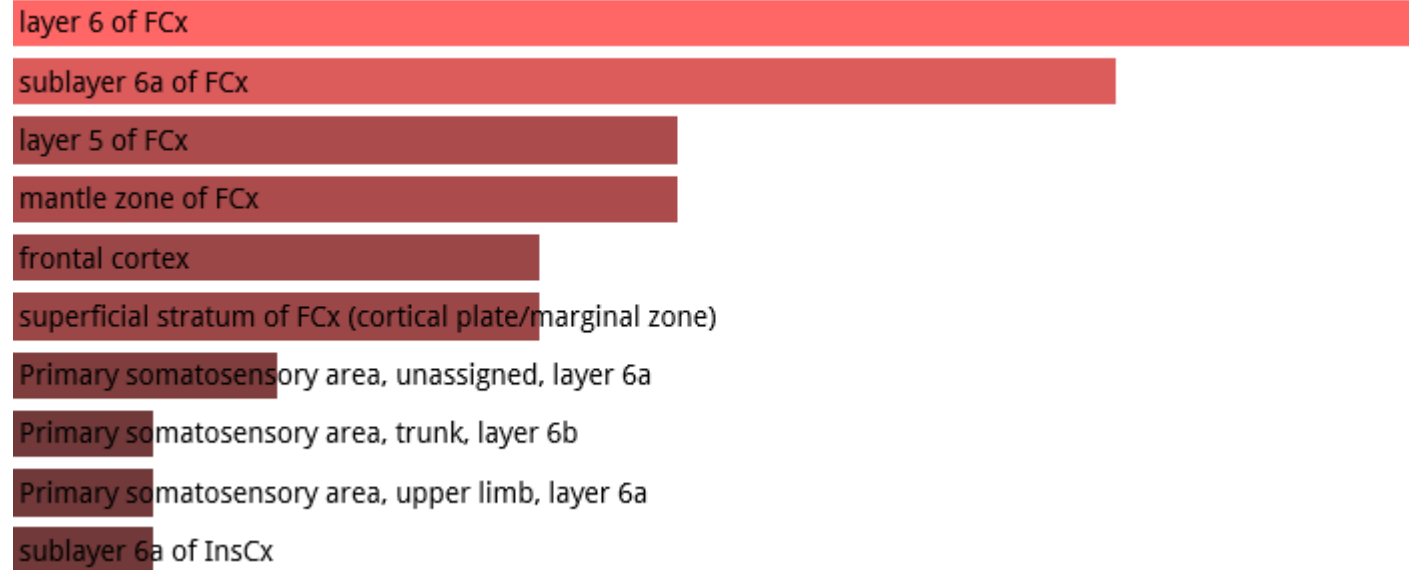


## Disease Perturbations from GEO down 1214 genes



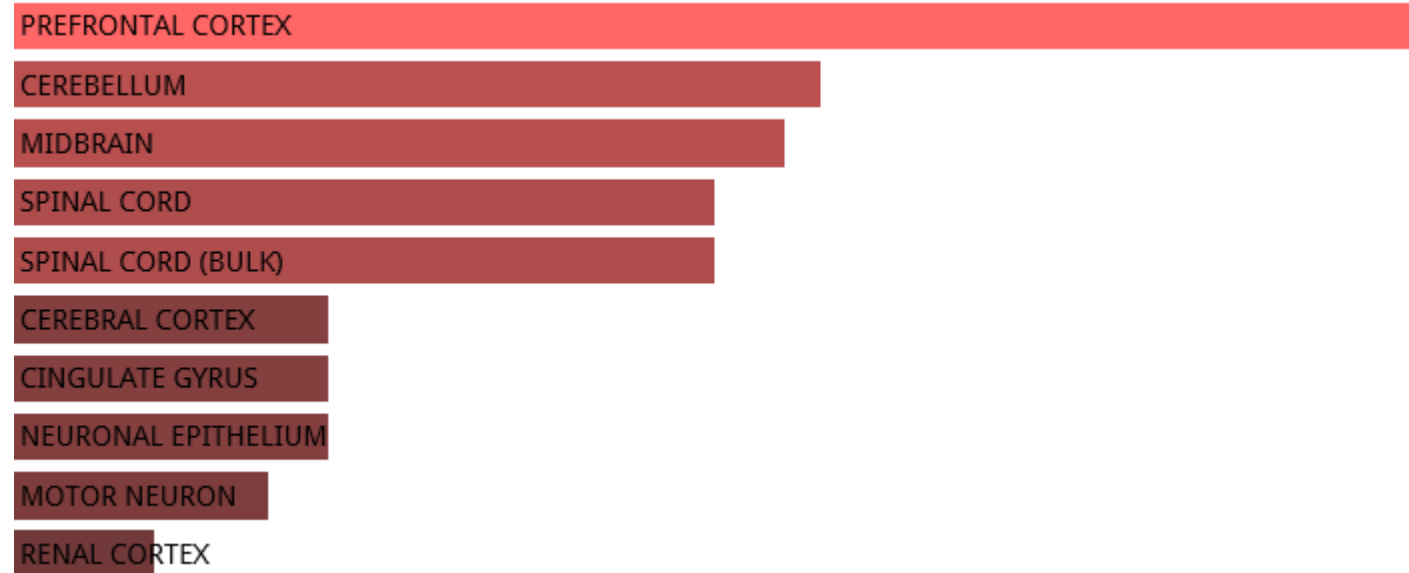
# 1,619 human-specific functional hESC enhancers

## Allen Brain Atlas up 1214 genes

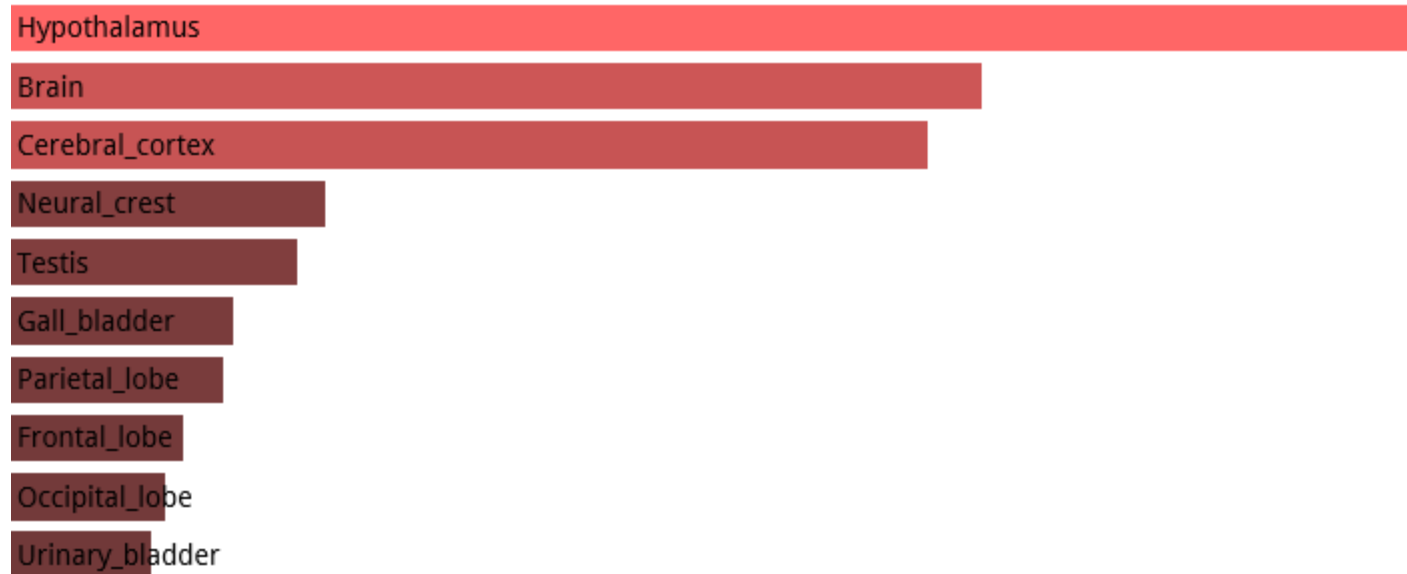


## 524 human-specific DHS

## ARCHS4 Tissues 747 genes

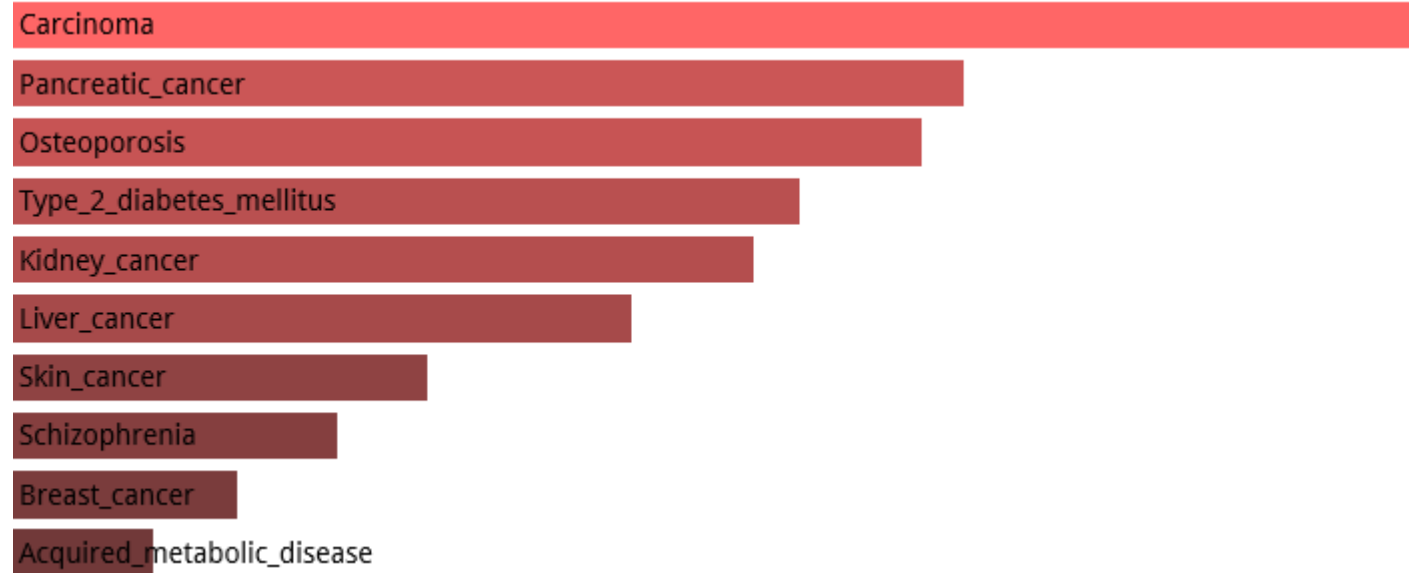


## Jensen TISSUES 747 genes

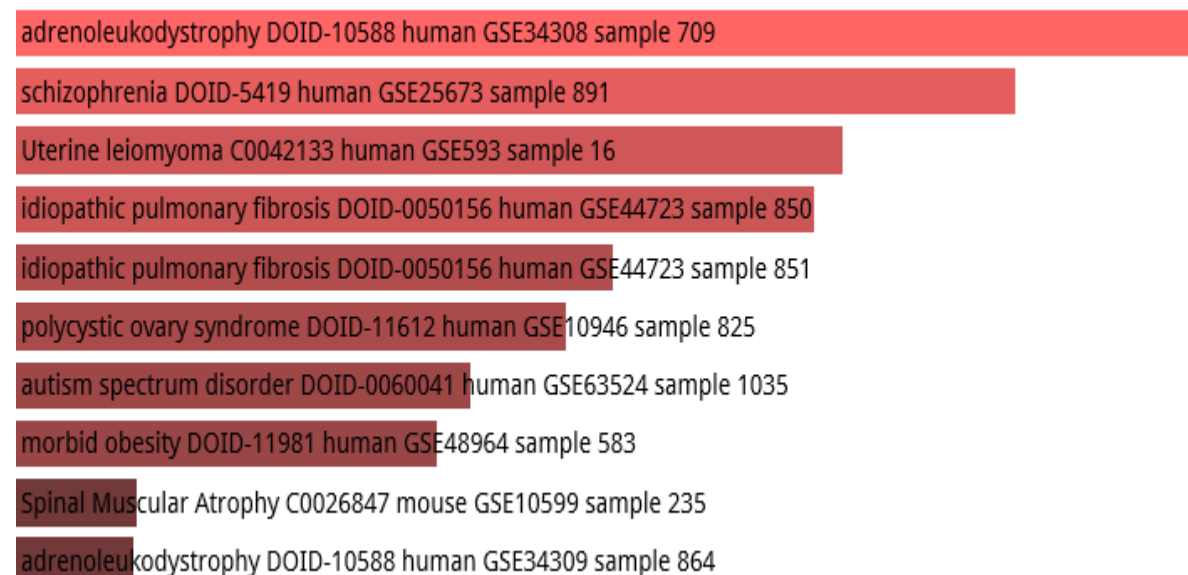


## 524 human-specific DHS

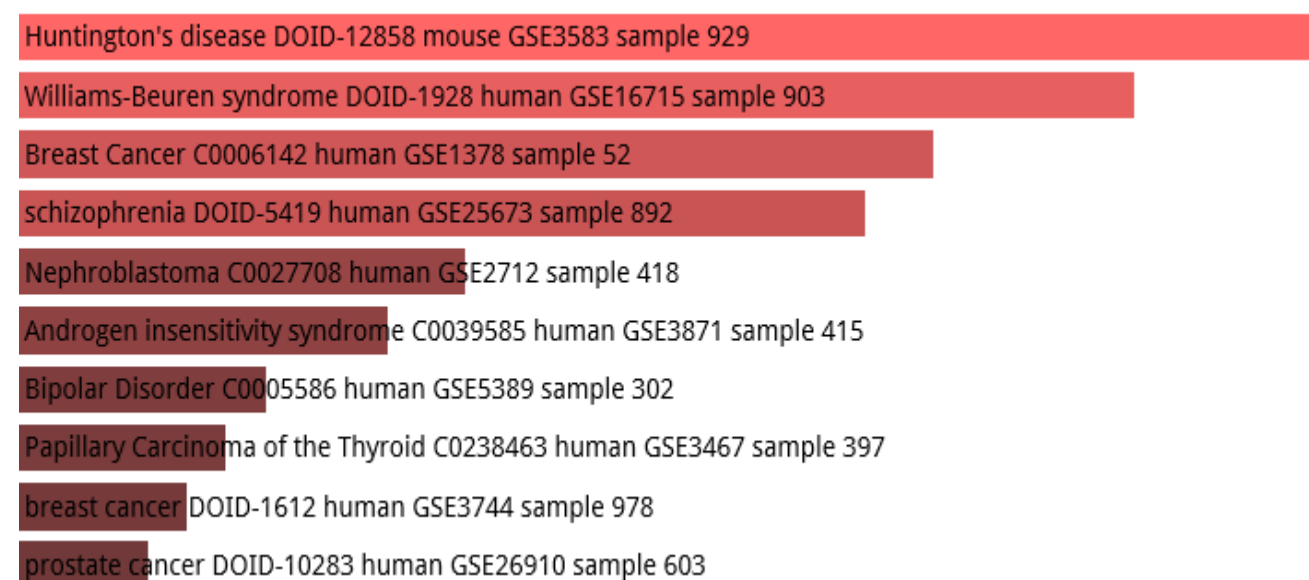
## Jensen DISEASES 747 genes



## Disease Perturbations from GEO up 747 genes



## Disease Perturbations from GEO down 747 genes





**42,847 human genes not linked by the GREAT algorithm with HSRS were randomly split into 21 control gene sets of various sizes ranging from 2,847 to 6,847 genes and subjected to the GSEA**

**GSEA OF TWENTY-ONE CONTROL GENE SETS NOT  
ASSOCIATED WITH HUMAN-SPECIFIC GENOMIC  
REGULATORY LOCI**

## Random Set 1

ASTROCYTE

KIDNEY (BULK TISSUE)

GASTRIC TISSUE (BULK)

SPINAL CORD (BULK)

SMALL INTESTINE (BULK TISSUE)

OVARY (BULK TISSUE)

SPINAL CORD

PLACENTA (BULK)

SKIN (BULK TISSUE)

HUMAN EMBRYO

ARCHS4 Tissues random 6000 genes

**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 6847 genes

**NO DATA AVAILABLE**

## Random Set 1

## Jensen DISEASES random 6000 genes

Reticular\_dysgenesis  
Amyotrophic\_lateral\_sclerosis\_type\_8  
Succinic\_semi-aldehyde\_dehydrogenase\_deficiency  
Rhizomelic\_chondrodysplasia\_punctata  
Biotinidase\_deficiency  
Intrahepatic\_cholestasis  
Conversion\_disorder  
Retinal\_ischemia  
Hypersensitivity\_reaction\_type\_II\_disease  
Pseudohermaphroditism

**NO SIGNIFICANT ENRICHMENT**

## Disease Perturbations from GEO up random 6000 genes

pancreatic ductal adenocarcinoma DOID-3498 mouse GSE53659 sample 699  
juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 772  
amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 679  
juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 771  
nemaline myopathy DOID-3191 mouse GSE3384 sample 971  
nemaline myopathy DOID-3191 mouse GSE3384 sample 971  
Nemaline Myopathy C0206157 mouse GSE3384 sample 317  
lupus erythematosus DOID-8857 human GSE30153 sample 739  
Duchenne muscular dystrophy (DMD) C0013264 mouse GSE1008 sample 298  
idiopathic pulmonary fibrosis DOID-0050156 human GSE24206 sample 872

**NO SIGNIFICANT ENRICHMENT**

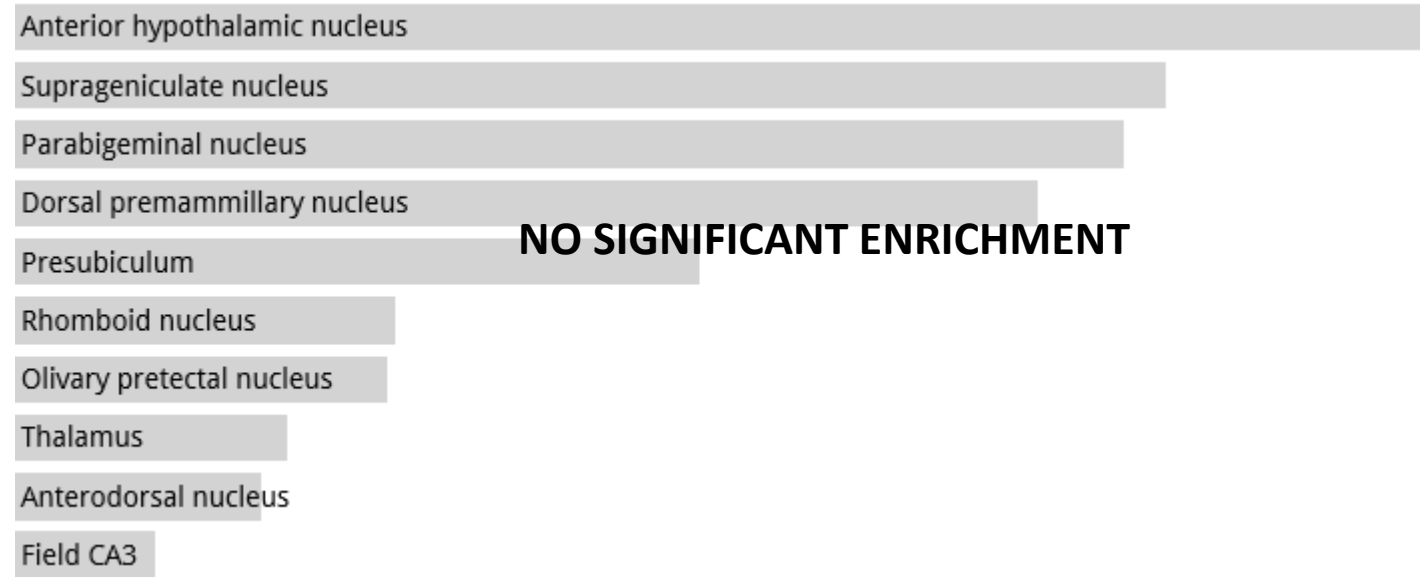
## Disease Perturbations from GEO down random 6000 genes

pulmonary tuberculosis DOID-2957 mouse GSE48027 sample 831  
polycystic ovary syndrome DOID-11612 human GSE48301 sample 558  
Nicotine addiction C0028043 human GSE11208 sample 325  
autism spectrum disorder DOID-0060041 human GSE62632 sample 1037  
mental retardation DOID-1059 human GSE3375 sample 1047  
Diabetic Neuropathy C0011882 mouse GSE11343 sample 7  
Alcohol poisoning C0392620 rat GSE3311 sample 288  
asthma DOID-2841 human GSE43696 sample 830  
systemic lupus erythematosus DOID-9074 human GSE55447 sample 1075  
Thymic Carcinoma C0205969 mouse GSE2501 sample 344

**NO SIGNIFICANT ENRICHMENT**

## Random Set 1

## Allen Brain Atlas up random 6000 genes



## Random Set 2

ARCHS4 Tissues random 6000 genes

GASTRIC TISSUE (BULK)

SKIN (BULK TISSUE)

OMENTUM

SMALL INTESTINE (BULK TISSUE)

CINGULATE GYRUS

COLON (BULK TISSUE)

NEURONAL EPITHELIUM

PREFRONTAL CORTEX

PANCREATIC ISLET

MYOBLAST

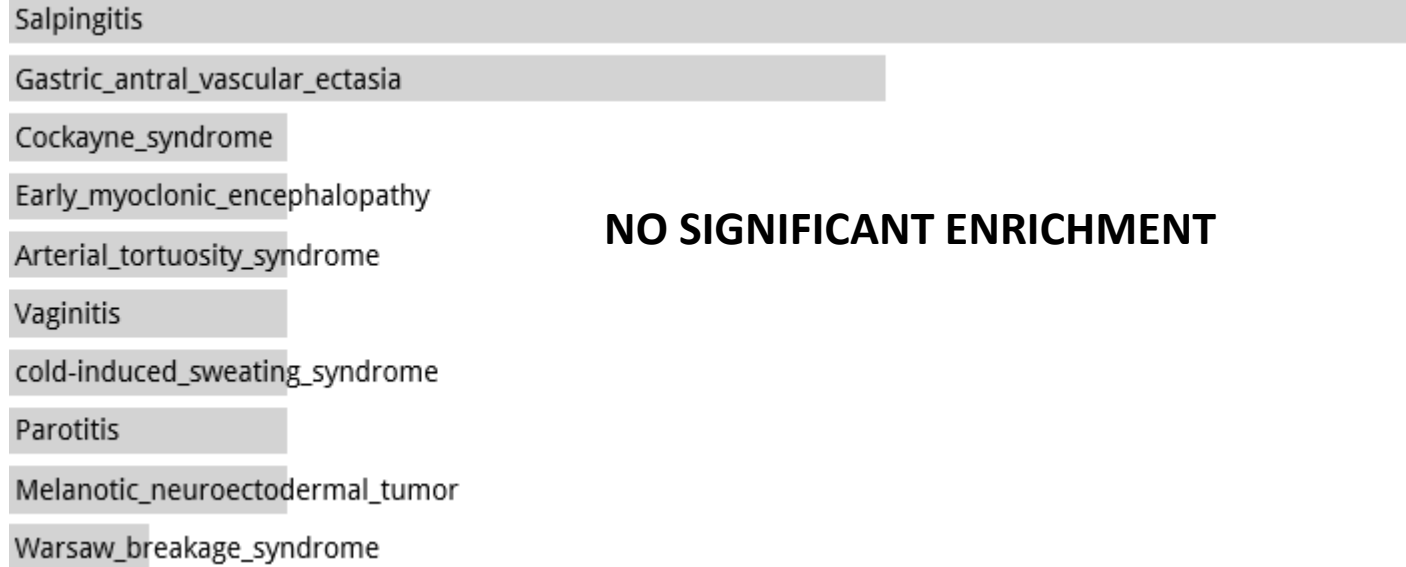
**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 6847 genes

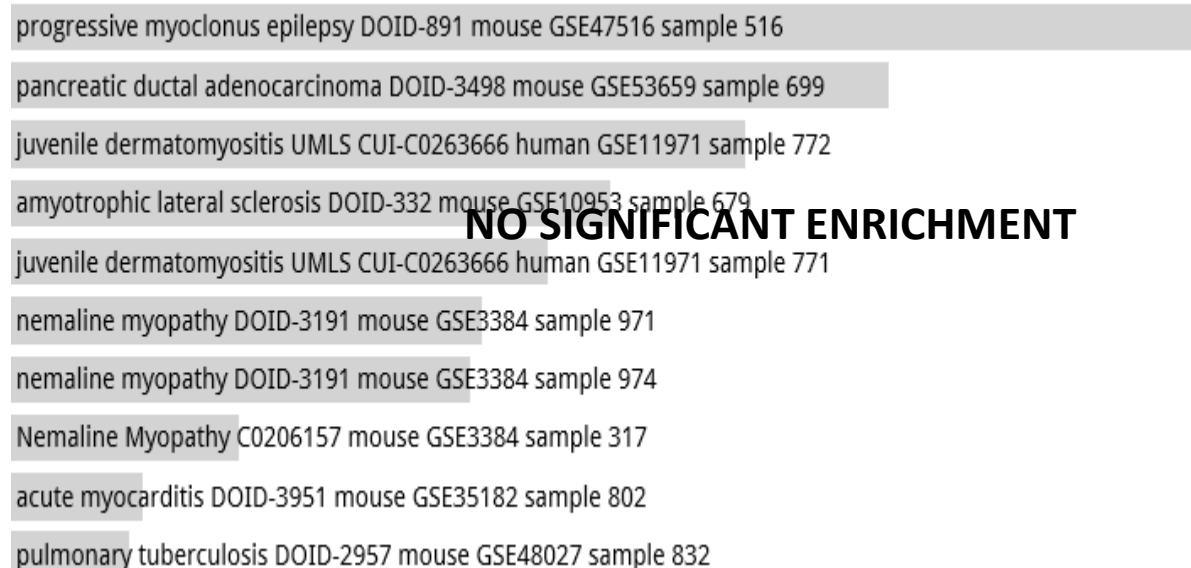
**NO DATA AVAILABLE**

## Random Set 2

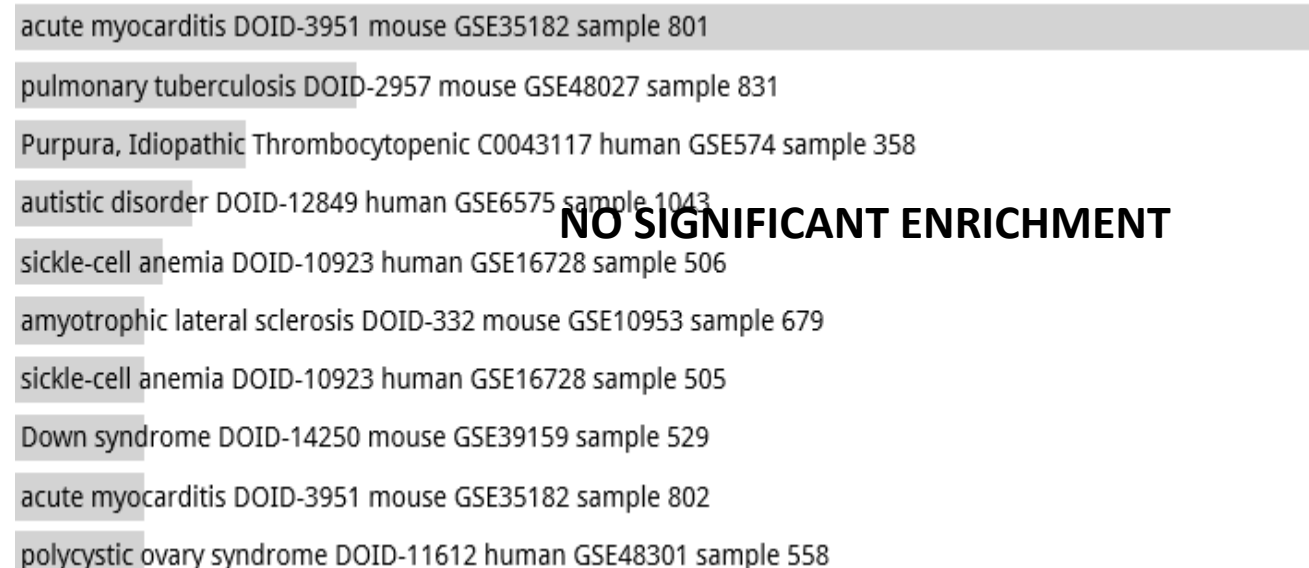
## Jensen DISEASES random 6000 genes



## Disease Perturbations from GEO up random 6000 genes

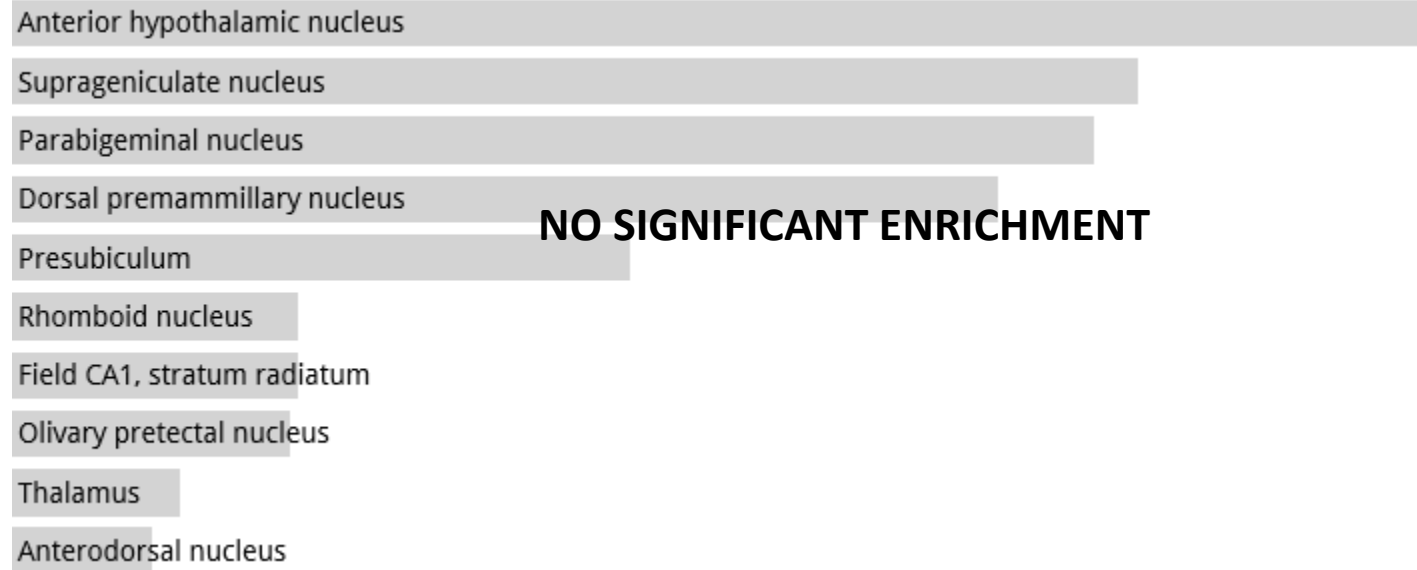


## Disease Perturbations from GEO down random 6000 genes



## Random Set 2

### Allen Brain Atlas up random 6000 genes





### Random Set 3

ARCHS4 Tissues random 6000 genes

COLONIC MUCOSA

VENTRICLE

PREFRONTAL CORTEX

BRAIN (BULK)

SUPERIOR FRONTAL GYRUS

NEURONAL EPITHELIUM

ADIPOSE (BULK TISSUE)

HUMAN ZYGOTE

RENAL CORTEX

MYOBLAST

**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 6847 genes

**NO DATA AVAILABLE**

### Random Set 3

### Jensen DISEASES random 6000 genes

Hereditary\_mucosal\_leukokeratosis

Echolalia

Dumping\_syndrome

X-linked\_endothelial\_corneal\_dystrophy

IMAGe\_syndrome

Clouston\_syndrome

Oculodentodigital\_dysplasia

Irritant\_dermatitis

Brain\_glioma

Pachyonychia\_congenita

**NO SIGNIFICANT ENRICHMENT**

### Disease Perturbations from GEO up random 6000 genes

Dental cavity, complex C0399396 human GSE1629 sample 175

Alzheimer's disease DOID-10652 human GSE36980 sample 519

actinic keratosis DOID-8866 human GSE2503 sample 628

Actinic keratosis C0022602 human GSE2503 sample 350

acute myeloid leukemia DOID-9119 human GSE9476 sample 782

ulcerative colitis DOID-8577 human GSE9452 sample 924

type 2 diabetes mellitus DOID-9352 human GSE23343 sample 895

acute myeloid leukemia DOID-9119 human GSE9476 sample 783

polycystic ovary syndrome DOID-11612 human GSE48301 sample 558

Epilepsy C0014544 human GSE7486 sample 417

**NO SIGNIFICANT ENRICHMENT**

### Disease Perturbations from GEO down random 6000 genes

multiple myeloma DOID-9538 human GSE6691 sample 787

acute myocarditis DOID-3951 mouse GSE35182 sample 801

Smoldering multiple myeloma C1531608 human GSE5900 sample 404

autistic disorder DOID-12849 human GSE6575 sample 1043

autistic disorder DOID-12849 human GSE6575 sample 1042

pulmonary tuberculosis DOID-2957 mouse GSE48027 sample 831

mental retardation DOID-1059 human GSE6575 sample 1044

Down syndrome DOID-14250 mouse GSE39159 sample 529

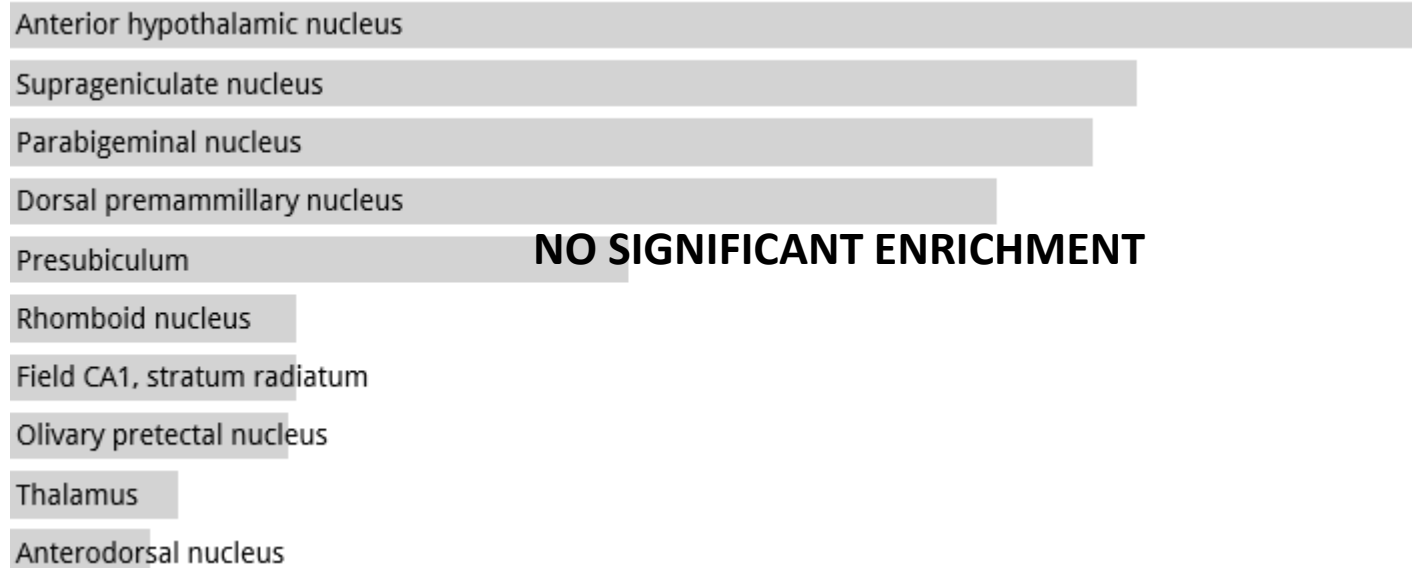
Septic Shock C0036983 human GSE9692 sample 307

West Nile fever DOID-2366 human GSE30719 sample 874

**NO SIGNIFICANT ENRICHMENT**

### Random Set 3

### Allen Brain Atlas up random 6000 genes



## Random Set 4

GASTRIC TISSUE (BULK)

PREFRONTAL CORTEX

COLON (BULK TISSUE)

LUNG (BULK TISSUE)

OMENTUM

NEURONAL EPITHELIUM

LIVER (BULK TISSUE)

ILEUM (BULK)

FETAL BRAIN CORTEX

SUPERIOR FRONTAL GYRUS

ARCHS4 Tissues random 6000 genes

**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 6847 genes

**NO DATA AVAILABLE**

## Random Set 4

## Jensen DISEASES random 6000 genes

Urinary\_schistosomiasis

Hermaphroditism

Zellweger\_syndrome

Latex\_allergy

Scimitar\_syndrome

Conversion\_disorder

Gastric\_lymphoma

Juvenile\_polyposis\_syndrome

Invasive\_lobular\_carcinoma

Acute\_promyelocytic\_leukemia

**NO SIGNIFICANT ENRICHMENT**

## Disease Perturbations from GEO up random 6000 genes

pancreatic ductal adenocarcinoma DOID-3498 mouse GSE53659 sample 699

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 772

amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 679

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 771

Cardiac Hypertrophy C1383860 rat GSE1055 sample 354

nemaline myopathy DOID-3191 mouse GSE3384 sample 974

nemaline myopathy DOID-3191 mouse GSE3384 sample 971

Nemaline Myopathy C0206157 mouse GSE3384 sample 317

Down syndrome DOID-14250 human GSE20910 sample 1063

idiopathic pulmonary fibrosis DOID-0050156 human GSE24206 sample 872

**NO SIGNIFICANT ENRICHMENT**

## Disease Perturbations from GEO down random 6000 genes

pulmonary tuberculosis DOID-2957 mouse GSE48027 sample 831

acute myocarditis DOID-3951 mouse GSE35182 sample 801

amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 679

nemaline myopathy DOID-3191 mouse GSE3384 sample 971

Purpura, Idiopathic Thrombocytopenic C0043117 human GSE574 sample 358

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 772

West Nile fever DOID-2366 human GSE30719 sample 874

Nemaline Myopathy C0206157 mouse GSE3384 sample 317

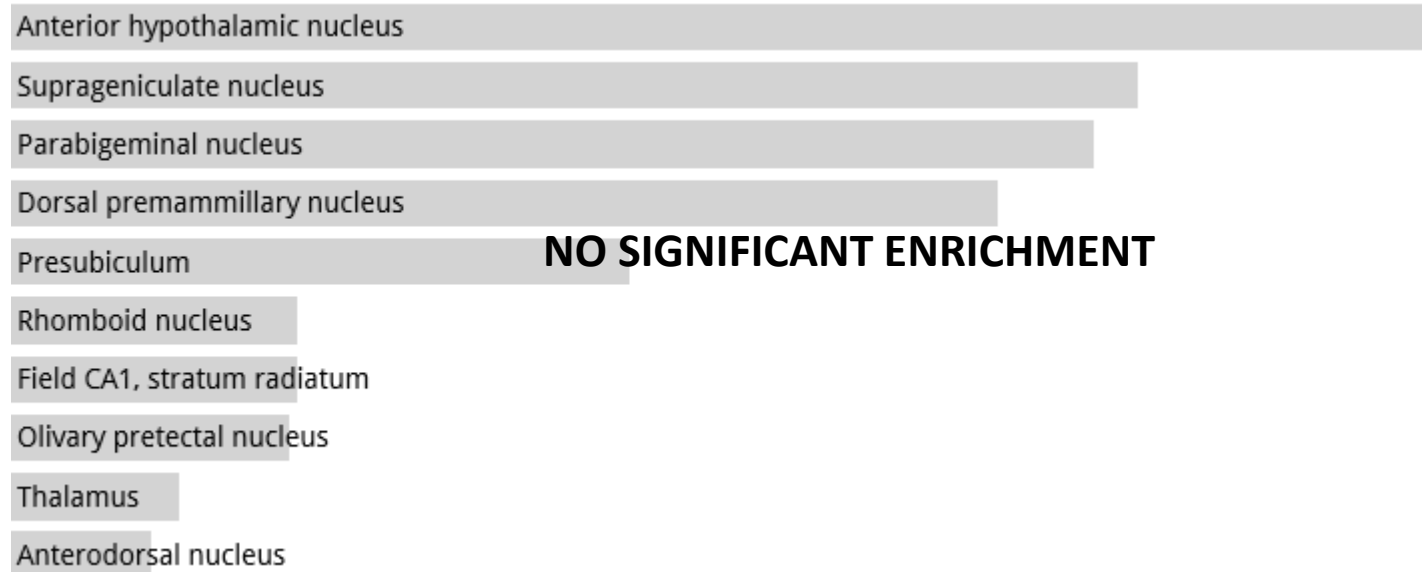
juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 771

acute myocardial infarction DOID-9408 mouse GSE775 sample 1000

**NO SIGNIFICANT ENRICHMENT**

## Random Set 4

## Allen Brain Atlas up random 6000 genes



## Random Set 5

ARCHS4 Tissues random 6000 genes

SMALL INTESTINE (BULK TISSUE)

VASCULAR SMOOTH MUSCLE

SKIN (BULK TISSUE)

THYROID (BULK TISSUE)

STROMAL CELL

RESPIRATORY SMOOTH MUSCLE

RENAL CORTEX

SKELETAL MUSCLE (BULK TISSUE)

REGULATORY T CELLS

PANCREATIC ISLET

**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 6847 genes

**NO DATA AVAILABLE**

## Random Set 5

## Jensen DISEASES random 6000 genes

Enthesopathy

Gyrate\_atrophy

Vitelliform\_macular\_dystrophy

Opisthorchiasis

Amelogenesis\_imperfecta

3-methylglutaconic\_aciduria

Mastitis

Cone\_dystrophy

Hydrocephalus

Fundus\_dystrophy

**NO SIGNIFICANT ENRICHMENT**

## Disease Perturbations from GEO up random 6000 genes

nemaline myopathy DOID-3191 mouse GSE3384 sample 974

Nemaline Myopathy C0206157 mouse GSE3384 sample 317

Depression DOID-1595 mouse GSE26025 sample 463

autistic disorder DOID-12849 human GSE6575 sample 1043

Retinal damage C0235272 rat GSE1001 sample 312 **NO SIGNIFICANT ENRICHMENT**

Retinitis Pigmentosa C0035334 mouse GSE128 sample 33

Alzheimer's disease DOID-10652 human GSE36980 sample 523

Alcohol poisoning C0392620 rat GSE3311 sample 288

non-systemic juvenile idiopathic arthritis (sJIA) (subgroup-RF+ polyarthritis) DOID-848 human GSE21521 sample

colitis DOID-0060180 mouse GSE34874 sample 800

## Disease Perturbations from GEO down random 6000 genes

polycystic ovary syndrome DOID-11612 human GSE48301 sample 558

glaucoma associated with systemic syndromes DOID-1686 mouse GSE26299 sample 491

MS (Multiple Sclerosis) C0026769 mouse GSE842 sample 381

breast adenocarcinoma DOID-3458 human GSE61304 sample 1071

Muscular Dystrophy C0026850 mouse GSE21521 sample 1043 **NO SIGNIFICANT ENRICHMENT**

Wolfram syndrome UMLS CUI-C0043207 mouse GSE15293 sample 695

amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 681

Huntington's disease DOID-12858 mouse GSE3248 sample 723

Duchenne muscular dystrophy (DMD) C0013264 mouse GSE466 sample 328

Emery-Dreifuss muscular dystrophy, autosomal recessive form DOID-11726 mouse GSE8000 sample 677



## Random Set 5

## Allen Brain Atlas up random 6000 genes

Presubiculum

Endopiriform nucleus, dorsal part

dorsal endopiriform nucleus

Copula pyramidis

bed nucleus of the external capsule

mantle zone of r7Co

Agranular insular area, posterior part, layer 6a

Entorhinal area, medial part, dorsal zone, layer 5

corticoid layer of TuStr

Infralimbic area, layer 6b

**NO SIGNIFICANT ENRICHMENT**

## Random Set 6

ARCHS4 Tissues random 6000 genes

**NO DATA AVAILABLE**

Jensen TISSUES random 6000 genes

**NO DATA AVAILABLE**

**Random Set 6**

Jensen DISEASES random 6000 genes

**NO DATA AVAILABLE**

Disease Perturbations from GEO up random 6000 genes

**NO DATA AVAILABLE**

Disease Perturbations from GEO down random 6000 genes

Systemic lupus erythematosus (SLE) C0024141 human GSE12374 sample 123

**NO SIGNIFICANT ENRICHMENT**

**Random Set 6**

Allen Brain Atlas up random 6000 genes

**NO DATA AVAILABLE**

## Random Set 7

ARCHS4 Tissues random 6847 genes

LYMPHOCYTE

OMENTUM

GASTRIC TISSUE (BULK)

PREFRONTAL CORTEX

SPINAL CORD (BULK)

SPINAL CORD

LUNG (BULK TISSUE)

CARDIAC MUSCLE FIBER

BREAST (BULK TISSUE)

NEURONAL EPITHELIUM

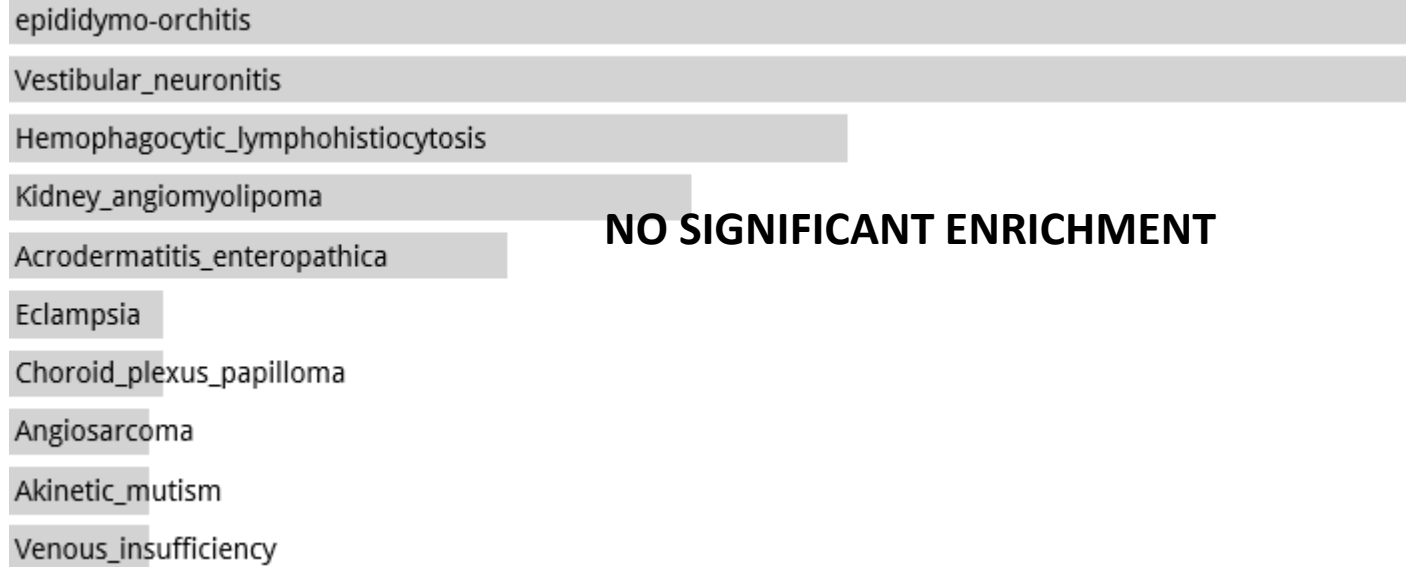
**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 6847 genes

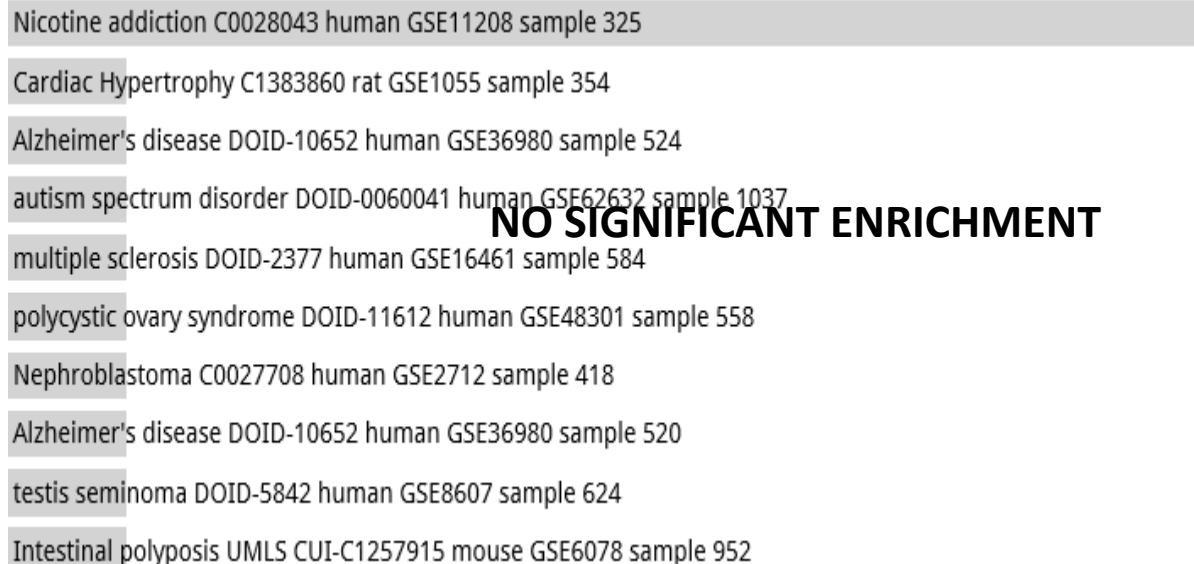
**NO DATA AVAILABLE**

## Random Set 7

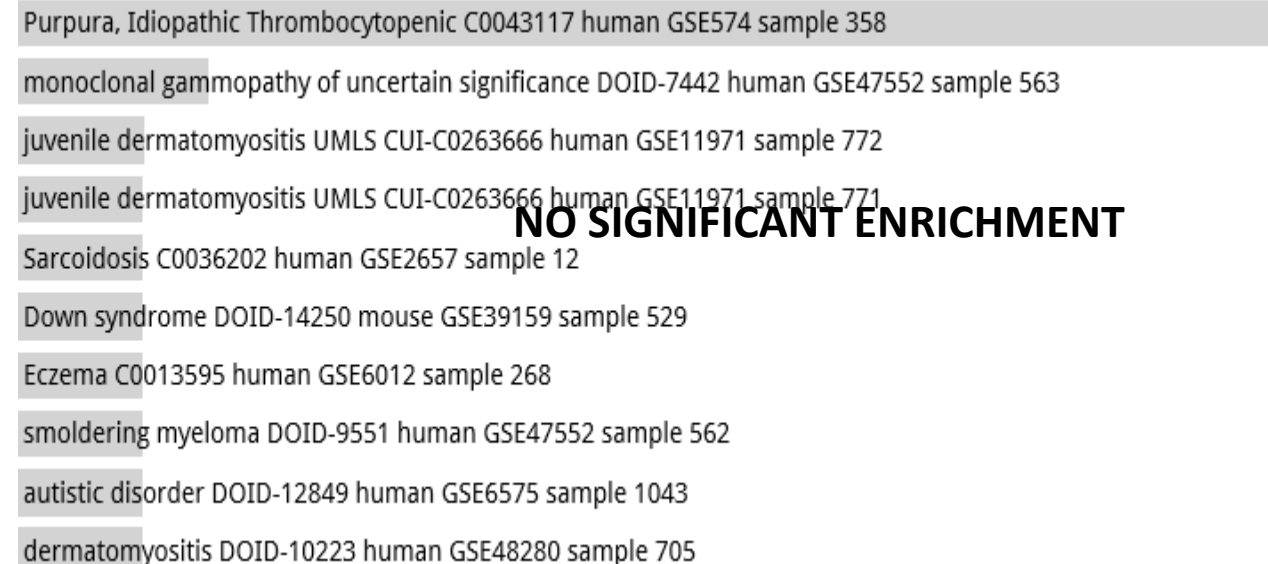
## Jensen DISEASES random 6847 genes



## Disease Perturbations from GEO up random 6000 genes

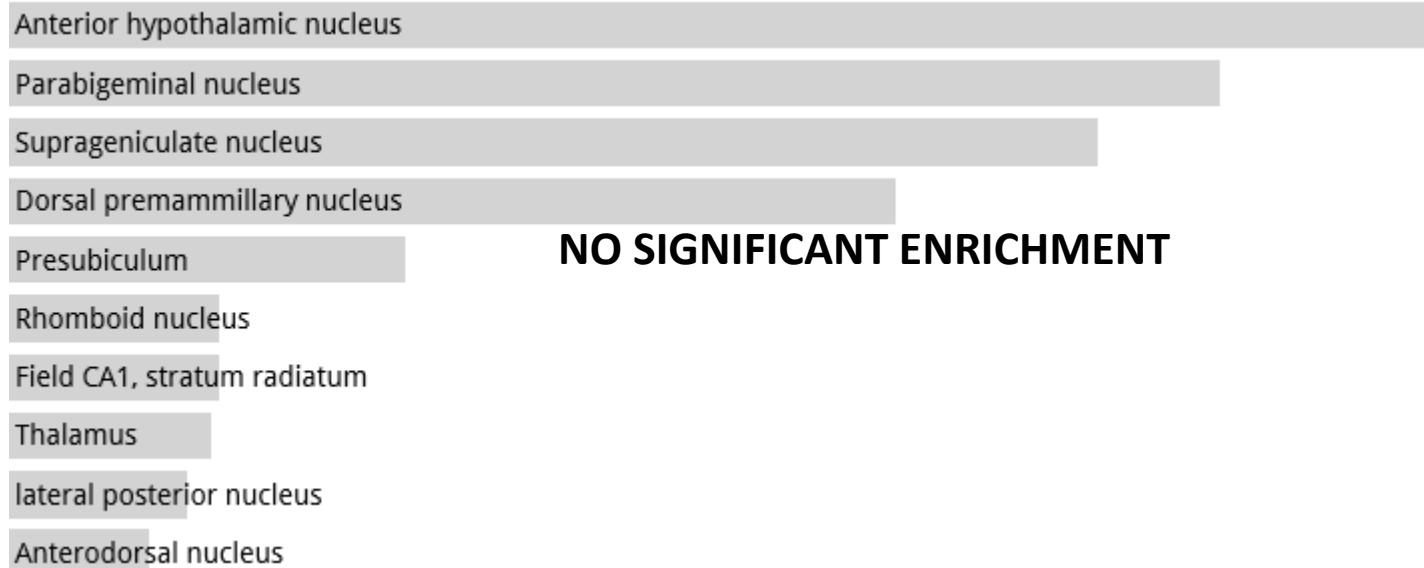


## Disease Perturbations from GEO down random 6000 genes



## Random Set 7

## Allen Brain Atlas up random 6847 genes



## Random Set 8

ARCHS4 Tissues random 6001 genes

GASTRIC TISSUE (BULK)

SKIN (BULK TISSUE)

RENAL CORTEX

KIDNEY (BULK TISSUE)

PREFRONTAL CORTEX

PANCREATIC ISLET

SPINAL CORD (BULK)

SMALL INTESTINE (BULK TISSUE)

COLON (BULK TISSUE)

SPINAL CORD

**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 6001 genes

**NO DATA AVAILABLE**



## Random Set 8

## Jensen DISEASES random 6001 genes

Early\_myoclonic\_encephalopathy

Salpingitis

Vaginitis

cold-induced\_sweating\_syndrome

Parotitis

Carnitine\_palmitoyltransferase\_II\_deficiency **NO SIGNIFICANT ENRICHMENT**

Poliomyelitis

Factor\_XI\_deficiency

cytochrome-c\_oxidase\_deficiency\_disease

Ascariasis

## Disease Perturbations from GEO up random 6000 genes

pancreatic ductal adenocarcinoma DOID-3498 mouse GSE53659 sample 699

Schistosomiasis C0036323 mouse GSE19525 sample 439

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 772

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 771

amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 679

pancreatic ductal adenocarcinoma DOID-3498 human GSE15471 sample 604

nemaline myopathy DOID-3191 mouse GSE3384 sample 974

asthma DOID-2841 human GSE31773 sample 714

glaucoma associated with systemic syndromes DOID-1686 mouse GSE26299 sample 488

Congestive heart disease C0018802 mouse GSE2236 sample 258

**NO SIGNIFICANT ENRICHMENT**

## Disease Perturbations from GEO down random 6000 genes

acute myocarditis DOID-3951 mouse GSE35182 sample 801

pulmonary tuberculosis DOID-2957 mouse GSE48027 sample 831

autistic disorder DOID-12849 human GSE6575 sample 1043

Purpura, Idiopathic Thrombocytopenic C0043117 human GSE574 sample 358

amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 679

Nicotine addiction C0028043 human GSE11208 sample 325

sickle-cell anemia DOID-10923 human GSE16728 sample 505

polycystic ovary syndrome DOID-11612 human GSE48301 sample 558

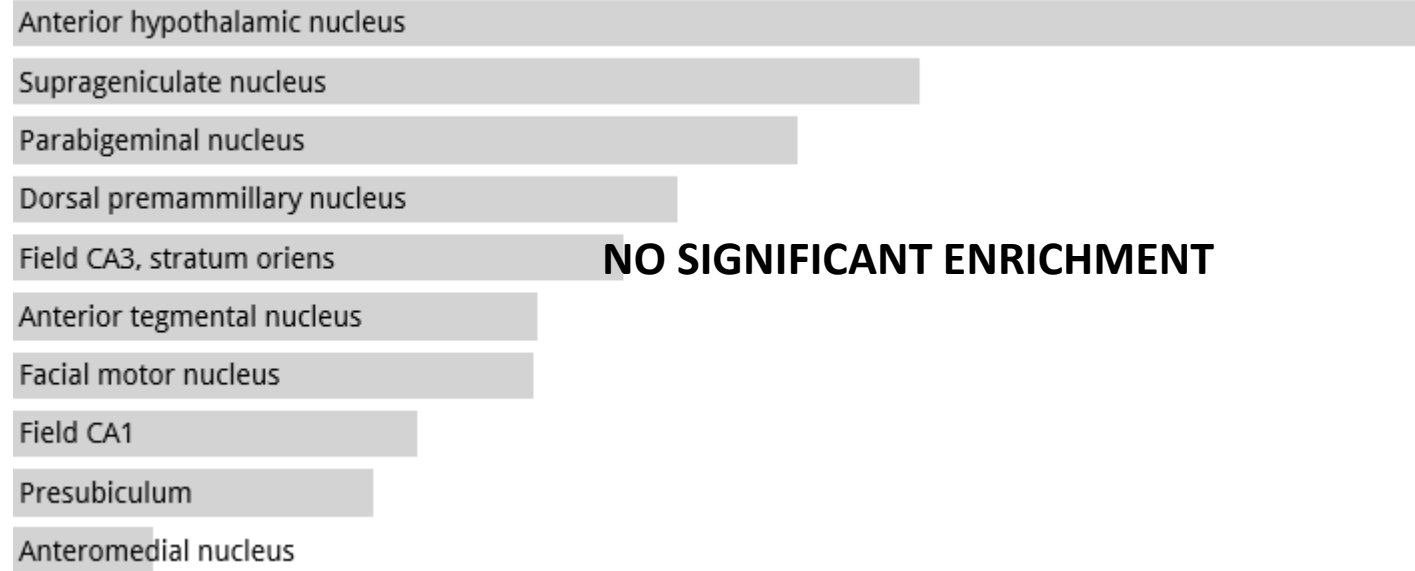
autism spectrum disorder DOID-0060041 human GSE62632 sample 1037

Cardiac Hypertrophy C1383860 rat GSE1055 sample 354

**NO SIGNIFICANT ENRICHMENT**

## Random Set 8

## Allen Brain Atlas up random 6001 genes



## Random Set 9

ARCHS4 Tissues random 6000 genes

COLONIC MUCOSA

PREFRONTAL CORTEX

VENTRICLE

ASTROCYTE

MYOBLAST

BRAIN (BULK)

FETAL BRAIN CORTEX

NEURONAL EPITHELIUM

ATRIUM

ADIPOSE (BULK TISSUE)

**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 6000 genes

**NO DATA AVAILABLE**

## Random Set 9

## Jensen DISEASES random 6000 genes

Hereditary\_mucosal\_leukokeratosis

Dumping\_syndrome

X-linked\_endothelial\_corneal\_dystrophy

Clouston\_syndrome

Enlarged\_vestibular\_aqueduct

Oculodentodigital\_dysplasia

Porphyria\_cutanea\_tarda

MEDNIK\_syndrome

Granulomatous\_amebic\_encephalitis

Pachyonychia\_congenita

**NO SIGNIFICANT ENRICHMENT**

**NO SIGNIFICANT ENRICHMENT**

## Disease Perturbations from GEO up random 6000 genes

Dental cavity, complex C0399396 human GSE1629 sample 175

actinic keratosis DOID-8866 human GSE2503 sample 628

Actinic keratosis C0022602 human GSE2503 sample 350

acute myeloid leukemia DOID-9119 human GSE9476 sample 782

Epilepsy C0014544 human GSE7486 sample 417

epidermolysis bullosa simplex DOID-4644 human GSE28315 sample 711

lupus erythematosus DOID-8857 human GSE30153 sample 739

cardiomyopathy DOID-0050700 human GSE9128 sample 781

type 2 diabetes mellitus DOID-9352 human GSE23343 sample 895

autism spectrum disorder DOID-0060041 human GSE25507 sample 1032

**NO SIGNIFICANT ENRICHMENT**

## Disease Perturbations from GEO down random 6000 genes

multiple myeloma DOID-9538 human GSE6691 sample 787

Smoldering multiple myeloma C1531608 human GSE5900 sample 404

autistic disorder DOID-12849 human GSE6575 sample 1043

acute myocarditis DOID-3951 mouse GSE35182 sample 801

pulmonary tuberculosis DOID-2957 mouse GSE48027 sample 831

autistic disorder DOID-12849 human GSE6575 sample 1042

mental retardation DOID-1059 human GSE6575 sample 1044

Septic Shock C0036983 human GSE9692 sample 307

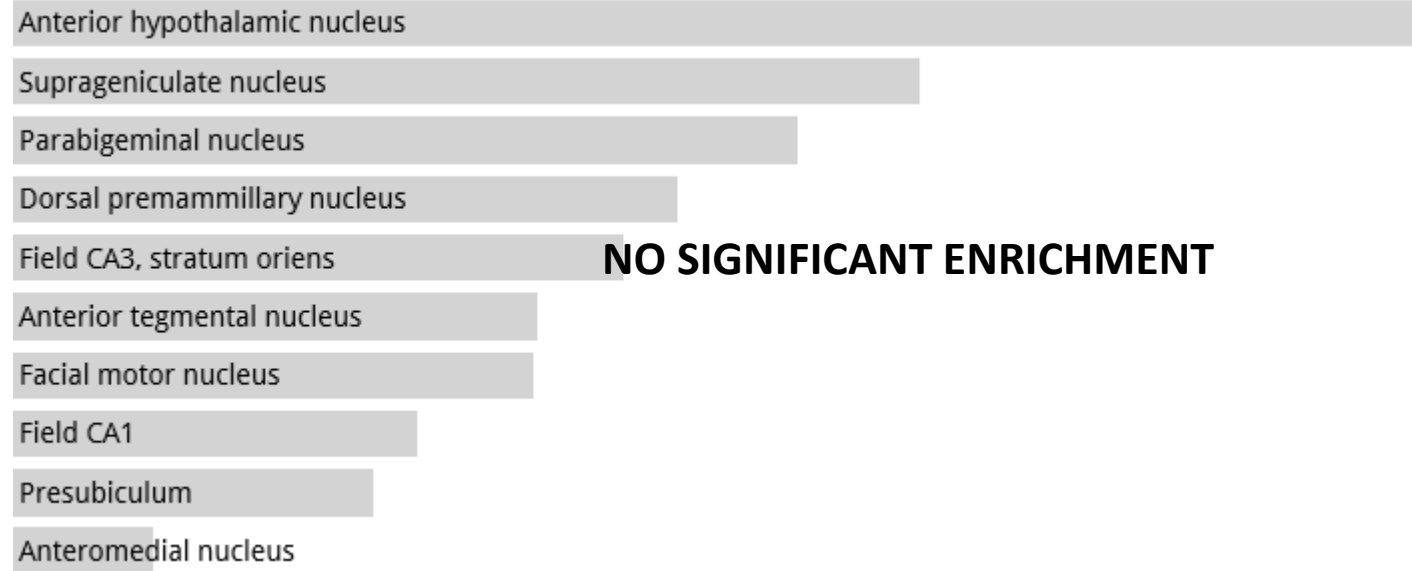
West Nile fever DOID-2366 human GSE30719 sample 874

Waldenstrom Macroglobulinemia UMLS CUI-C0024419 human GSE6691 sample 785

**NO SIGNIFICANT ENRICHMENT**

## Random Set 9

## Allen Brain Atlas up random 6000 genes



## Random Set 10

ARCHS4 Tissues random 6000 genes

GASTRIC TISSUE (BULK)

PREFRONTAL CORTEX

OMENTUM

COLON (BULK TISSUE)

LUNG (BULK TISSUE)

NEURONAL EPITHELIUM

LIVER (BULK TISSUE)

FETAL BRAIN CORTEX

RENAL CORTEX

AMNIOTIC FLUID

**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 6000 genes

**NO DATA AVAILABLE**

## Random Set 10

## Jensen DISEASES random 6000 genes

Laryngeal\_squamous\_cell\_carcinoma

Urinary\_schistosomiasis

Ebola\_hemorrhagic\_fever

Hermaphroditism

Irritant\_dermatitis

Zellweger\_syndrome

Aortic\_valve\_insufficiency

Scimitar\_syndrome

Gastric\_lymphoma

Multiple\_chemical\_sensitivity

**NO SIGNIFICANT ENRICHMENT**

## Disease Perturbations from GEO up random 6000 genes

Alzheimer's disease DOID-10652 human GSE36980 sample 519

polycystic ovary syndrome DOID-11612 human GSE48301 sample 558

Cardiac Hypertrophy C1383860 rat GSE1055 sample 354

ulcerative colitis DOID-8577 human GSE37283 sample 594

pancreatic ductal adenocarcinoma DOID-3498 mouse GSE53659 sample 699

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 772

amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 679

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 771

nemaline myopathy DOID-3191 mouse GSE3384 sample 974

nemaline myopathy DOID-3191 mouse GSE3384 sample 971

**NO SIGNIFICANT ENRICHMENT**

## Disease Perturbations from GEO down random 6000 genes

pulmonary tuberculosis DOID-2957 mouse GSE48027 sample 831

acute myocarditis DOID-3951 mouse GSE35182 sample 801

amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 679

nemaline myopathy DOID-3191 mouse GSE3384 sample 971

Down syndrome DOID-14250 mouse GSE39159 sample 529

Cardiomyopathy, Dilated C0007193 human GSE3586 sample 323

Purpura, Idiopathic Thrombocytopenic C0043117 human GSE574 sample 358

acute myocardial infarction DOID-9408 mouse GSE775 sample 1000

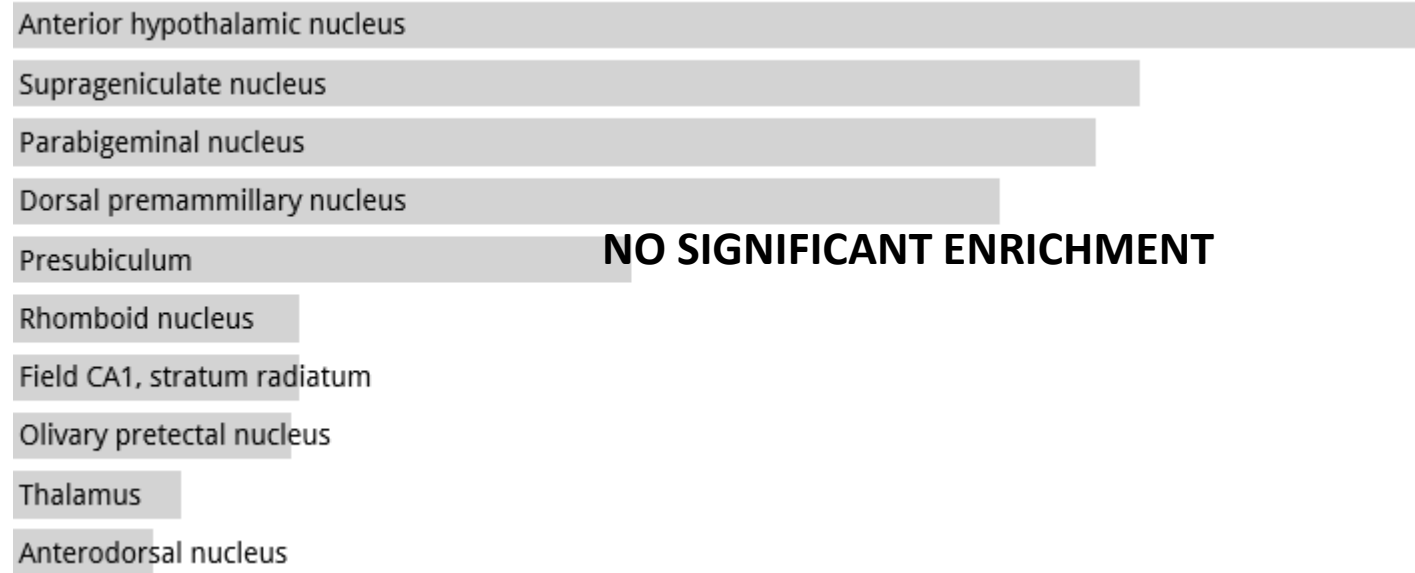
papillary thyroid carcinoma DOID-3969 human GSE54958 sample 652

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 772

**NO SIGNIFICANT ENRICHMENT**

## Random Set 10

## Allen Brain Atlas up random 6000 genes





## Random Set 11

ARCHS4 Tissues random 4000 genes

OMENTUM

VENTRICLE

SKIN (BULK TISSUE)

WHARTONS JELLY

VASCULAR SMOOTH MUSCLE

OVARY (BULK TISSUE)

CEREBELLUM

CHONDROCYTE

THYROID (BULK TISSUE)

SPINAL CORD (BULK)

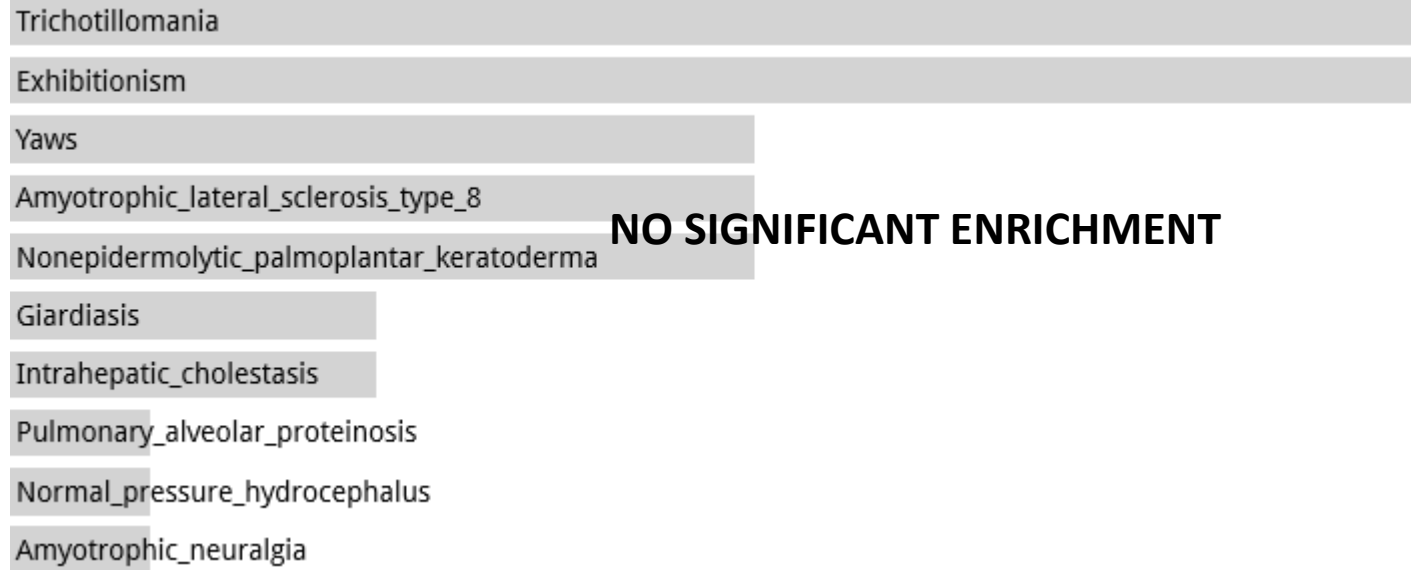
**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 4000 genes

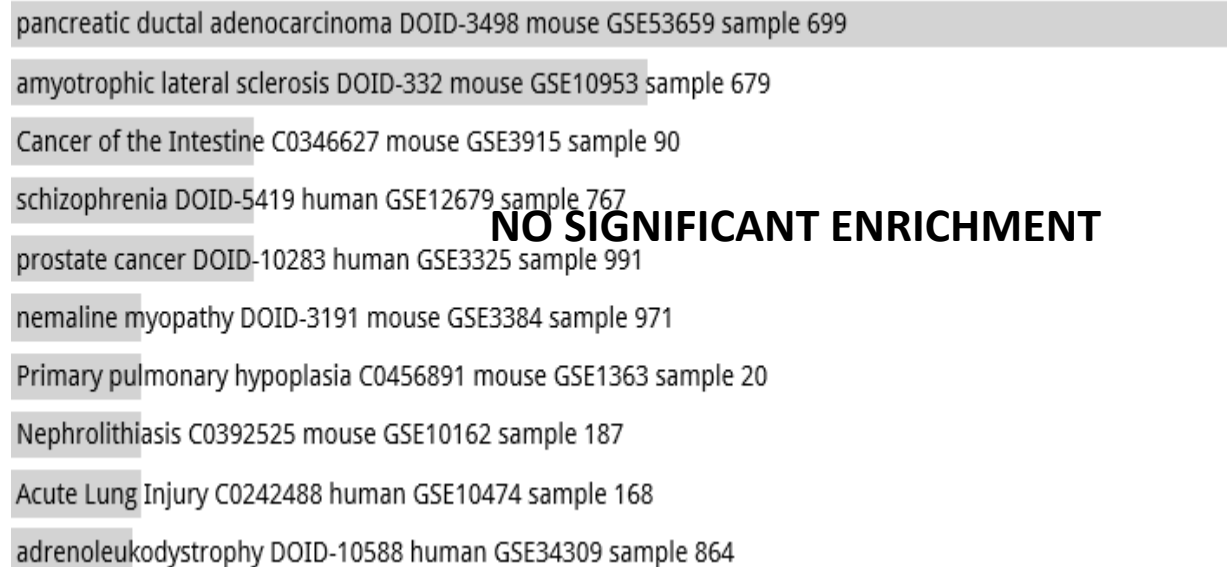
**NO DATA AVAILABLE**

## Random Set 11

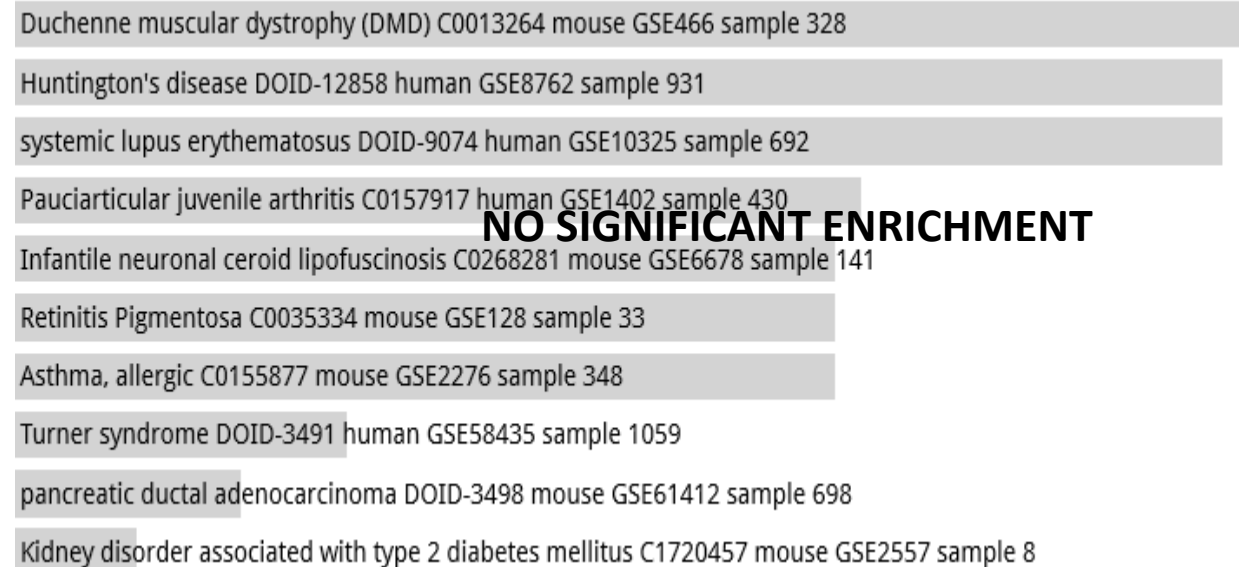
## Jensen DISEASES random 4000 genes



## Disease Perturbations from GEO up random 4000 genes

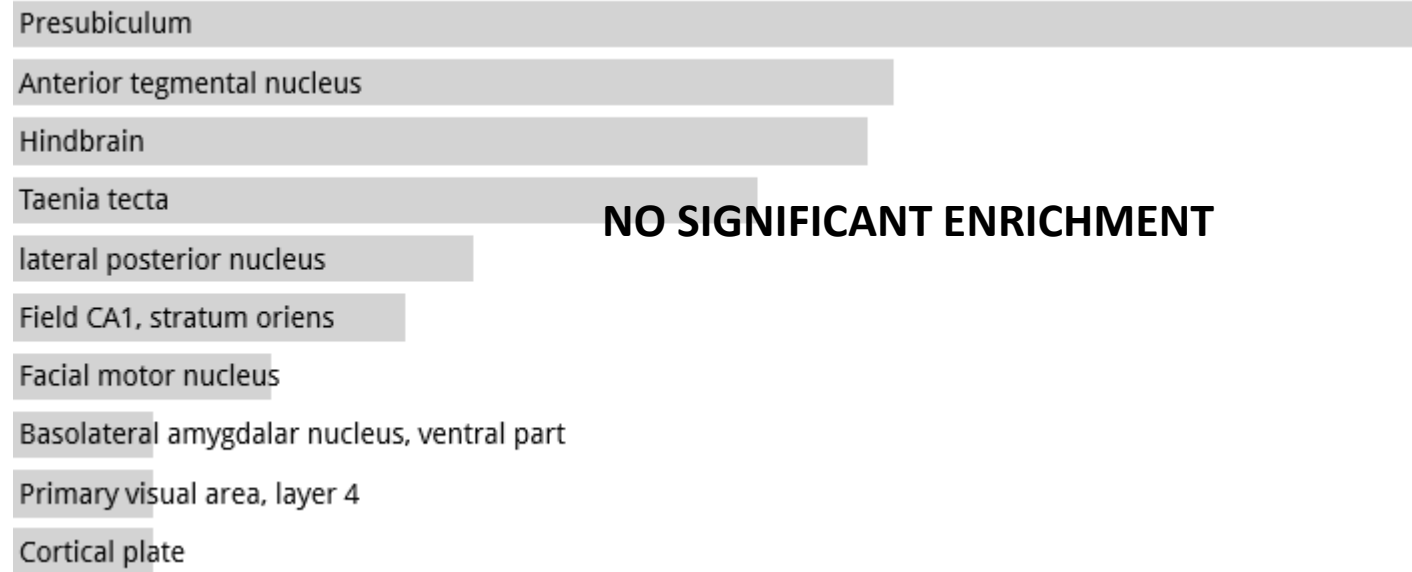


## Disease Perturbations from GEO down random 4000 genes



## Random Set 11

## Allen Brain Atlas up random 4000 genes



## Random Set 12

ARCHS4 Tissues random 4000 genes

GASTRIC TISSUE (BULK)

RENAL CORTEX

SKIN (BULK TISSUE)

SMALL INTESTINE (BULK TISSUE)

PANCREATIC ISLET

PREFRONTAL CORTEX

KIDNEY (BULK TISSUE)

SPINAL CORD (BULK)

PLACENTA (BULK)

SPINAL CORD

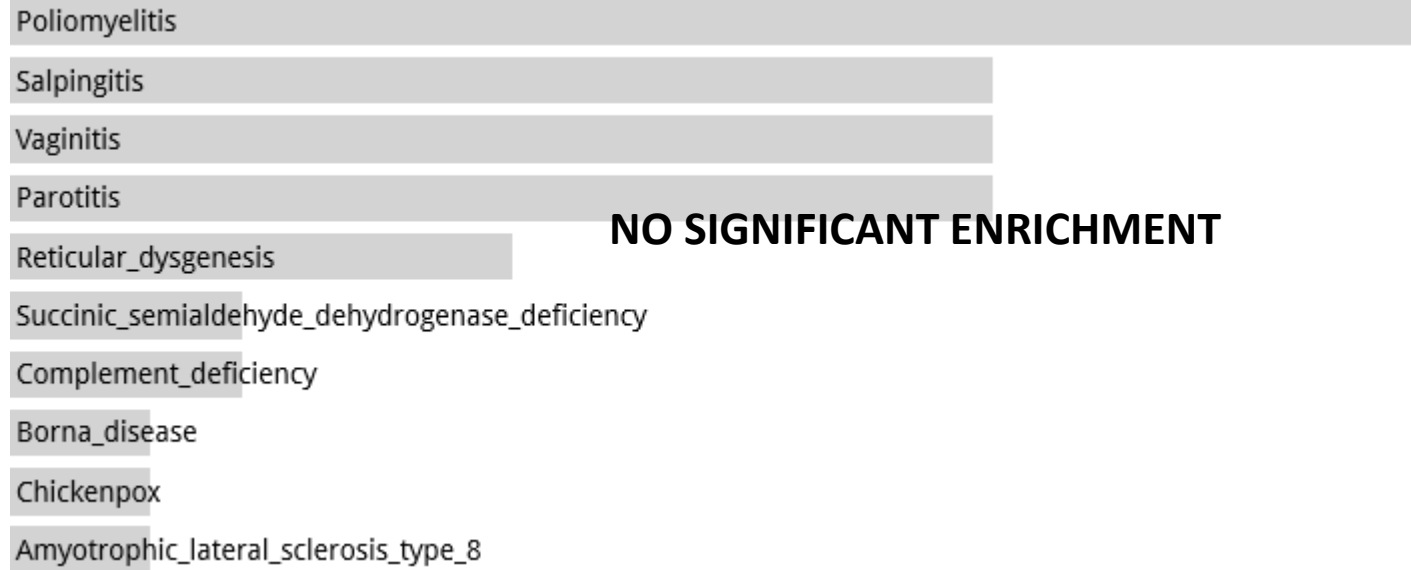
**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 4000 genes

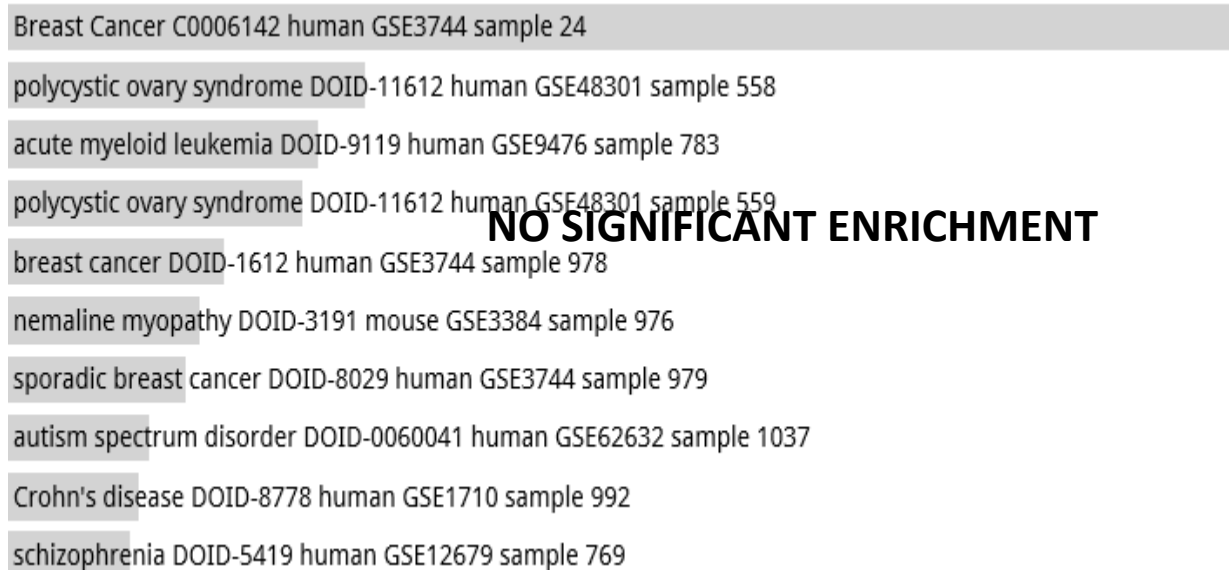
**NO DATA AVAILABLE**

## Random Set 12

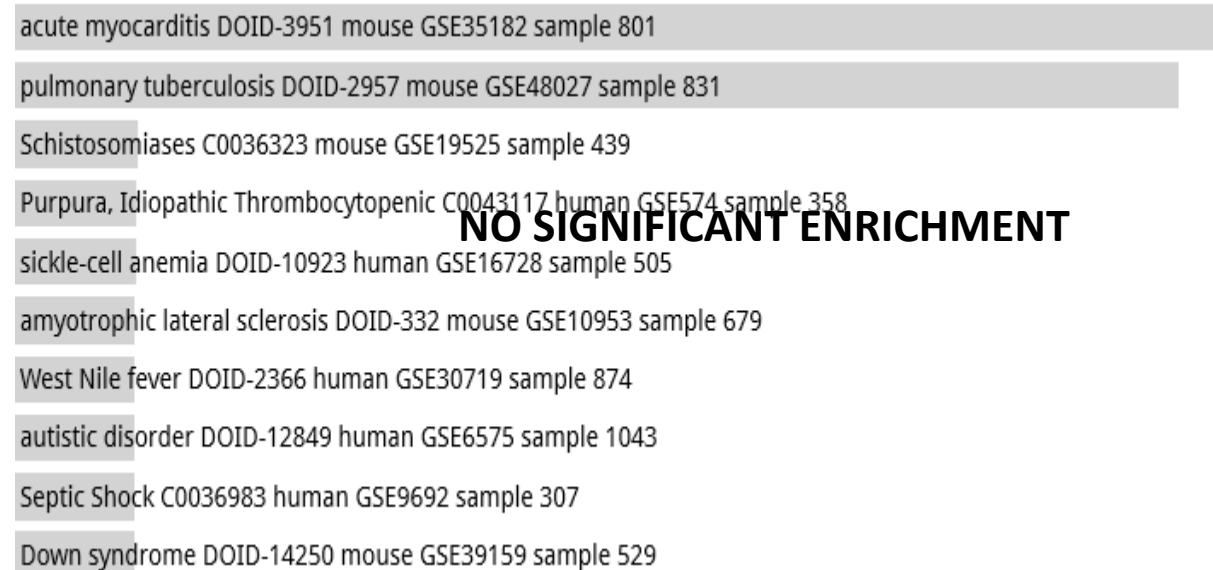
## Jensen DISEASES random 4000 genes



## Disease Perturbations from GEO up random 4000 genes



## Disease Perturbations from GEO down random 4000 genes



## Random Set 12

## Allen Brain Atlas up random 4000 genes

dorsal juxtacommissural pretectal nucleus

periventricular stratum of JcPL

JcPL part of the periaqueductal gray

periventricular stratum of PcPL

mantle zone of JcPD

PcPL part of the periaqueductal gray

intermediate stratum of JcPD

dorsal part of JcP

CoPV part of the periaqueductal gray

periventricular stratum of CoPV

**NO SIGNIFICANT ENRICHMENT**

## Random Set 13

ARCHS4 Tissues random 4000 genes

COLON (BULK TISSUE)

ILEUM (BULK)

SMALL INTESTINE (BULK TISSUE)

GASTRIC TISSUE (BULK)

OMENTUM

NEURONAL EPITHELIUM

SKIN (BULK TISSUE)

CEREBELLUM

PREFRONTAL CORTEX

MYOBLAST

**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 4000 genes

**NO DATA AVAILABLE**

## Random Set 13

## Jensen DISEASES random 4000 genes

cytochrome-c\_oxidase\_deficiency\_disease

Cockayne\_syndrome

Arterial\_tortuosity\_syndrome

cold-induced\_sweating\_syndrome

**NO SIGNIFICANT ENRICHMENT**

Esotropia

Xeroderma\_pigmentosum

Favism

Placenta\_praevia

Carnitine\_palmitoyltransferase\_II\_deficiency

Patent\_foramen\_ovale

## Disease Perturbations from GEO up random 4000 genes

autism spectrum disorder DOID-0060041 human GSE62632 sample 1037

Severe acute respiratory syndrome (SARS) C1175175 human GSE1739 sample 165

Nicotine addiction C0028043 human GSE11208 sample 325

ulcerative colitis DOID-8577 human GSE37283 sample 594

**NO SIGNIFICANT ENRICHMENT**

COPD - Chronic obstructive pulmonary disease C0024117 human GSE3320 sample 255

asthma DOID-2841 human GSE43696 sample 827

Idiopathic fibrosing alveolitis C0085786 human GSE10921 sample 272

Cancer of the testis C0153594 human GSE1818 sample 246

bipolar disorder DOID-3312 human GSE5392 sample 573

Ulcerative Colitis C0009324 human GSE1710 sample 264

## Disease Perturbations from GEO down random 4000 genes

autistic disorder DOID-12849 human GSE6575 sample 1043

Purpura, Idiopathic Thrombocytopenic C0043117 human GSE574 sample 358

pulmonary tuberculosis DOID-2957 mouse GSE48027 sample 831

sickle-cell anemia DOID-10923 human GSE16728 sample 506

**NO SIGNIFICANT ENRICHMENT**

acute myocarditis DOID-3951 mouse GSE35182 sample 802

amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 679

Septic Shock C0036983 human GSE9692 sample 307

Down syndrome DOID-14250 mouse GSE39159 sample 529

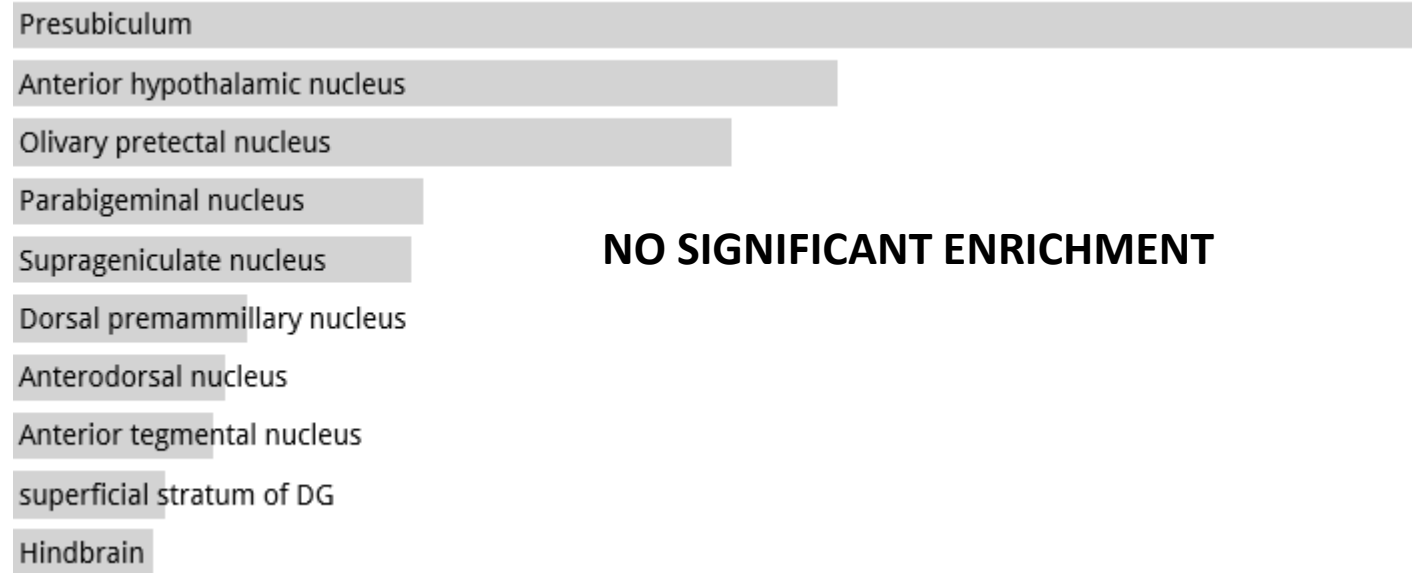
juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 772

amyotrophic lateral sclerosis DOID-332 human GSE833 sample 1007



## Random Set 13

## Allen Brain Atlas up random 4000 genes



## Random Set 14

ARCHS4 Tissues random 4000 genes

COLONIC MUCOSA

p-value: 0.0048; Adjusted p-value: 0.5188

CD19+ B CELLS

BLYMPHOCYTE

ILEUM (BULK)

**NO SIGNIFICANT ENRICHMENT**

COLON (BULK TISSUE)

PLASMA CELL

GASTRIC TISSUE (BULK)

BLASTOCYST

PERIPHERAL BLOOD

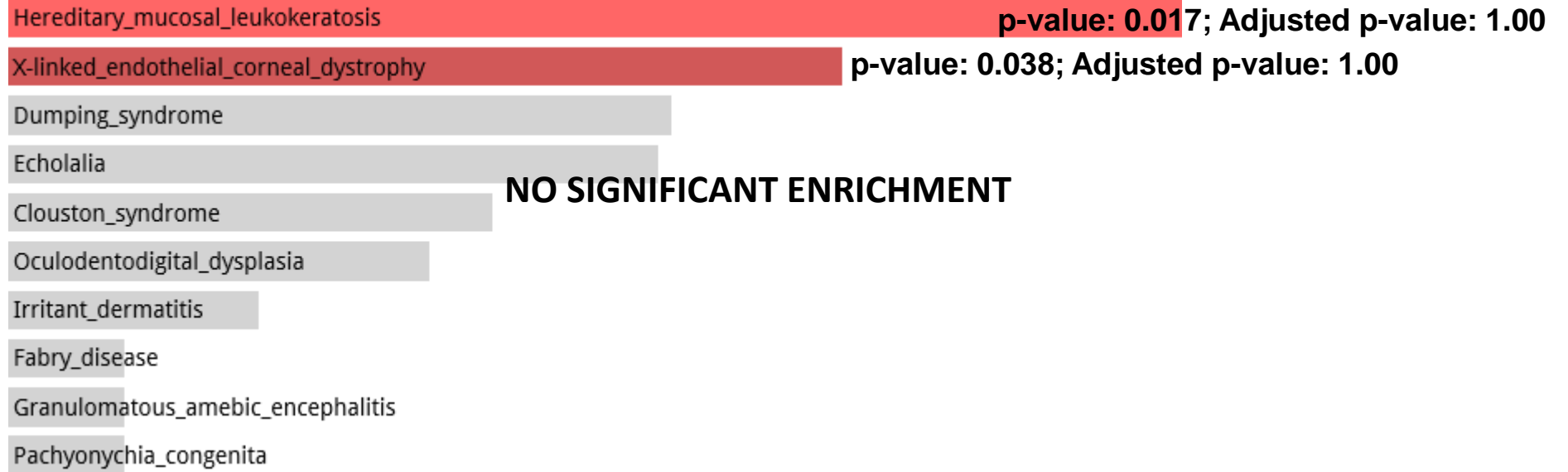
SPLEEN (BULK TISSUE)

Jensen TISSUES random 4000 genes

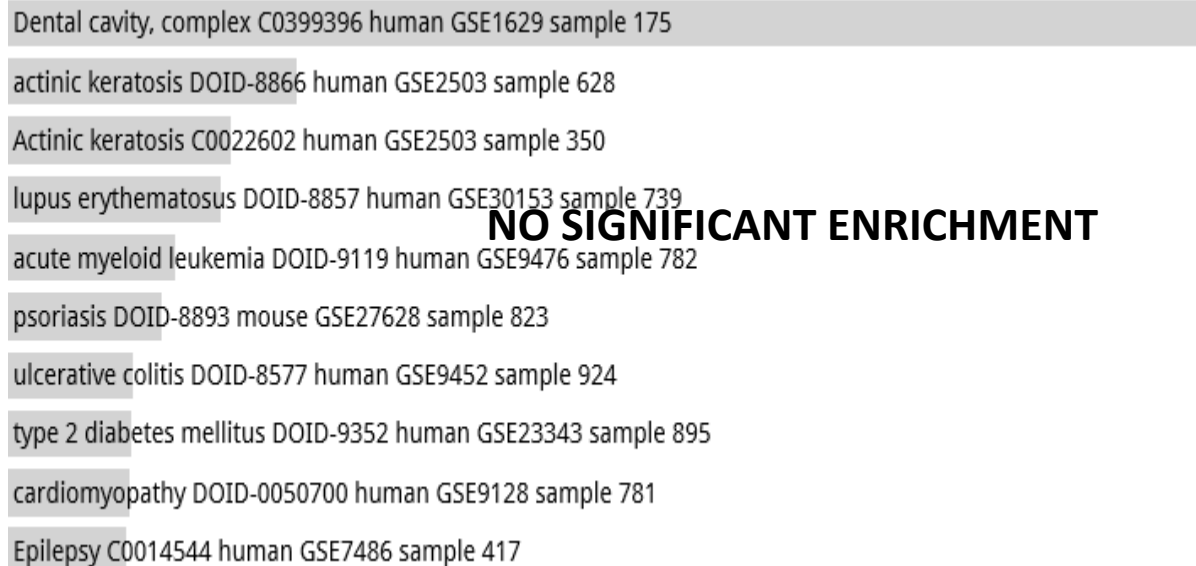
**NO DATA AVAILABLE**

## Random Set 14

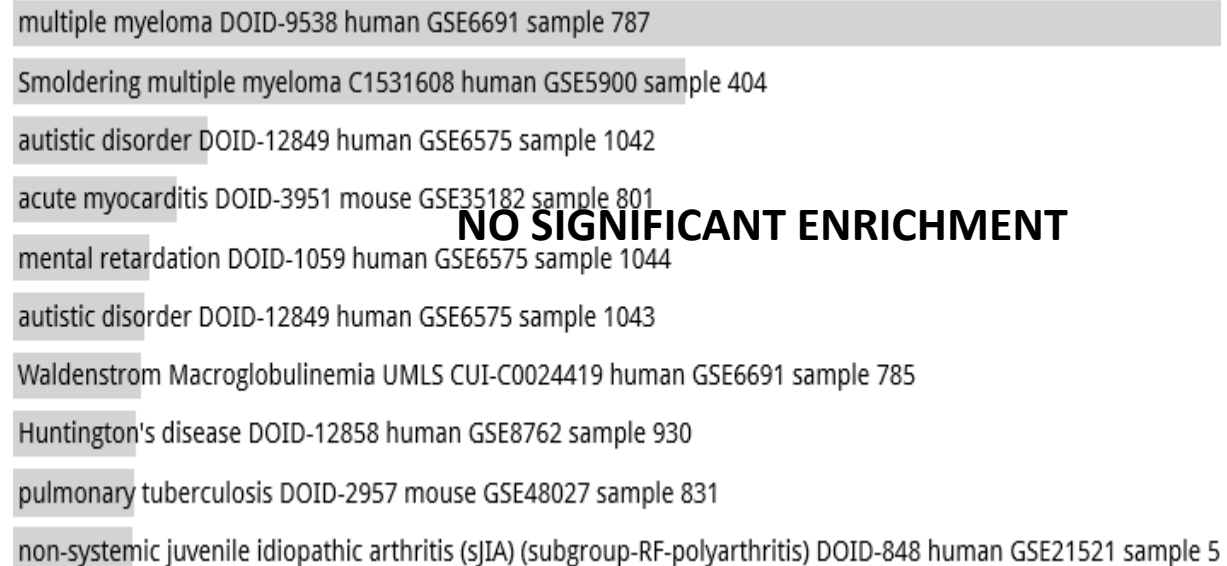
## Jensen DISEASES random 4000 genes



## Disease Perturbations from GEO up random 4000 genes



## Disease Perturbations from GEO down random 4000 genes



## Random Set 14

## Allen Brain Atlas up random 4000 genes

periventricular stratum of the VAP  
Nucleus of the solitary tract, lateral part  
Lateral amygdalar nucleus  
intermediate stratum of r4Ve  
r9 part of the vestibular column  
Lateral reticular nucleus, parvicellular part  
mantle zone of r9Ve  
r4 part of magnocellular medial vestibular nucleus  
periventricular stratum of r8Tr  
superficial stratum of r7Lim

**NO SIGNIFICANT ENRICHMENT**

## Random Set 15

COLON (BULK TISSUE)

GASTRIC TISSUE (BULK)

PREFRONTAL CORTEX

ILEUM (BULK)

LUNG (BULK TISSUE)

NEURONAL EPITHELIUM

OMENTUM

LIVER (BULK TISSUE)

AMNIOTIC FLUID

RENAL CORTEX

ARCHS4 Tissues random 4000 genes

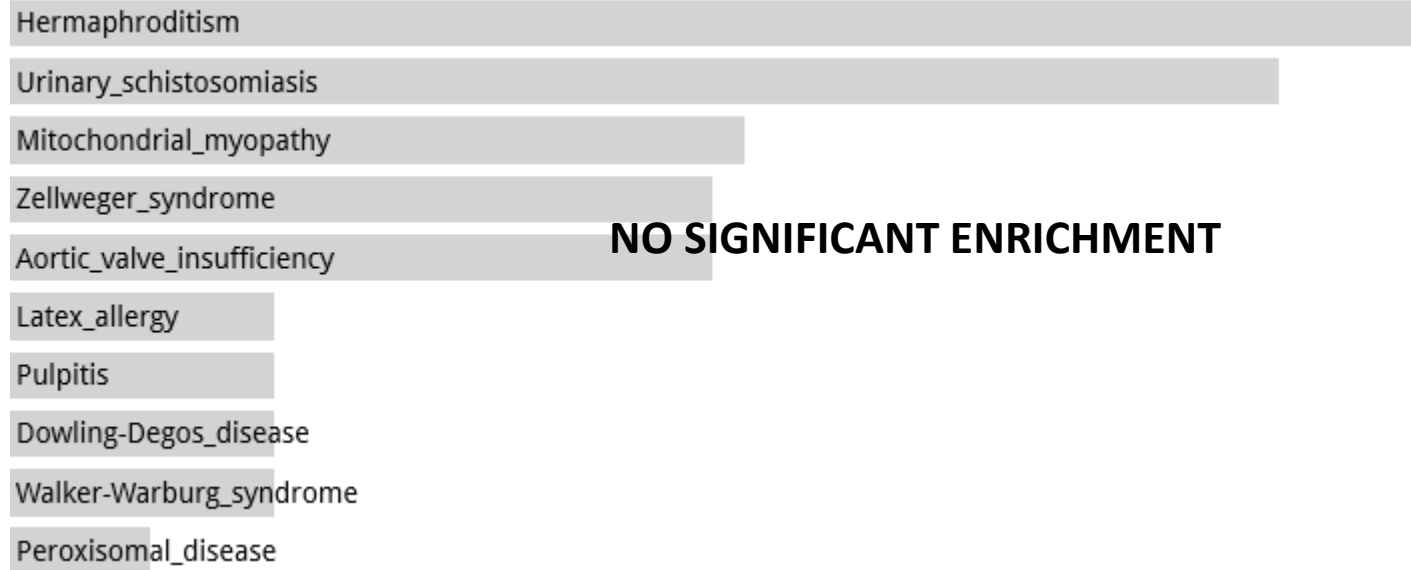
**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 4000 genes

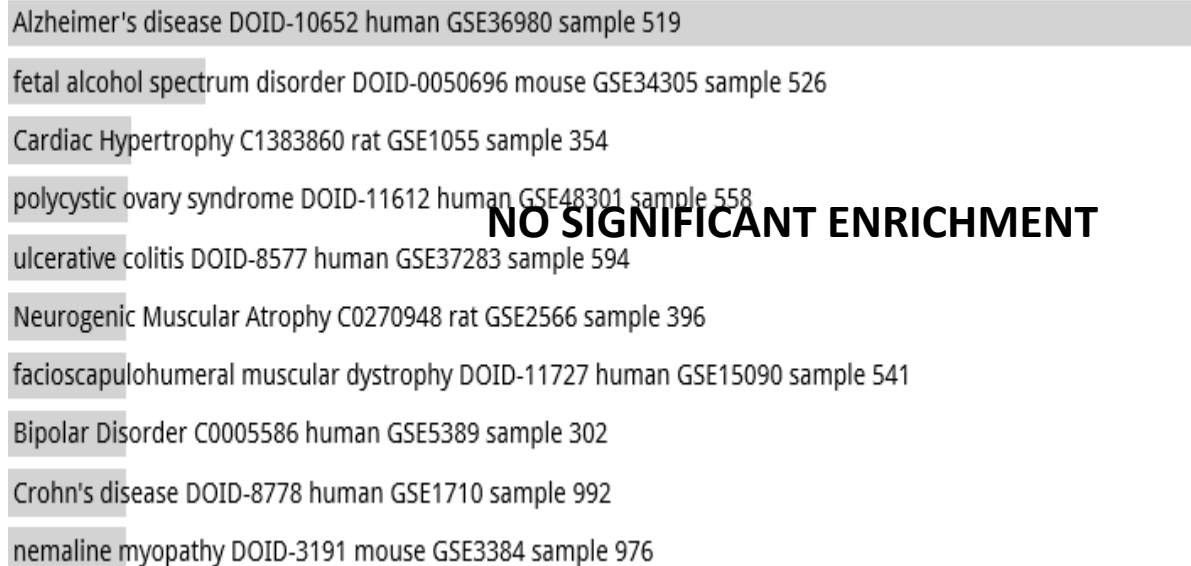
**NO DATA AVAILABLE**

## Random Set 15

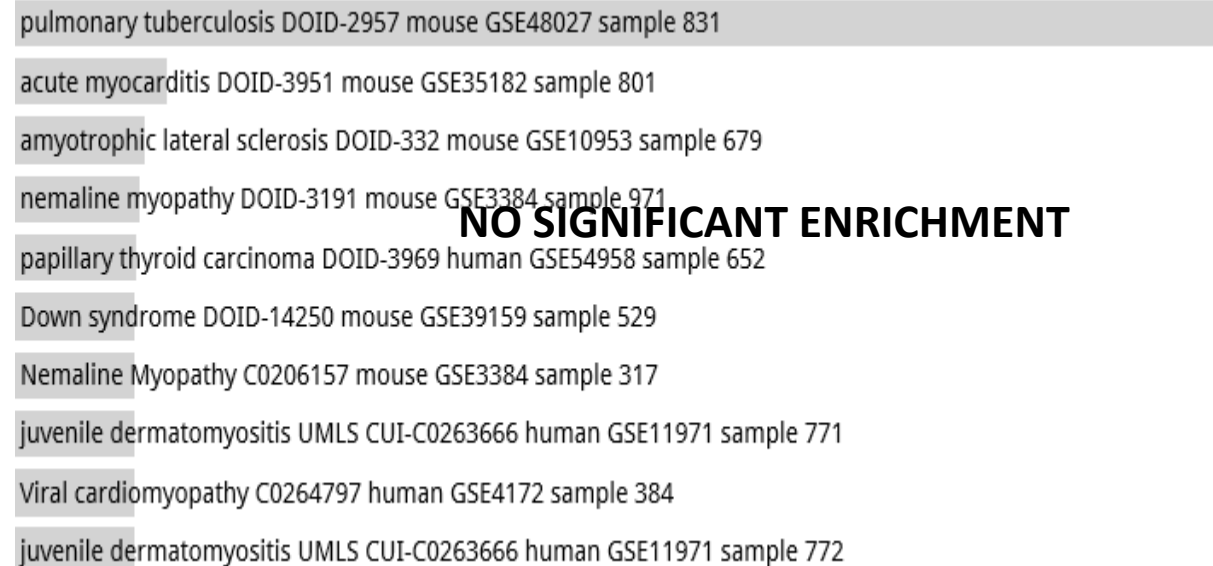
## Jensen DISEASES random 4000 genes



## Disease Perturbations from GEO up random 4000 genes

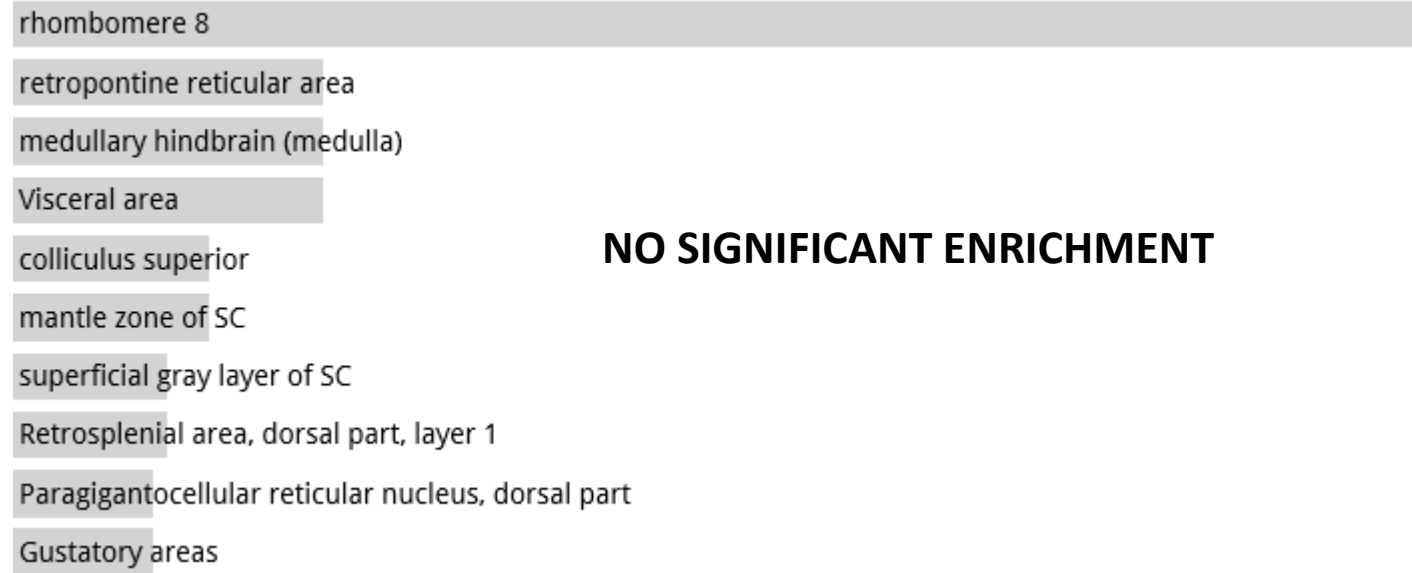


## Disease Perturbations from GEO down random 4000 genes



## Random Set 15

## Allen Brain Atlas up random 4000 genes



## Random Set 16

ARCHS4 Tissues random 4000 genes

PREFRONTAL CORTEX

SPINAL CORD (BULK)

SUPERIOR FRONTAL GYRUS

OMENTUM

SPINAL CORD

LUNG (BULK TISSUE)

NEURONAL EPITHELIUM

FETAL BRAIN CORTEX

ASTROCYTE

BRAIN (BULK)

**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 4000 genes

**NO DATA AVAILABLE**



## Random Set 16

## Jensen DISEASES random 4000 genes

Scimitar\_syndrome

Western\_equine\_encephalitis

Opitz-GBBB\_syndrome

Conversion\_disorder

Gastric\_lymphoma

Cutaneous\_porphyria

Leiomyomatosis

Epididymitis

DNA\_ligase\_IV\_deficiency

Juvenile\_polyposis\_syndrome

**NO SIGNIFICANT ENRICHMENT**

## Disease Perturbations from GEO up random 4000 genes

pancreatic ductal adenocarcinoma DOID-3498 mouse GSE53659 sample 699

amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 679

autism spectrum disorder DOID-0060041 human GSE62632 sample 1037

RA (rheumatoid arthritis) C0003873 human GSE3592 sample 183

schizophrenia DOID-5419 human GSE27383 sample 548

Cancer of the Intestine C0346627 mouse GSE3915 sample 90

schizophrenia DOID-5419 human GSE12679 sample 767

Duchenne muscular dystrophy (DMD) C0013264 mouse GSE1472 sample 62

swine influenza DOID-0050211 human GSE48466 sample 498

psoriasis DOID-8893 human GSE26952 sample 982

**NO SIGNIFICANT ENRICHMENT**

## Disease Perturbations from GEO down random 4000 genes

pulmonary tuberculosis DOID-2957 mouse GSE48027 sample 831

Purpura, Idiopathic Thrombocytopenic C0043117 human GSE574 sample 358

West Nile fever DOID-2366 human GSE30719 sample 874

amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 679

Down syndrome DOID-14250 mouse GSE39159 sample 529

pancreatic ductal adenocarcinoma DOID-3498 mouse GSE53659 sample 699

Cardiomyopathy, Dilated C0007193 human GSE3586 sample 323

nemaline myopathy DOID-3191 mouse GSE3384 sample 971

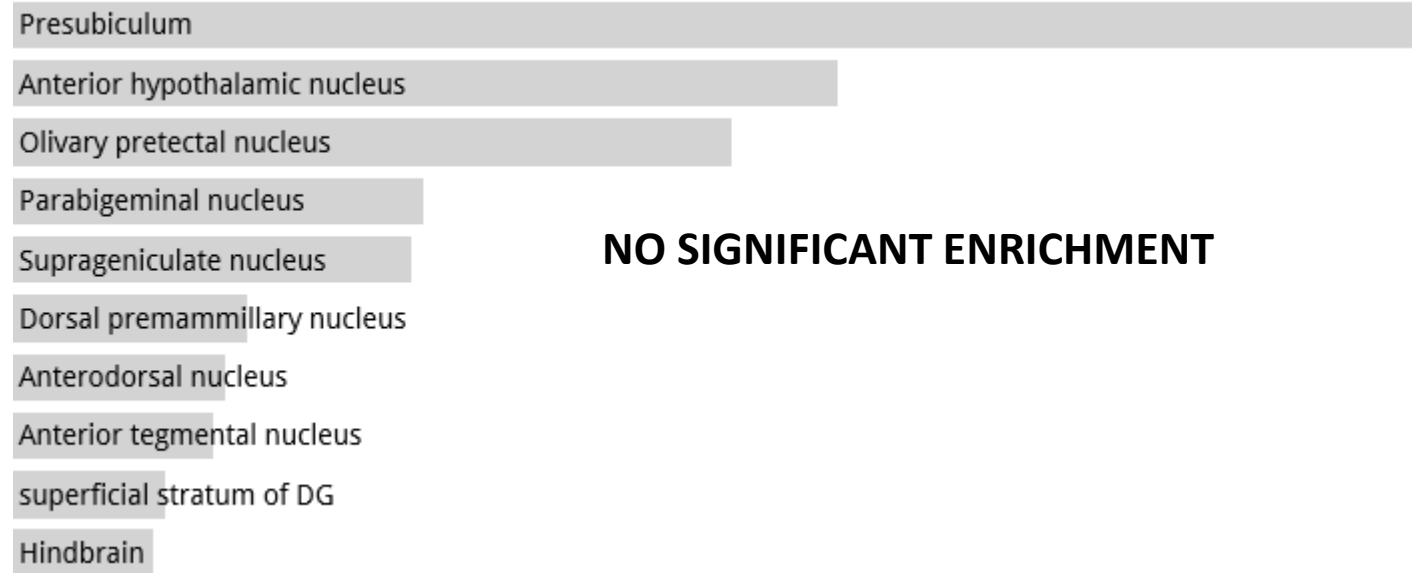
juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 772

Schistosomiasis C0036323 mouse GSE19525 sample 439

**NO SIGNIFICANT ENRICHMENT**

## Random Set 16

## Allen Brain Atlas up random 4000 genes



## Random Set 17

ARCHS4 Tissues random 4000 genes

SMALL INTESTINE (BULK TISSUE)

VASCULAR SMOOTH MUSCLE

SKIN (BULK TISSUE)

THYROID (BULK TISSUE)

STROMAL CELL

RESPIRATORY SMOOTH MUSCLE

RENAL CORTEX

SKELETAL MUSCLE (BULK TISSUE)

REGULATORY T CELLS

PANCREATIC ISLET

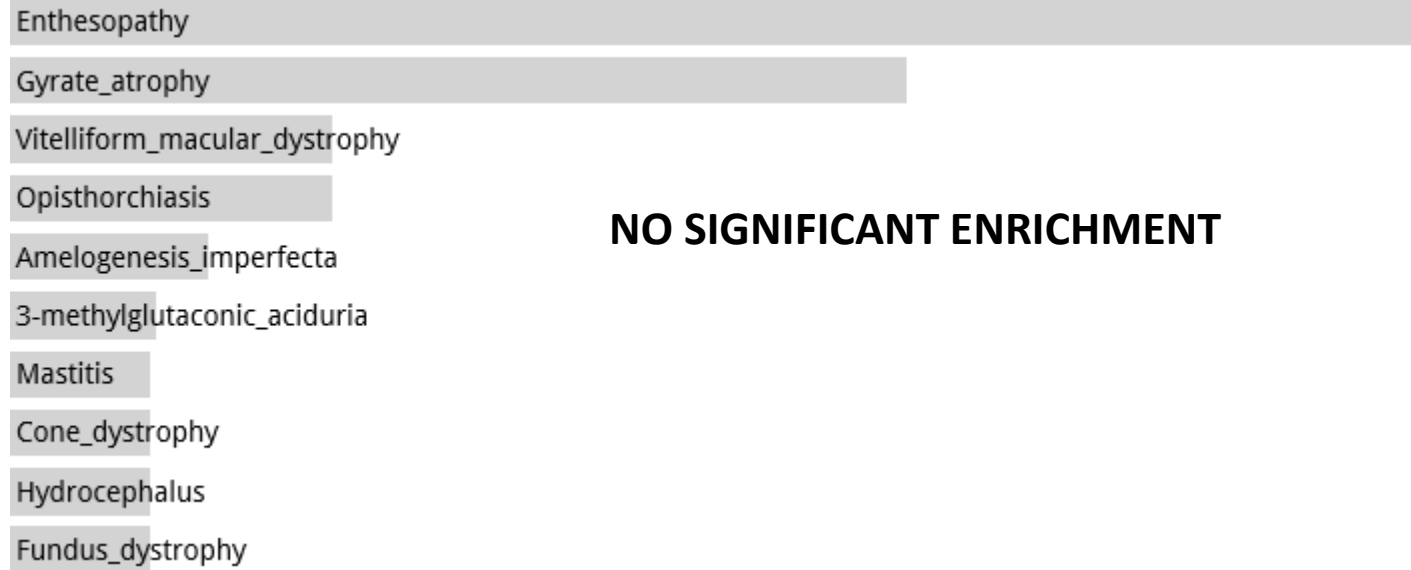
**NO SIGNIFICANT ENRICHMENT**

Jensen TISSUES random 4000 genes

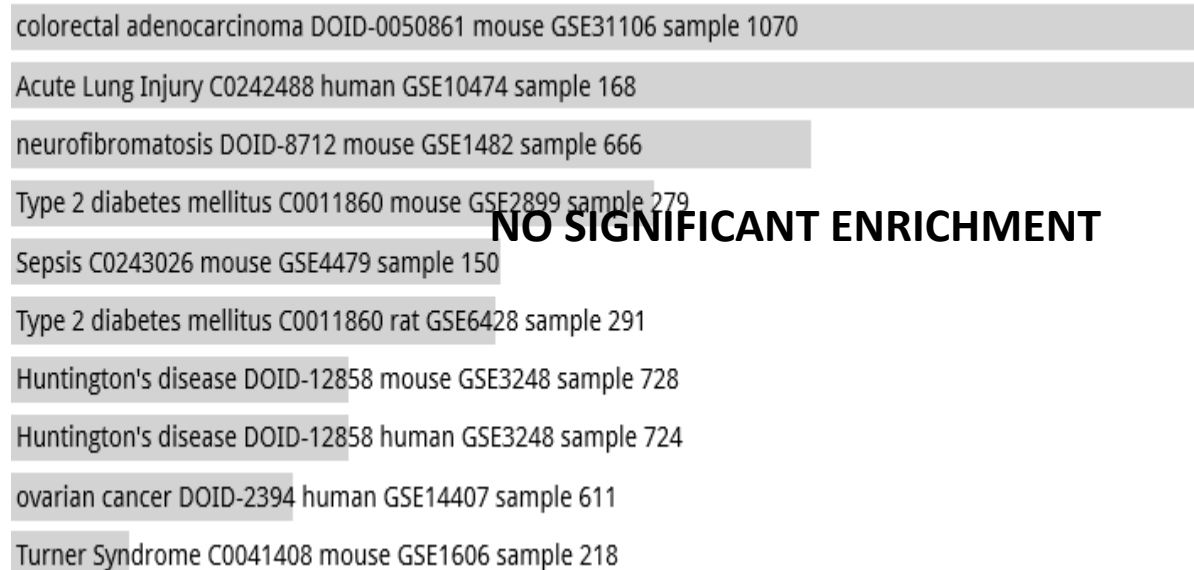
**NO DATA AVAILABLE**

## Random Set 17

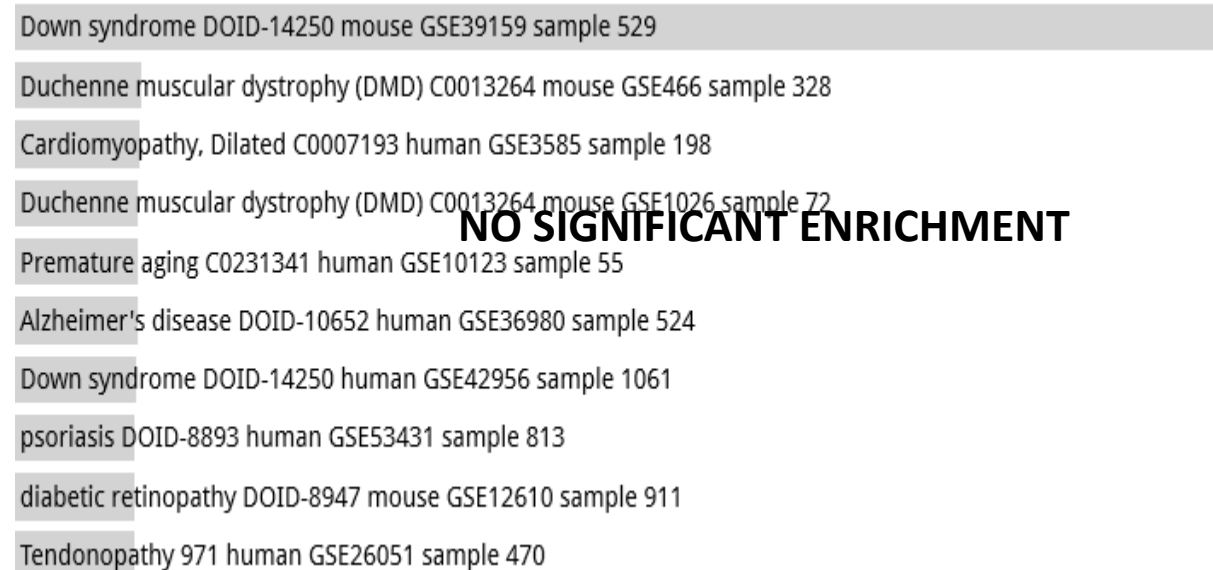
## Jensen DISEASES random 4000 genes



## Disease Perturbations from GEO up random 4000 genes



## Disease Perturbations from GEO down random 4000 genes



## Random Set 17

## Allen Brain Atlas up random 4000 genes

Presubiculum

Endopiriform nucleus, dorsal part

dorsal endopiriform nucleus

Copula pyramidis

bed nucleus of the external capsule

mantle zone of r7Co

Agranular insular area, posterior part, layer 6a

Entorhinal area, medial part, dorsal zone, layer 5

corticoid layer of TuStr

Infralimbic area, layer 6b

**NO SIGNIFICANT ENRICHMENT**

**Random Set 18**

ARCHS4 Tissues random 4000 genes

**NO DATA AVAILABLE**

Jensen TISSUES random 4000 genes

**NO DATA AVAILABLE**

**Random Set 18**

Jensen DISEASES random 4000 genes

**NO DATA AVAILABLE**

Disease Perturbations from GEO up random 4000 genes

Disease Perturbations from GEO down random 4000 genes

**NO DATA AVAILABLE**

**NO DATA AVAILABLE**

**Random Set 18**

Allen Brain Atlas up random 4000 genes

**NO DATA AVAILABLE**



**Random Set 19**

ARCHS4 Tissues random 4000 genes

**NO DATA AVIALABLE**

Jensen TISSUES random 4000 genes

**NO DATA AVAILABLE**

**Random Set 19**

Jensen DISEASES random 4000 genes

**NO DATA AVAILABLE**

Disease Perturbations from GEO up random 4000 genes

Disease Perturbations from GEO down random 4000 genes

**NO DATA AVAILABLE**

**NO DATA AVAILABLE**

**Random Set 19**

Allen Brain Atlas up random 4000 genes

**NO DATA AVAILABLE**

## Random Set 20

### ARCHS4 Tissues random 4000 genes

LYMPHOCYTE

BLASTOCYST

BLOOD DENDRITIC CELLS

BONE MARROW (BULK TISSUE)

SPINAL CORD (BULK)

GASTRIC TISSUE (BULK)

SPINAL CORD

MOTOR NEURON

PREFRONTAL CORTEX

NEURONAL EPITHELIUM

**NO SIGNIFICANT ENRICHMENT**

### Jensen TISSUES random 4000 genes

**NO DATA AVAILABLE**

## Random Set 20

## Jensen DISEASES random 4000 genes

Acrodermatitis\_enteropathica

Vestibular\_neuronitis

Acute\_poststreptococcal\_glomerulonephritis

Diastrophic\_dysplasia

Eclampsia

Biotin\_deficiency

Schimke\_immuno-osseous\_dysplasia

Bowen-Conradi\_syndrome

Chikungunya

Western\_equine\_encephalitis

**NO SIGNIFICANT ENRICHMENT**

## Disease Perturbations from GEO up random 4000 genes

Nicotine addiction C0028043 human GSE11208 sample 325

Alzheimer's disease DOID-10652 human GSE36980 sample 524

multiple sclerosis DOID-2377 human GSE16461 sample 584

Alzheimer's disease DOID-10652 human GSE36980 sample 520

Intestinal polyposis UMLS CUI-C1257915 mouse GSE6078 sample 952

Nephroblastoma C0027708 human GSE2712 sample 418

Alzheimer's disease DOID-10652 human GSE36980 sample 522

testis seminoma DOID-5842 human GSE8607 sample 624

polycystic ovary syndrome DOID-11612 human GSE48301 sample 558

esophagus squamous cell carcinoma DOID-3748 human GSE63941 sample 658

**NO SIGNIFICANT ENRICHMENT**

## Disease Perturbations from GEO down random 4000 genes

monoclonal gammopathy of uncertain significance DOID-7442 human GSE47552 sample 563

Purpura, Idiopathic Thrombocytopenic C0043117 human GSE574 sample 358

smoldering myeloma DOID-9551 human GSE47552 sample 562

Eczema C0013595 human GSE6012 sample 268

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 771

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 772

Sarcoidosis C0036202 human GSE2657 sample 12

systemic lupus erythematosus DOID-9074 human GSE10325 sample 691

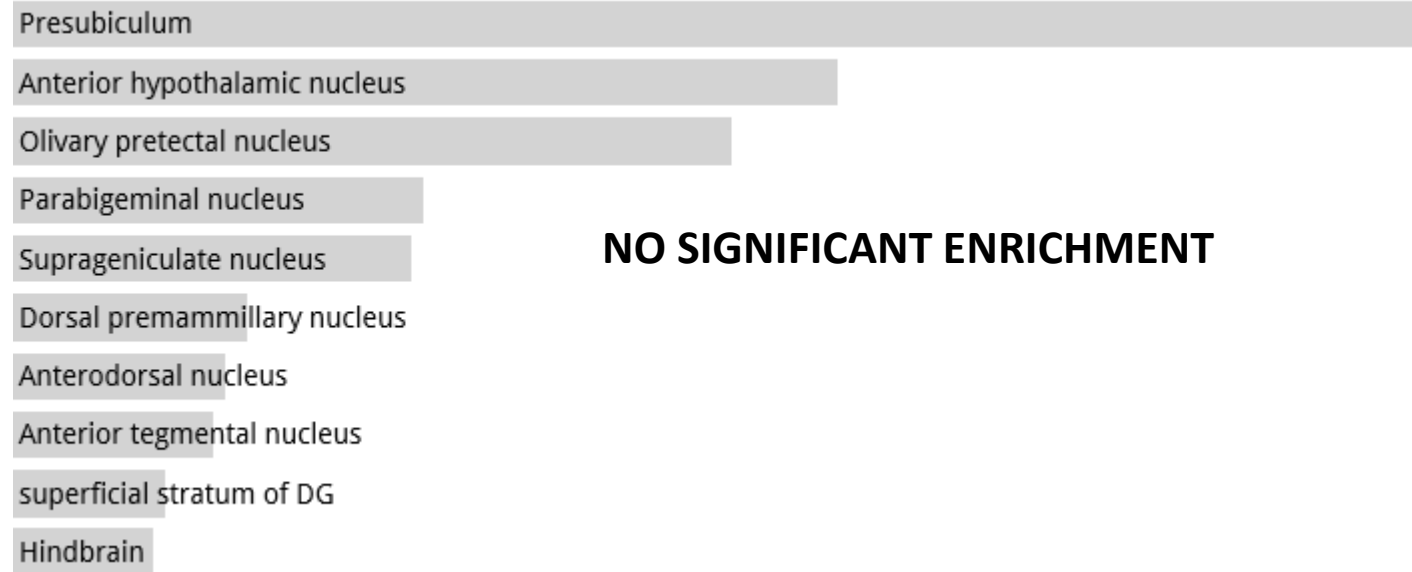
Pulmonary Hypertension C0020542 human GSE703 sample 241

multiple myeloma DOID-9538 human GSE47552 sample 561

**NO SIGNIFICANT ENRICHMENT**

## Random Set 20

## Allen Brain Atlas up random 4000 genes



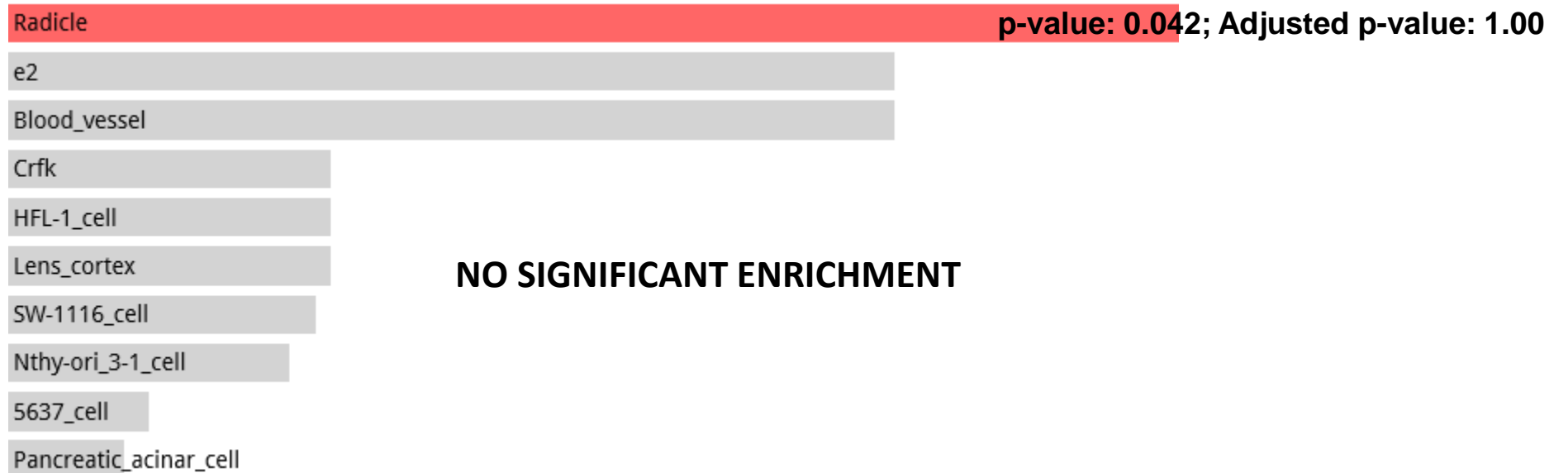
**NO SIGNIFICANT ENRICHMENT**

## Random Set 21

### ARCHS4 Tissues random 2847 genes



### Jensen TISSUES random 2847 genes

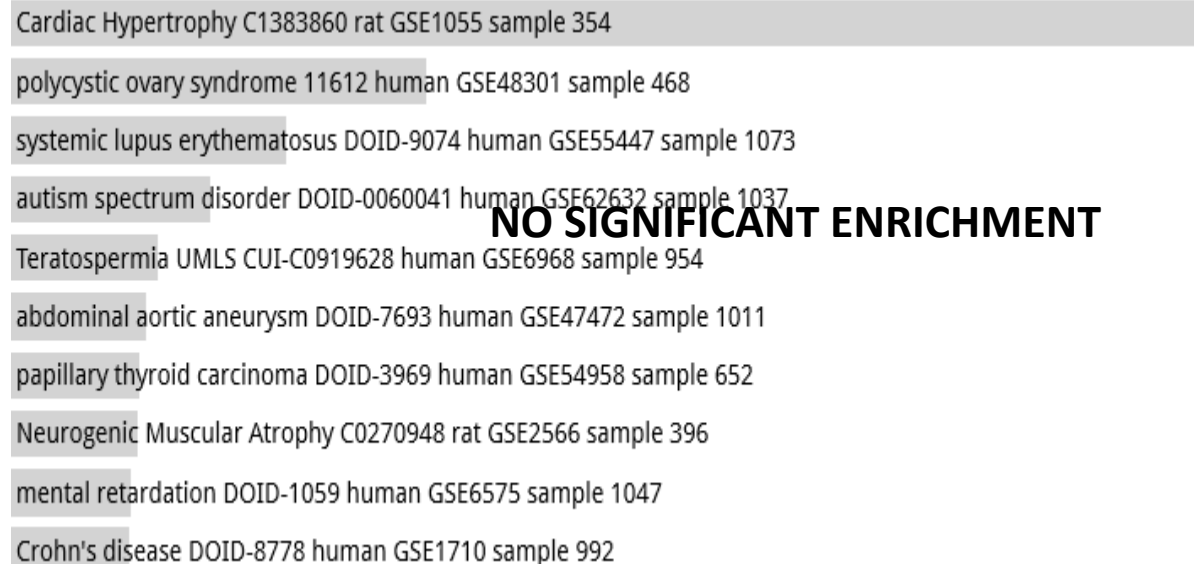


## Random Set 21

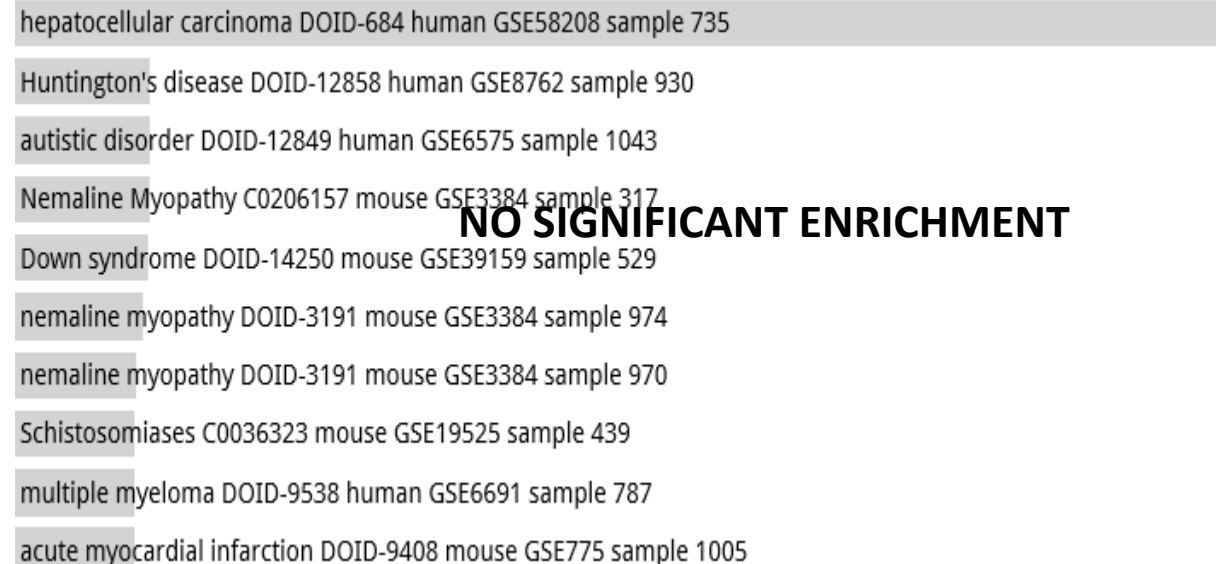
## Jensen DISEASES random 2847 genes



## Disease Perturbations from GEO up random 2847 genes



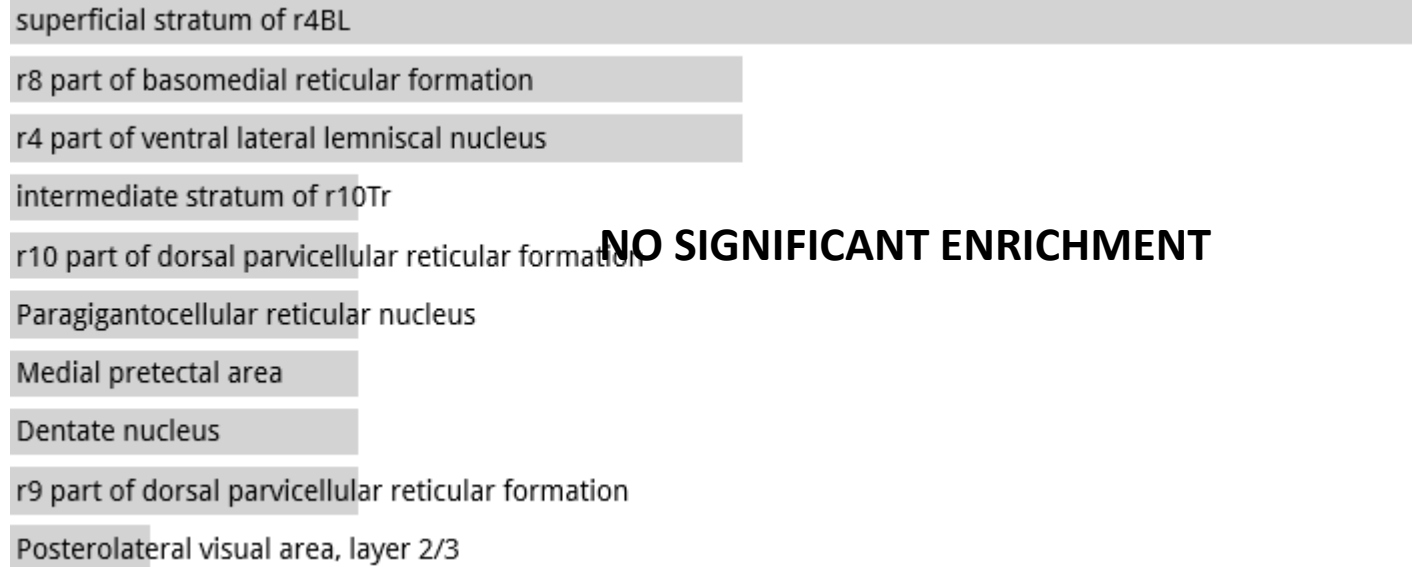
## Disease Perturbations from GEO down random 2847 genes



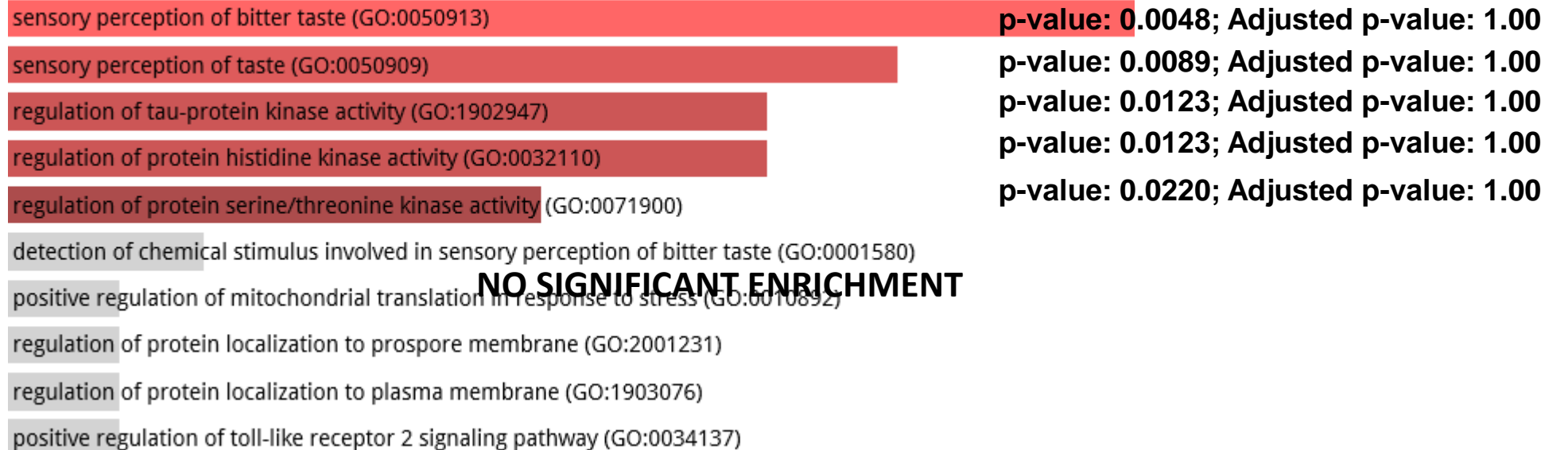


## Random Set 21

### Allen Brain Atlas up random 2847 genes

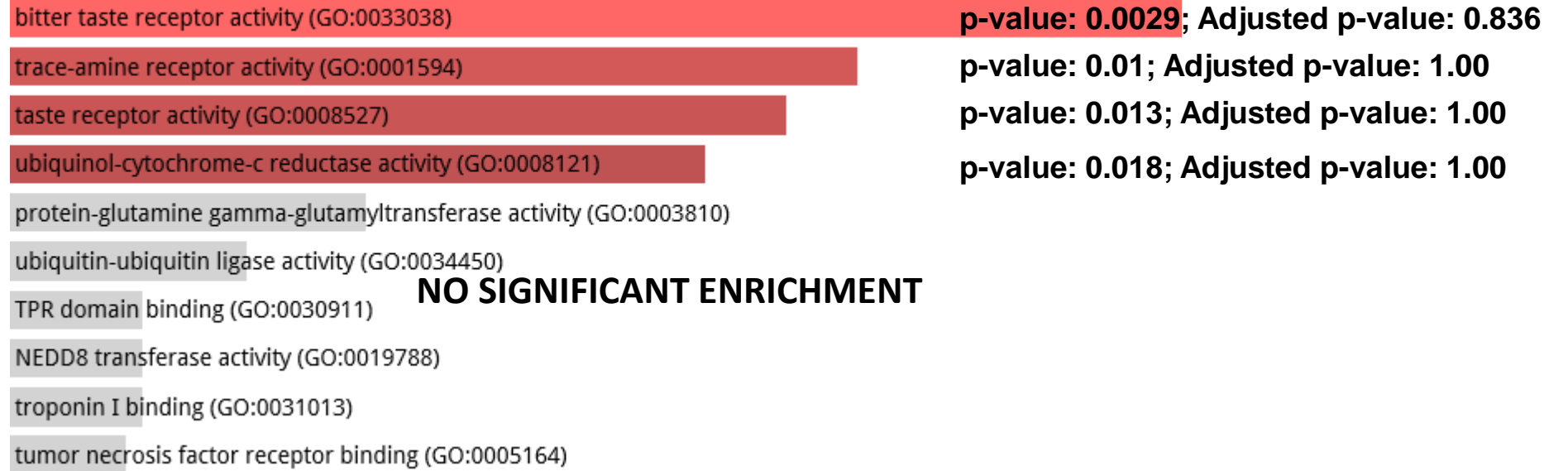


### GO Biological Process 2017b 2847 genes

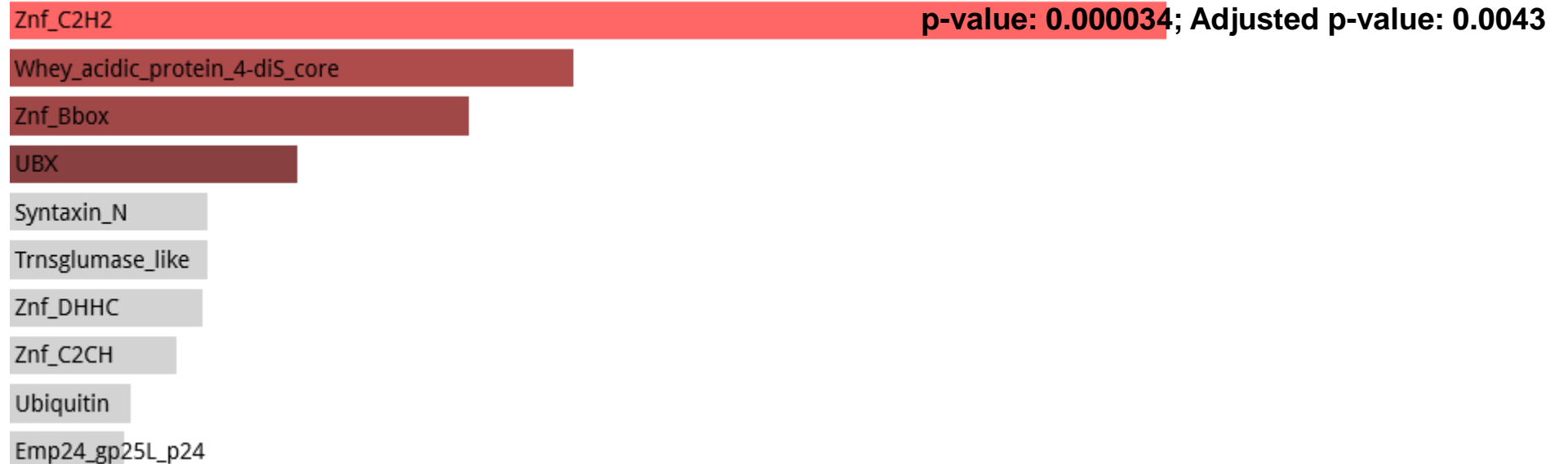


## Random Set 21

## GO Molecular Function 2017 2847 genes



## Pfam InterPro Domains 2847 genes



## Random Set 21

## GWAS Catalog 2019 2847 genes

Alcoholic chronic pancreatitis	p-value: 0.00049; Adjusted p-value: 0.31
Plasminogen activator inhibitor type 1 levels (PAI-1)	
Total bilirubin levels in HIV-1 infection	
Cerebellum growth	
Response to simvastatin treatment (PCSK9 protein level change)	
<b>NO SIGNIFICANT ENRICHMENT</b>	
Pancreatic ductal adenocarcinoma	
Hepcidin/transferrin saturation ratio	
Renal function-related traits (sCR)	
Proteinuria and chronic kidney disease	
Coffee consumption (cups per day)	

## GO Biological Process 2018 2847 genes

bitter taste receptor activity (GO:0033038)	p-value: 0.00047; Adjusted p-value: 0.1939
taste receptor activity (GO:0008527)	p-value: 0.0034; Adjusted p-value: 0.7151
trace-amine receptor activity (GO:0001594)	p-value: 0.018; Adjusted p-value: 1.00
oxidoreductase activity, acting on diphenols and related substances as donors, cytochrome as acceptor (GO:0004711)	p-value: 0.028; Adjusted p-value: 1.00
ubiquinol-cytochrome-c reductase activity (GO:0008121)	p-value: 0.028; Adjusted p-value: 1.00
<b>NO SIGNIFICANT ENRICHMENT</b>	
uridylyltransferase activity (GO:0070569)	
protein-glutamine gamma-glutamyltransferase activity (GO:0003810)	
ubiquitin-ubiquitin ligase activity (GO:0034450)	
MHC protein binding (GO:0042287)	
small protein activating enzyme activity (GO:0008641)	

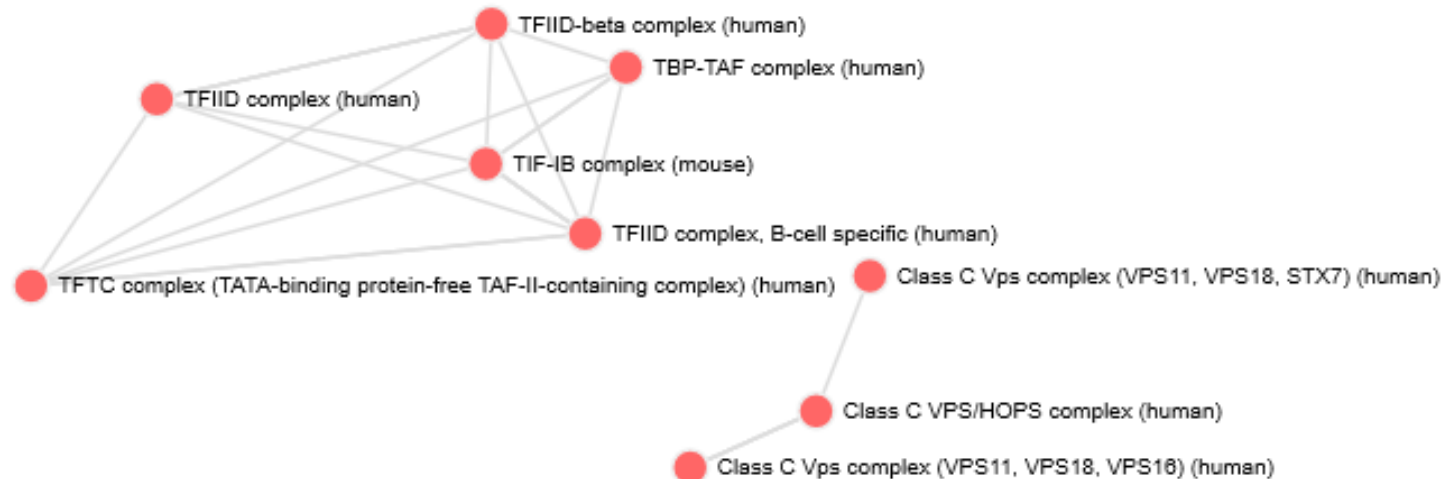
## Random Set 21

## CORUM 2847 genes

Class C Vps complex (VPS11, VPS18, VPS16) (human)	p-value: 0.0029; Adjusted p-value: 0.2781
Class C Vps complex (VPS11, VPS18, STX7) (human)	p-value: 0.0029; Adjusted p-value: 0.2781
Class C VPS/HOPS complex (human)	p-value: 0.0048; Adjusted p-value: 0.3109
TBP-TAF complex (human)	p-value: 0.01; Adjusted p-value: 0.3975
TIF-IB complex (mouse)	p-value: 0.01; Adjusted p-value: 0.3975
TFIID complex (human)	
TFIID complex, B-cell specific (human)	
TRAF2-TRADD complex (human)	
TSC1-TSC2 complex (human)	
TFTC complex (TATA-binding protein-free TAF-II-containing complex) (human)	

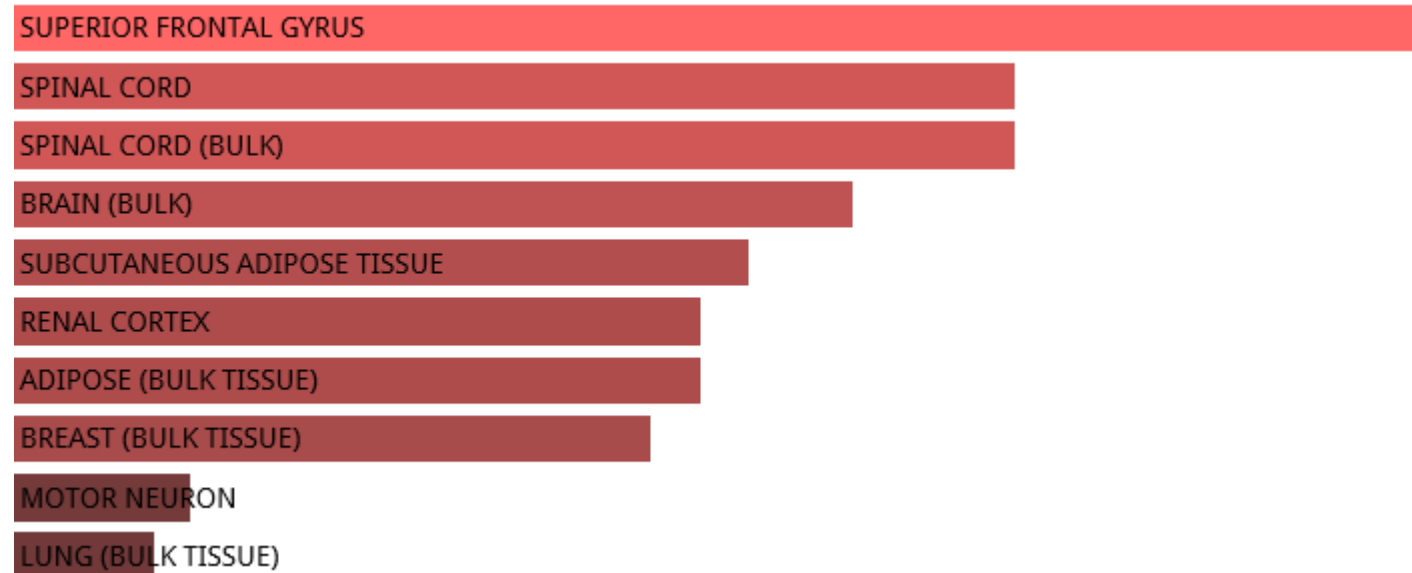
**NO SIGNIFICANT ENRICHMENT**

## CORUM database of protein complexes identified by mass spectrometry

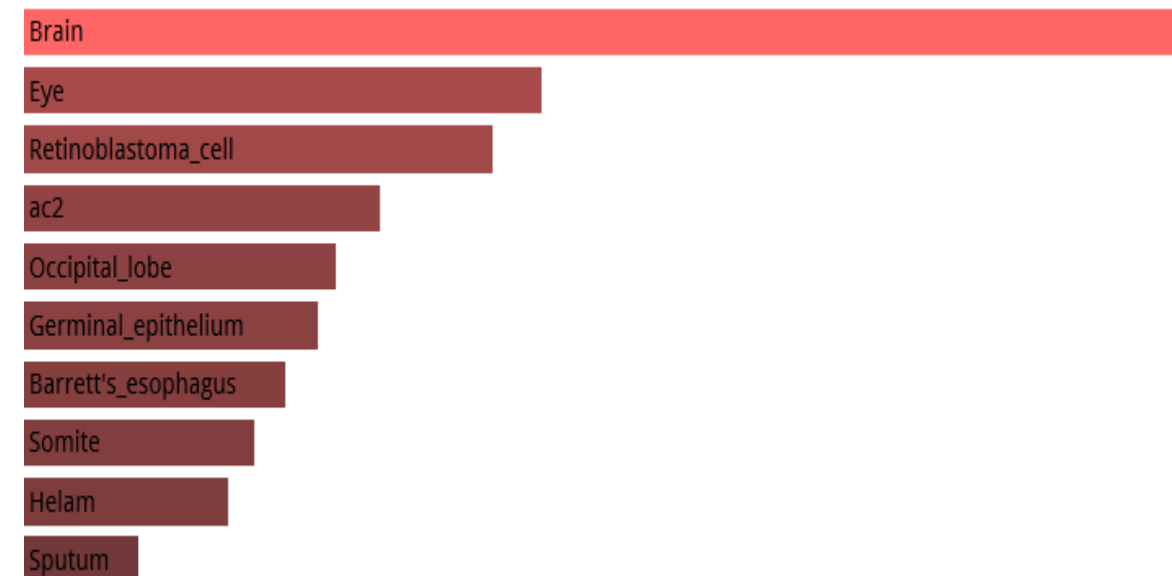


# 1,279 human-specific STR contractions

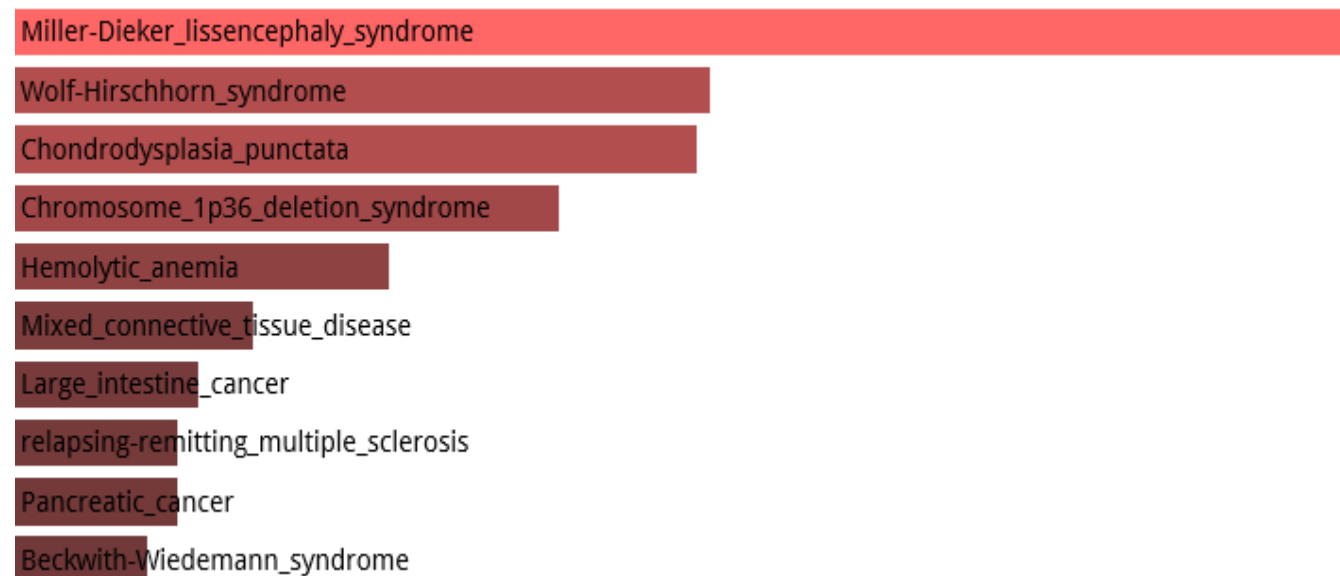
## ARCHS4 Tissues 973 genes



## Jensen TISSUES 973 genes

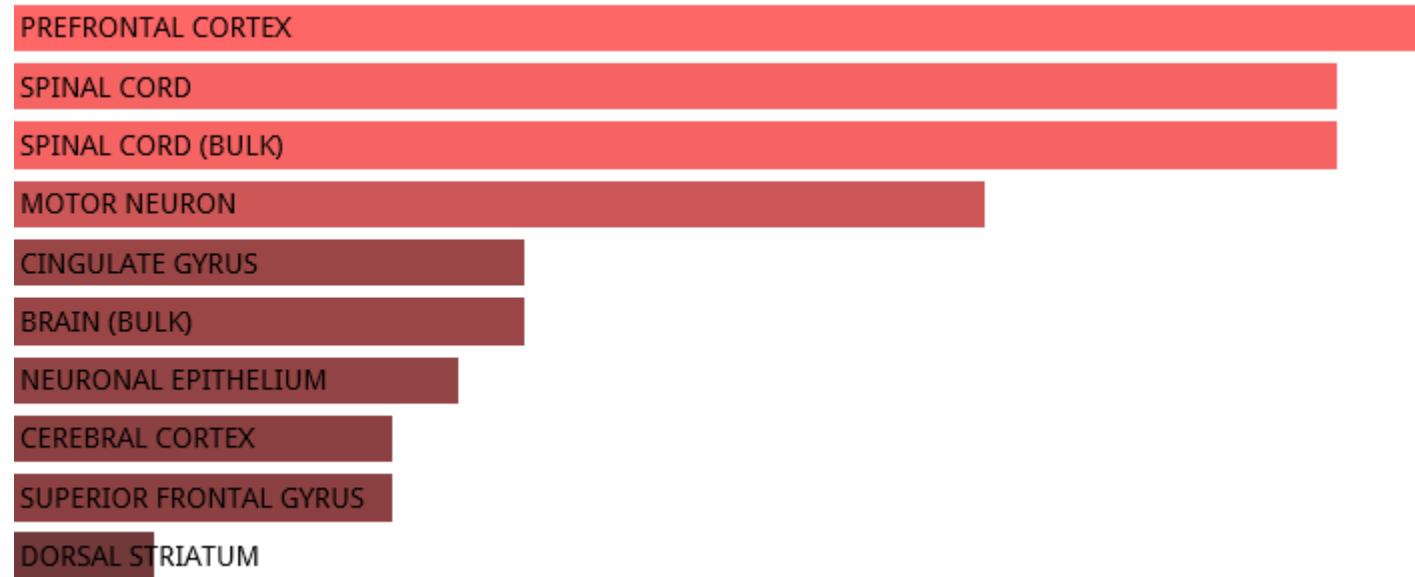


## Jensen DISEASES 973 genes



**2,118 fixed human-specific regulatory regions  
(non-hESC DHS)**

ARCHS4 Tissues 552 genes



**410 H3K4me3 sites with human-specific  
enrichment in prefrontal cortex neurons**

ARCHS4 Tissues 578 genes

