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.

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

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)

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

11,878 fixed human-specific insertions

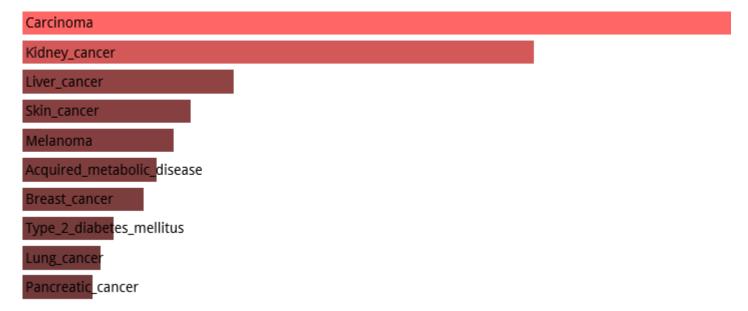
ARCHS4 Tissues 7979 genes

PREFRONTAL CORTEX	
NEURONAL EPITHELIUM	
SPINAL CORD	
SPINAL CORD (BULK)	
CEREBELLUM	
CINGULATE GYRUS	
CEREBRAL CORTEX	
RENAL CORTEX	
MOTOR NEURON	
BRAIN (BULK)	
	Jensen TISSUES 7979 genes
Hypothalamus	
Brain	
Cerebral_cortex	
Heart	
Lung	
Adipose_tissue	
Adrenal_gland	
Gall_bladder	
Colon	
Kidney	

11,878 fixed human-specific insertions

schizophrenia DOID-5419 human GSE25673 sample 891

Jensen DISEASES 7979 genes



Disease Perturbations from GEO up

Disease Perturbations from GEO down

schizophrenia DOID-5419 human GSF25673 sample 892

Schizophilenia Dolb-5415 human dSc25075 Sample 051	Schizophrenia Dolb-5419 human daezao/3 sample 692
Spinal Muscular Atrophy C0026847 mouse GSE10599 sample 235	Crohn's disease DOID-8778 human GSE6731 sample 757
schizophrenia DOID-5419 human GSE25673 sample 892	Bipolar Disorder C0005586 human GSE5389 sample 302
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850	adrenoleukodystrophy DOID-10588 human GSE34309 sample 864
morbid obesity DOID-11981 human GSE48964 sample 583	ulcerative colitis DOID-8577 human GSE6731 sample 759
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 851	Primary open angle glaucoma C0339573 human GSE2705 sample 257
Primary open angle glaucoma C0339573 human GSE2705 sample 257	Ulcerative Colitis C0009324 human GSE6731 sample 249
cardiomyopathy DOID-0050700 human GSE9128 sample 780	idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850
adrenoleuko dystrophy DOID-10588 human GSE34309 sample 864	Breast Cancer C0006142 human GSE1378 sample 52
Cardiomyopathy, Dilated C0007193 human GSE3585 sample 198	Crohn's disease DOID-8778 human GSE6731 sample 758

7,599 duplicated regions in the human genome

ARCHS4 Tissues 6618 genes

NEURONAL EPITHELIUM PREFRONTAL CORTEX FETAL BRAIN CORTEX SPINAL CORD SPINAL CORD (BULK) CEREBELLUM CINGULATE GYRUS MOTOR NEURON **HUMAN EMBRYO** CEREBRAL CORTEX Jensen TISSUES 6618 genes Hypothalamus Frontal_lobe Heart Retina Trachea Adipose_tissue Occipital_lobe B-lymphocyte Natural_killer_cell

7,599 duplicated regions in the human genome

Jensen DISEASES 6618 genes

Kidney_cancer
Liver_cancer
Breast_cancer
Skin_cancer
Melanoma
Lung_cancer
Endometrial_cancer
Alzheimer's_disease
Pancreatic_cancer

Disease Perturbations from GEO up

Disease Perturbations from GEO down

myotonic dystrophy type 1 DOID-11722 human GSE7177 sample 1053	Bipolar Disorder C0005586 human GSE5389 sample 302
schizophrenia DOID-5419 human GSE25673 sample 891	Crohn's disease DOID-8778 human GSE6731 sample 757
mytonic dystrophy Type 1 DOID-11722 human GSE7179 sample 1049	ulcerative colitis DOID-8577 human GSE6731 sample 759
breast cancer DOID-1612 human GSE14943 sample 504	autism spectrum disorder DOID-0060041 human GSE62632 sample 1037
anaplastic thyroid carcinoma UMLS CUI-C0238461 human GSE65144 sample 654	Nephroblastoma C0027708 human GSE2712 sample 418
myotonic dystrophy type 1 DOID-11722 human GSE7179 sample 1050	Primary open angle glaucoma C0339573 human GSE2705 sample 257
Asthma, allergic CO155877 human GSE3004 sample 360	Crohn's disease DOID-8778 human GSE6731 sample 758
Schizophrenia C0036341 human GSE12649 sample 261	schizophrenia DOID-5419 human GSE25673 sample 892
endometrial cancer DOID-1380 human GSE17025 sample 840	Ulcerative Colitis C0009324 human GSE6731 sample 249
cardiomyopathy DOID-0050700 human GSE9128 sample 780	myotonic dystrophy type 1 DOID-11722 human GSE7178 sample 1051

5,883 fixed human-specific deletions

ARCHS4 Tissues 5489 genes

PREFRONTAL CORTEX	
NEURONAL EPITHELIUM	
SPINAL CORD	
SPINAL CORD (BULK)	
CEREBELLUM	
CINGULATE GYRUS	
CEREBRAL CORTEX	
RENAL CORTEX	
MOTOR NEURON	
BRAIN (BULK)	
	Jensen TISSUES 5489 genes
	-
Hypothalamus	
Hypothalamus Cerebral_cortex	
Cerebral_cortex	
Cerebral_cortex Brain	
Cerebral_cortex Brain Gall_bladder	
Cerebral_cortex Brain Gall_bladder Urinary_bladder	
Cerebral_cortex Brain Gall_bladder Urinary_bladder Testis	
Cerebral_cortex Brain Gall_bladder Urinary_bladder Testis Stomach	

5,883 fixed human-specific deletions

schizonhrenia DOID-5/119 human GSE25673 sample 891

Jensen DISEASES 5489 genes



Disease Perturbations from GEO up

Disease Perturbations from GEO down

schizonhrenia DOID-5/19 human GSF25673 sample 892

schizophrenia DOID-5419 human G5E25675 sample 691	Schizophrenia DOID-5419 human GSE25673 Sample 692
Spinal Muscular Atrophy C0026847 mouse GSE10599 sample 235	idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850
schizophrenia DOID-5419 human GSE25673 sample 892	Alzheimer's disease DOID-10652 human GSE36980 sample 520
morbid obesity DOID-11981 human GSE48964 sample 583	schizophrenia DOID-5 <mark>4</mark> 19 human GSE25673 sample 891
adrenoleukodystrophy DOID-10588 human GSE34309 sample 865	ulcerative colitis DOID-8577 human GSE6731 sample 759
autism spectrum disorder DOID-0060041 human GSE62632 sample 1036	Crohn's disease DOID-8778 human GSE6731 sample 757
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 851	Huntington's disease DOID-12858 mouse GSE3583 sample 929
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850	Breast Cancer C0006142 human GSE1378 sample 52
Down syndrome DOID-14250 human GSE42956 sample 1060	Lewy body dementia DOID-12217 human GSE49036 sample 1068
Pauciarticular juvenile arthritis C0157917 human GSE1402 sample 430	adrenoleukodystrophy DOID-10588 human GSE34309 sample 864

4,875 human-specific STR expansions

ARCHS4 Tissues 4844 genes

SPINAL CORD	
SPINAL CORD (BULK)	
PREFRONTAL CORTEX	
CINGULATE GYRUS	
BRAIN (BULK)	
CEREBRAL CORTEX	
MOTOR NEURON	
CEREBELLUM	
RENAL CORTEX	
SUPERIOR FRONTAL GYRUS	
	Jensen TISSUES 4844 genes
	•
Hypothalamus	
Hypothalamus Brain	
Brain	
Brain Cerebral_cortex	
Brain Cerebral_cortex Adrenal_gland	
Brain Cerebral_cortex Adrenal_gland Lung	
Brain Cerebral_cortex Adrenal_gland Lung Occipital_lobe	
Brain Cerebral_cortex Adrenal_gland Lung Occipital_lobe Urinary_bladder	

4,875 human-specific STR expansions

Jensen DISEASES 4844 genes

Kidney_cancer
Liver_cancer

Skin_cancer

Pancreatic_cancer

Melanoma

Schizophrenia

Breast_cancer

Acquired_metabolic_disease

Mental_depression

Disease Perturbations from GEO up

Disease Perturbations from GEO down

schizophrenia DOID-5419 human GSE25673 sample 891	schizophrenia DOID-5419 human GSE25673 sample 892
Spinal Muscular Atrophy C0026847 mouse GSE10599 sample 235	idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 851	Huntington's disease DOID-12858 mouse GSE3583 sample 929
morbid obesity DOID-11981 human GSE48964 sample 583	Primary open angle glaucoma C0339573 human GSE2705 sample 257
Diamond-Blackfan anaemia DOID-1339 human GSE14335 sample 472	Bipolar Disorder C0005586 human GSE5389 sample 302
adrenoleukodystrophy DOID-10588 human GSE34308 sample 709	Androgen insensitivity syndrome C0039585 human GSE3871 sample 415
Down Syndrome C0013080 human GSE5390 sample 277	idiopathic pulmonary fibrosis DOID-0050156 human GSE24206 sample 871
schizophrenia DOID-5419 human GSE25673 sample 892	Alzheimer's disease DOID-10652 human GSE36980 sample 520
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850	Gastrointestinal stromal tumor C0238198 human GSE15966 sample 76
intracrania <mark>l aneurysm DOID-10941 human GSE26969 sample 744</mark>	Crohn's disease DOID-8778 human GSE6731 sample 757

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

PREFRONT

CINGULAT

ARCHS4 Tissues 4051 genes

PREFRONTAL CORTEX

CINGULATE GYRUS

NEURONAL EPITHELIUM

CEREBELLUM

CEREBRAL CORTEX

SPINAL CORD

SPINAL CORD (BULK)

DORSAL STRIATUM

MOTOR NEURON

FETAL BRAIN CORTEX

Jensen TISSUES 4051 genes

Brain

Cerebral_cortex

Frontal_lobe

Occipital_lobe

Heart

Ovary

Cerebellum

Parietal_lobe

Gall_bladder

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

Spinal Muscular Atrophy C0026847 mouse GSE10599 sample 235

Jensen DISEASES 4051 genes

Kidney_cancer
Liver_cancer

Melanoma
Skin_cancer

Breast_cancer

Acquired_metabolic_disease
Lung_cancer

Endometrial_cancer

Attention_deficit_hyperactivity_disorder

Disease Perturbations from GEO up

Disease Perturbations from GEO down

Bipolar Disorder C0005586 human GSE5389 sample 302

adrenoleukodystrophy DOID-10588 human GSE343 <mark>09 sample 864</mark>	autism spectrum disorder DOID-0060041 human GSE62632 sample 1037
morbid obesity DOID-11981 human GSE48964 sample 583	schizophrenia DOID-5419 human GSE25673 sample 892
schizophrenia DOID-5419 human GSE25673 sample 892	idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850
schizophrenia DOID-5419 human GSE25673 sample 891	Crohn's disease DOID-8778 human GSE6731 sample 757
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 851	schizophrenia DOID-5419 human GSE25673 sample 891
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850	ulcerative colitis DOID-8577 human GSE6731 sample 759
adrenoleukodystrophy DOID-10588 human GSE34308 sample 709	Crohn's disease DOID-8778 human GSE6731 sample 758
smoldering myeloma DOID-9551 human GSE47552 sample 562	Alzheimer's disease DOID-10652 human GSE36980 sample 520
melanoma DOID-1909 human GSE6887 sample 948	Nephroblastoma C0027708 human GSE2712 sample 418

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

Allen Brain Atlas up 4051 genes

Dentate gyrus, molecular layer

molecular layer of the DG

Dentate gyrus

superficial stratum of DG

granule cell layer of the DG

Hippocampal region

Dentate gyrus, granule cell layer

bed nucleus of the external capsule

Field CA1, stratum pyramidale

medial pallium (hippocampal allocortex)

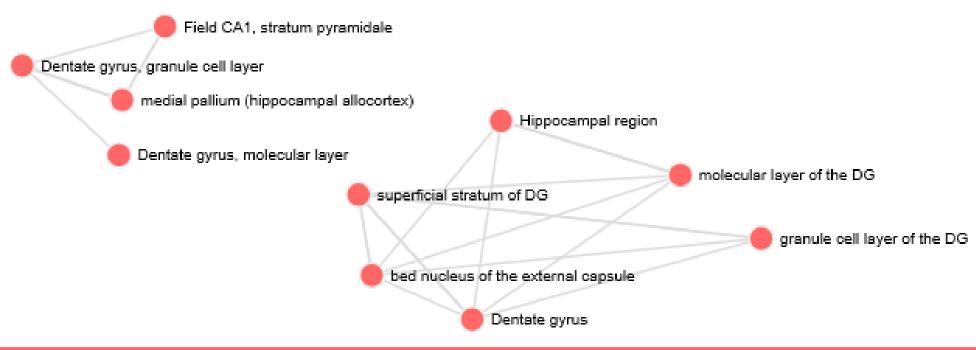
sorted by p-value ranking

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



Dentate gyrus

Dentate gyrus, molecular layer

granule cell layer of the DG

molecular layer of the DG

Hippocampal region

Dentate gyrus, granule cell layer

bed nucleus of the external capsule

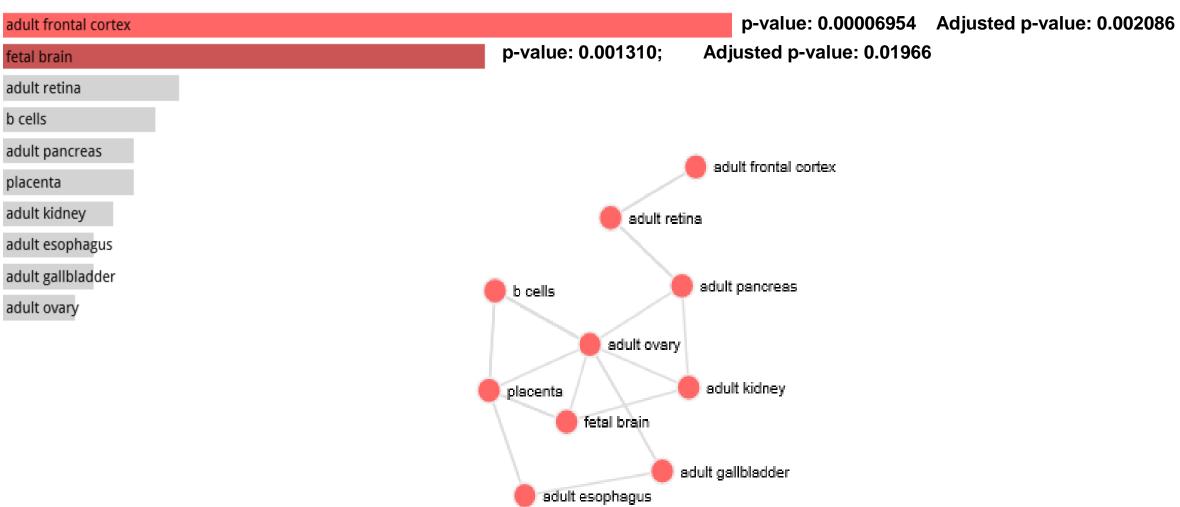
Field CA1, stratum pyramidale

medial pallium (hippocampal allocortex)

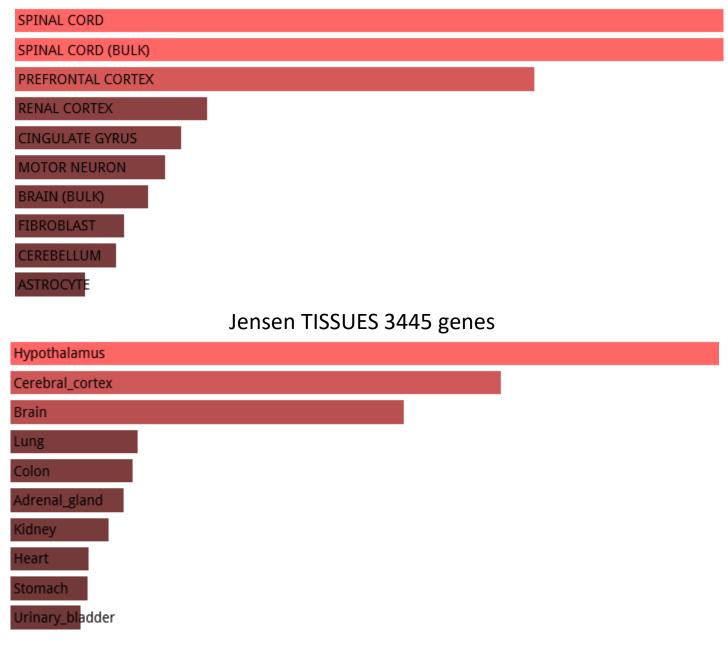
superficial stratum of DG

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

Tissue Protein Expression from Human Proteome Map: 4051 genes



ARCHS4 Tissues 3445 genes



3,538 human-specific ace-DHS

Jensen DISEASES 3445 genes

Kidney_cancer

Type_2_diabetes_mellitus

Liver_cancer

Breast_cancer

Acquired_metabolic_disease

Melanoma

Lymphoid_leukemia

Skin_cancer

Immune_system_cancer

Disease Perturbations from GEO up

funtington's disease DOID-12858 mouse GSE3583 sample 929

Huntington's disease DOID-12858 mouse GSE3583 sample 929

pe 2 diabetes mellitus C0011860 human GSE12643 sample 274

drenoleukodystrophy DOID-10588 human GSE34309 sample 864

sophagus squamous cell carcinoma DOID-3748 human GSE63941 sample 658

diopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850

schizophrenia DOID-5419 human GSE25673 sample 892

ipolar Disorder C0005586 human GSE5389 sample 302

rohn's disease DOID-8778 human GSE6731 sample 757

reast cancer DOID-1612 human GSE9574 sample 448

ystonia C0393593 human GSE3064 sample 329

Disease Perturbations from GEO down

Diamond-Blackfan anaemia DOID-1339 human GSE14335 sample 472
morbid obesity DOID-11981 human GSE48964 sample 583
pancreatic cancer DOID-1793 human GSE23952 sample 798
GERD - Gastro-esophageal reflux disease C0017168 human GSE2144 sample 27
diabetes mellitus type 2 DOID-9352 human GSE12643 sample 766
Spinal Muscular Atrophy C0026847 mouse GSE10599 sample 235
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 851
adrenoleukodystrophy DOID-10588 human GSE34309 sample 865
Down Syndrome C0013080 human GSE5390 sample 277

Allen Brain Atlas up 3445 genes

Field CA2, stratum oriens

medial pallium (hippocampal allocortex)

Hippocampal formation

Field CA2, stratum radiatum

granule cell layer of the DG

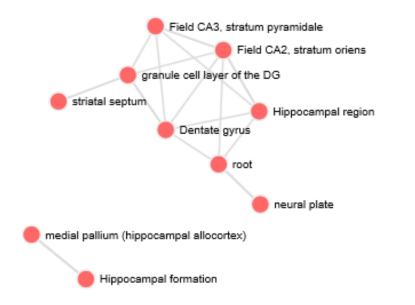
Hippocampal region

Field CA2

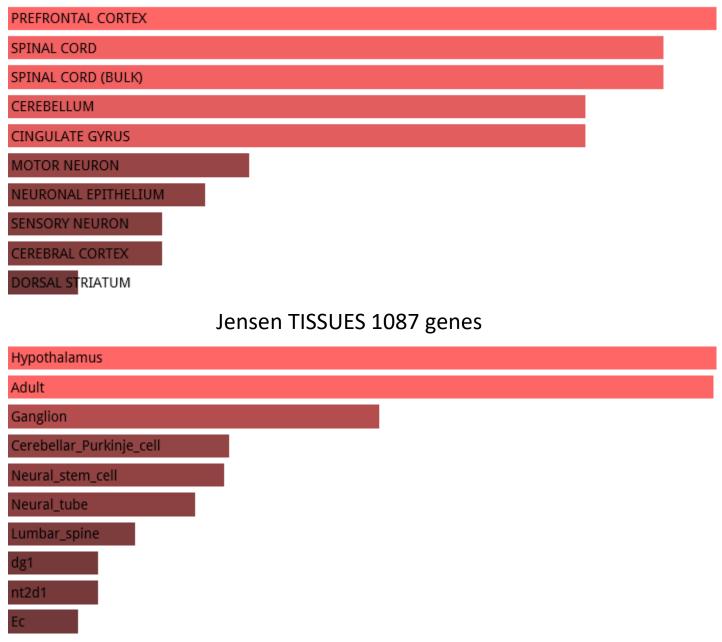
neural plate

striatal septum

Field CA3, stratum lucidum

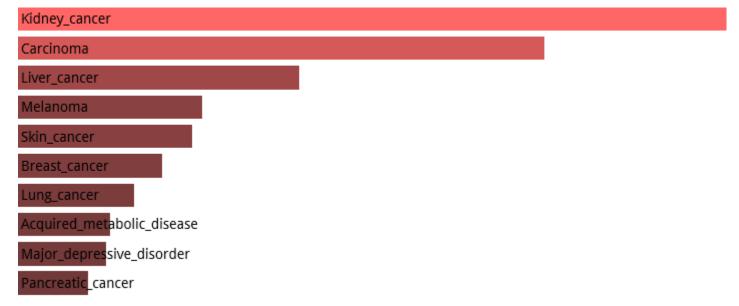


ARCHS4 Tissues 1087 genes



3,803 human-specific TFBS in hESC

Jensen DISEASES 1087 genes



Disease Perturbations from GEO up

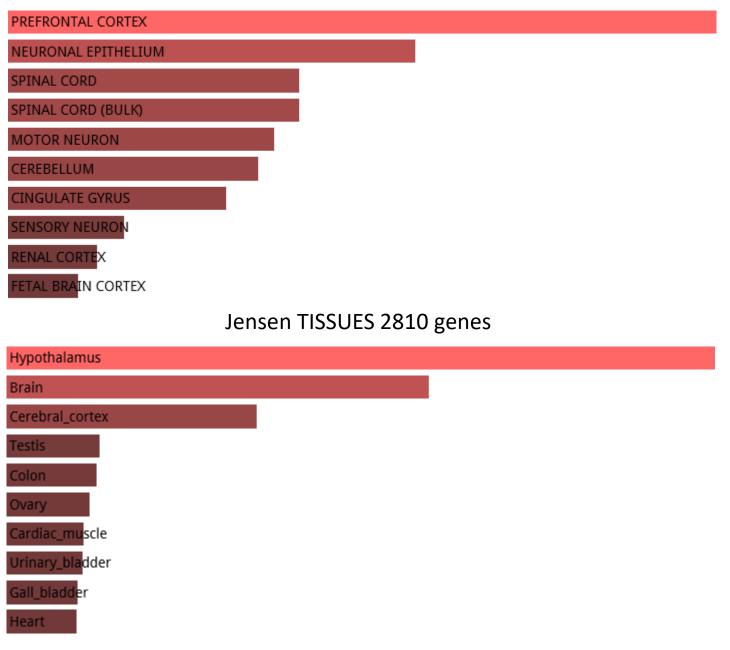
Spinal Muscular Atrophy C0026847 mouse GSE10599 sample 235

Disease Perturbations from GEO down

autism spectrum disorder DOID-0060041 human GSE62632 sample 1037

idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850	Bipolar Disorder C0005586 human GSE5389 sample 302
schizophrenia DOID-5419 human GSE25673 sample 891	allergic contact dermatitis DOID-3042 human GSE6281 sample 928
schizophrenia DOID-5419 human GSE25 <mark>6</mark> 73 sample 892	schizophrenia DOID-5419 human GSE25673 sample 892
Pauciarticular juvenile arthritis C0157917 human GSE1402 sample 430	idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850
Anterior Horn Cell Disease C0154681 human GSE1481 sample 208	prostate cancer DOID-10283 human GSE26910 sample 603
adrenoleukodystrophy DOID-10588 human GSE34309 sample 865	Non-syndromic cleft lip and palate DOID-9296 human GSE42589 sample 618
adrenoleukodystrophy DOID-10588 human GSE34308 sample 709	Alzheimer's disease DOID-10652 human GSE36980 sample 523
Gastrointestinal str <mark>o</mark> mal tumor C0238198 human GSE15966 sample 76	Bipolar Disorder C0005586 human GSE5388 sample 107
psoriasis DOID-8893 human GSE61281 sample 697	fragile X syndrome DOID-14261 human GSE7329 sample 809

ARCHS4 Tissues 2810 genes



4,249 fixed human-specific regulatory loci

Jensen DISEASES 2810 genes

Carcinoma

Kidney_cancer

Skin_cancer

Liver_cancer

Melanoma

Type_2_diabetes_mellitus

Breast_cancer

Acquired_metabolic_disease

Attention_deficit_hyperactivity_disorder

Endometrial_cancer

Disease Perturbations from GEO up

Disease Perturbations from GEO down

Spinal Muscular Atrophy C0026847 mouse GSE10599 sample 235	Crohn's disease DOID-8778 human GSE6731 sample 757
schizophrenia DOID-5419 human GSE25673 sample 892	adrenoleukodystrophy DOID-10588 human GSE34309 sample 864
Cardiomyopathy, Dilated C0007193 human GSE3585 sample 198	Ulcerative Colitis C0009324 human GSE6731 sample 249
multiple myeloma DOID-9538 human GSE47552 sample 561	autism spectrum disorder DOID-0060041 human GSE62632 sample 1037
adrenoleukodystrophy DOID-10588 human GSE34309 sample 864	ulcerative colitis DOID-8577 human GSE6731 sample 759
adrenoleukodystrophy DOID-10588 human GSE34308 sample 709	schizophrenia DOID-5419 human GSE25673 sample 892
Papillary Carcinoma of the Thyroid C0238463 human GSE3678 sample 306	Crohn's disease DOID-8778 human GSE6731 sample 758
oligodendroglioma DOID-3181 human GSE15824 sample 858	schizophrenia DOID-5419 human GSE25673 sample 891
Uterine leiomyoma C0042133 human GSE593 sample 16	ulcerative colitis DOID-8577 human GSE6731 sample 760
morbid obesity DOID-11981 human GSE48964 sample 583	colitis DOID-0060180 human GSE6731 sample 761

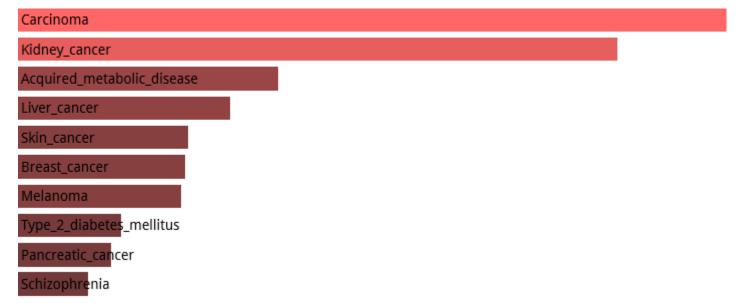
2,745 human accelerated regions

ARCHS4 Tissues 2281 genes

PREFRONTAL CORTEX	
SPINAL CORD	
SPINAL CORD (BULK)	
CEREBELLUM	
NEURONAL EPITHELIUM	
MOTOR NEURON	
CINGULATE GYRUS	
RENAL CORTEX	
SENSORY NEURON	
CEREBRAL CORTEX	
	Jensen TISSUES 2281 genes
Hypothalamus	
Hypothalamus Cerebral_cortex	
Cerebral_cortex	
Cerebral_cortex Brain	
Cerebral_cortex Brain Testis	
Cerebral_cortex Brain Testis Neural_crest	
Cerebral_cortex Brain Testis Neural_crest Bud	
Cerebral_cortex Brain Testis Neural_crest Bud Neural_tube	

2,745 human accelerated regions

Jensen DISEASES 2281 genes



Disease Perturbations from GEO up

Disease Perturbations from GEO down

Spinal Muscular Atrophy C0026847 mouse GSE10599 sample 235	schizophrenia DOID-5419 human GSE25673 sample 892
schizophrenia DOID-5419 human GSE25673 sample 891	Huntington's disease DOID-12858 mouse GSE3583 sample 929
schizophrenia DOID-5419 human GSE25673 sample 892	Crohn's disease DOID-8778 human GSE6731 sample 757
Huntington's disease DOID-12858 mouse GSE3583 sample 929	ulcerative colitis DOID-8577 human GSE6731 sample 759
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850	Down syndrome DOID-14250 human GSE42956 sample 1060
fragile X syndrome DOID-14261 human GSE7329 sample 515	schizophrenia DOID-5419 human GSE25673 sample 891
adrenoleukodystrophy DOID-10588 human GSE34308 sample 709	Ulcerative Colitis C0009324 human GSE6731 sample 249
adrenoleukodystrophy DOID-10588 human GSE34309 sample 864	autism spectrum disorder DOID-0060041 human GSE62632 sample 1036
Cardiomyopathy, Dilated C0007193 human GSE3585 sample 198	Breast Cancer C0006142 human GSE1378 sample 52
Down Syndrome C0013080 human GSE10758 sample 310	Bipolar Disorder C0005586 human GSE5389 sample 302

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

ARCHS4 Tissues 1458 genes

PREFRONTAL CORTEX CINGULATE GYRUS CEREBELLUM SPINAL CORD SPINAL CORD (BULK) MOTOR NEURON DORSAL STRIATUM SENSORY NEURON CEREBRAL CORTEX DENTATE GRANULE CELL Jensen TISSUES 1458 genes Hypothalamus Cerebral_cortex Brain Ganglion Neural_crest Mesenchyme Adult Mesenchymal_stem_cell 3T3-L1_cel

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

Spinal Muscular Atrophy C0026847 mouse GSE10599 sample 235

Jensen DISEASES 1458 genes

Kidney_cancer

Melanoma

Skin_cancer

Liver_cancer

Acquired_metabolic_disease

Breast_cancer

Lung_cancer

Type_2_dia_betes_mellitus

Attention_deficit_hyperactivity_disorder

Disease Perturbations from GEO up

Disease Perturbations from GEO down

schizophrenia DOID-5419 human GSE25673 sample 892

Spirial Muscular Acrophy Coozoo47 Mouse as E10333 sumple 233	Schizophi chia Dolb-5415 haman dee25075 sample 052
schizophrenia DOID-5419 human GSE25673 sample 891	adrenoleukodystrophy DOID-10588 human GSE34309 sample 864
schizophrenia DOID-5419 human GSE25673 sample 892	Crohn's disease DOID-8778 human GSE6731 sample 757
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850	schizophrenia DOID-5419 human GSE25673 sample 891
pulmonary sarcoidosis DOID-13406 human GSE16538 sample 916	Bipolar Disorder C0005586 human GSE5389 sample 302
smoldering myeloma DOID-9551 human GSE47552 sample 562	autism spectrum disorder DOID-0060041 human GSE62632 sample 1037
morbid obesity DOID-11981 human GSE48964 sample 583	glaucoma associated with systemic syndromes DOID-1686 mouse GSE26299 sample 488
multiple myeloma DOID-9538 human GSE47552 sample 561	allergic contact dermatitis DOID-3042 human GSE6281 sample 928
adrenoleukodystrophy DOID-10588 human GSE34308 sample 709	Gastrointestinal stromal tumor C0238198 human GSE15966 sample 76
adrenoleukodystrophy DOID-10588 human GSE34309 sample 864	idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850

1,000 human-biased cranial neural crest ARCHS4 Tissues 1439 genes cells (CNCCs) enhancers PREFRONTAL CORTEX SPINAL CORD SPINAL CORD (BULK) CEREBELLUM CINGULATE GYRUS CEREBRAL CORTEX MOTOR NEURON RENAL CORTEX DORSAL STRIATUM NEURONAL EPITHELIUM Jensen TISSUES 1439 genes Hypothalamus Brain Cerebral_cortex Uterus Esophagus Adrenal_gland Frontal_lobe Urinary_bladder Heart

Cerebellun

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

Jensen DISEASES 1439 genes

Carcinoma

Kidney_cancer

Neurogenic_bowel

Multiple_sclerosis

Type_2_diabetes_mellitus

Schizophrenia

Obesity

Liver_cancer

Bipolar_disorder

Heart_conduction_disease

Disease Perturbations from GEO up

Disease Perturbations from GEO down

Spinal Muscular Atrophy C0026847 mouse GSE10599 sample 235	Androgen insensitivity syndrome C0039585 human GSE3871 sample 415
schizophrenia DOID-5419 human GSE25673 sample 891	schizophrenia DOID-5419 human GSE25673 sample 892
adrenoleukodystrophy DOID-10588 human GSE34309 sample 864	idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850
schizophrenia DOID-5419 human GSE25673 sample 892	Huntington's disease DOID-12858 mouse GSE3583 sample 929
autism spectrum disorder DOID-0060041 human GSE63524 sample 1035	Down syndrome DOID-14250 human GSE42956 sample 1060
Acne C0702166 human GSE10432 sample 297	Crohn's disease DOID-8778 human GSE6731 sample 757
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 851	Alzheimer's disease DOID-10652 human GSE4757 sample 592
Cardiomyopathy, Dilated C0007193 human GSE3585 sample 198	idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 851
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850	ulcerative colitis DOID-8577 human GSE6731 sample 759
prostate cancer DOID-10283 human GSE26910 sample 603	Primary open angle glaucoma C0339573 human GSE2705 sample 257

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

ARCHS4 Tissues 1445 genes

SPINAL CORD SPINAL CORD (BULK) PREFRONTAL CORTEX RENAL CORTEX SMALL INTESTINE (BULK TISSUE) CEREBELLUM CINGULATE GYRUS NEURONAL EPITHELIUM ADIPOSE (BULK TISSUE)

LUNG (BULK TISSUE)

Spleen

Jensen TISSUES 1445 genes

Hypothalamus Brain Cerebral_cortex Uterus Heart Prostate_gland Ovary Placenta Esophagus

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

Jensen DISEASES 1445 genes

Kidney_cancer
Liver_cancer
Type_2_diabetes_mellitus
Melanoma
Skin_cancer
Acquired_metabolic_disease
Endometrial_cancer
Iridogoniodysgenesis_syndrome
Alzheimer's_disease

Disease Perturbations from GEO up

Acne C0702166 human GSE10432 sample 297

Primary open angle glaucoma C0339573 human GSE2705 sample 257

morbid obesity DOID-11981 human GSE48964 sample 583

Uterine leiomyoma C0042133 human GSE593 sample 16

adrenoleukodystrophy DOID-10588 human GSE34309 sample 864

schizophrenia DOID-5419 human GSE25673 sample 891

schizophrenia DOID-5419 human GSE25673 sample 892

pulmonary sarcoidosis DOID-13406 human GSE16538 sample 916

Type 2 diabetes mellitus C0011860 human GSE12643 sample 274

idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 851

Disease Perturbations from GEO down

Down syndrome DOID-14250 human GSE42956 sample 1060

autism spectrum disorder DOID-0060041 human GSE62632 sample 1036

schizophrenia DOID-5419 human GSE25673 sample 892

breast cancer DOID-1612 human GSE3744 sample 978

sporadic breast cancer DOID-8029 human GSE3744 sample 979

Breast Cancer C0006142 human GSE3744 sample 24

Primary open angle glaucoma C0339573 human GSE2705 sample 257

Androgen insensitivity syndrome C0039585 human GSE3871 sample 415

diabetes mellitus type 2 DOID-9352 human GSE12643 sample 766

melanoma in situ UMLS CUI-C0346040 human GSE4587 sample 980

1,619 human-specific functional hESC enhancers

ARCHS4 Tissues 1214 genes

PREFRONTAL CORTEX CEREBELLUM **NEURONAL EPITHELIUM** FETAL BRAIN CORTEX CINGULATE GYRUS CEREBRAL CORTEX MOTOR NEURON SPINAL CORD SPINAL CORD (BULK) BRAIN (BULK) Jensen TISSUES 1214 genes Hypothalamus Brain Heart Occipital_lobe Frontal_lobe Saliva Parietal_lobe Cerebral_cortex JEG-3_cell Temporal_lobe

1,619 human-specific functional hESC enhancers

Jensen DISEASES 1214 genes

Carcinoma

Liver_cancer

Skin_cancer

Melanoma

Breast_cancer

Acquired_metabolic_disease

Attention_deficit_hyperactivity_disorder

Lung_cancer

Major_dep_ressive_disorder

Disease Perturbations from GEO up 1214 genes

morbid obesity DOID-11981 human GSE48964 sample 583

Down Syndrome C0013080 human GSE10758 sample 310

Spinal Muscular Atrophy C0026847 mouse GSE10599 sample 235

schizophrenia DOID-5419 human GSE25673 sample 891

idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850

adenosquamous cell lung carcinoma DOID-4829 human GSE51852 sample 731

Urothelial carcinoma in situ C0334267 human GSE3167 sample 229

Malignant mesothelioma of pleura C0812413 human GSE2549 sample 118

asthma DOID-2841 human GSE31773 sample 715

smoldering myeloma DOID-9551 human GSE47552 sample 562

Disease Perturbations from GEO down 1214 genes

Gastrointestinal stromal tumor C0238198 human GSE15966 sample 76

autism spectrum disorder DOID-0060041 human GSE62632 sample 1037

schizophrenia DOID-5419 human GSE25673 sample 892

Bipolar Disorder C0005586 human GSE5389 sample 302

idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850

Nephroblastoma C0027708 human GSE2712 sample 418

autism spectrum disorder DOID-0060041 human GSE62632 sample 1036

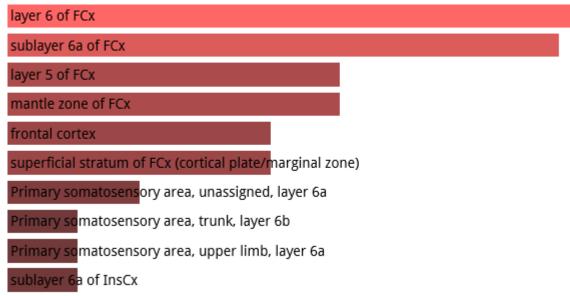
glaucoma associated with systemic syndromes DOID-1686 mouse GSE26299 sample 488

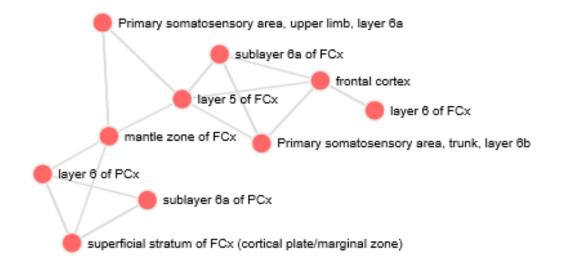
Bipolar Disorder C0005586 human GSE5388 sample 107

precursor B lymphoblastic lymphoma/leukemia DOID-7061 human GSE29639 sample 803

1,619 human-specific functional hESC enhancers

Allen Brain Atlas up 1214 genes



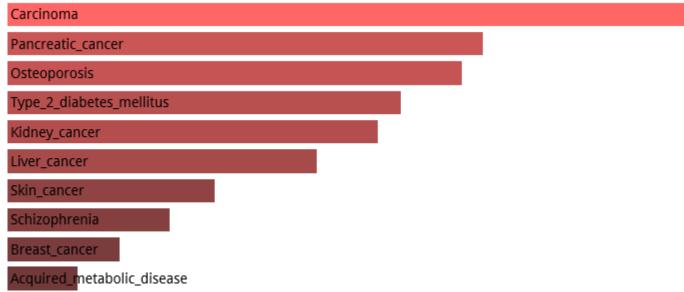


ARCHS4 Tissues 747 genes

PREFRONTAL CORTEX			
CEREBELLUM			
MIDBRAIN			
SPINAL CORD			
SPINAL CORD (BULK)			
CEREBRAL CORTEX			
CINGULATE GYRUS			
NEURONAL EPITHELIUM			
MOTOR NEURON			
RENAL CORTEX			
	Jensen TISSU	ES 747 genes	
Hypothalamus			
Brain			
Cerebral_cortex			
Neural_crest			
Testis			
Gall_bladder			
Parietal_lobe			
Frontal_lobe			
Occipital_lo <mark>be</mark>			
Urinary_bladder			

524 human-specific DHS

Jensen DISEASES 747 genes



Disease Perturbations from GEO up 747 genes

Disease Perturbations from GEO down 747 genes

adrenoleukodystrophy DOID-10588 human GSE34308 sample 709	Huntington's disease DOID-12858 mouse GSE3583 sample 929		
schizophrenia DOID-5419 human GSE25673 sample 891	Williams-Beuren syndrome DOID-1928 human GSE16715 sample 903		
Uterine leiomyoma C0042133 human GSE593 sample 16	Breast Cancer C0006142 human GSE1378 sample 52		
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 850	schizophrenia DOID-5419 human GSE25673 sample 892		
idiopathic pulmonary fibrosis DOID-0050156 human GSE44723 sample 851	Nephroblastoma C0027708 human GSE2712 sample 418		
polycystic ovary syndrome DOID-11612 human GSE10946 sample 825	Androgen insensitivity syndrome C0039585 human GSE3871 sample 415		
autism spectrum disorder DOID-0060041 human GSE63524 sample 1035	Bipolar Disorder C0005586 human GSE5389 sample 302		
morbid obesity DOID-11981 human GSE48964 sample 583	Papillary Carcinoma of the Thyroid C0238463 human GSE3467 sample 397		
Spinal Muscular Atrophy C0026847 mouse GSE10599 sample 235	breast cancer DOID-1612 human GSE3744 sample 978		
adrenoleukodystrophy DOID-10588 human GSE34309 sample 864	prostate cancer DOID-10283 human GSE26910 sample 603		

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

ARCHS4 Tissues random 6000 genes

ASTROCYTE

KIDNEY (BULK TISSUE)

GASTRIC TISSUE (BULK)

SPINAL CORD (BULK)

SMALL INTESTINE (BULK TISSUE)

NO SIGNIFICANT ENRICHMENT

OVARY (BULK TISSUE)

SPINAL CORD

PLACENTA (BULK)

SKIN (BULK TISSUE)

HUMAN EMBRYO

Jensen TISSUES random 6847 genes

Jensen DISEASES random 6000 genes

Reticular_dysgenesis

Amyotrophic_lateral_sclerosis_type_8

Succinic_semialdehyde_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

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 772

pancreatic ductal adenocarcinoma DOID-3498 mouse GSE53659 sample 699

amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 679

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 771

nemaline myopathy DOID-3191 mouse GSE3384 NOIS GNIFICANT ENRICHMENT

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

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 hMQh SIGNIFICANT7 ENRICHMENT

Disease Perturbations from GEO down random 6000 genes

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

Allen Brain Atlas up random 6000 genes

Anterior hypothalamic nucleus

Suprageniculate nucleus

Parabigeminal nucleus

Dorsal premammillary nucleus

NO SIGNIFICANT ENRICHMENT

Rhomboid nucleus

Presubiculum

Olivary pretectal nucleus

Thalamus

Anterodorsal nucleus

Field CA3

ARCHS4 Tissues random 6000 genes

GASTRIC TISSUE (BULK)

SKIN (BULK TISSUE)

OMENTUM

SMALL INTESTINE (BULK TISSUE)

CINGULATE GYRUS

NO SIGNIFICANT ENRICHMENT

COLON (BULK TISSUE)

NEURONAL EPITHELIUM

PREFRONTAL CORTEX

PANCREATIC ISLET

MYOBLAST

Jensen TISSUES random 6847 genes

Jensen DISEASES random 6000 genes

Salpingitis

Gastric antral vascular ectasia

Cockayne_syndrome

Early_myoclonic_encephalopathy

Arterial tortuosity syndrome

Vaginitis

cold-induced_sweating_syndrome

Parotitis

Melanotic_neuroectodermal_tumor

Warsaw_breakage_syndrome

NO SIGNIFICANT ENRICHMENT

Disease Perturbations from GEO up random 6000 genes

progressive myoclonus epilepsy DOID-891 mouse GSE47516 sample 516

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
NO SIGNIFICANT ENRICHMENT

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 771

nemaline myopathy DOID-3191 mouse GSE3384 sample 971

nemaline myopathy DOID-3191 mouse GSE3384 sample 974

Nemaline Myopathy C0206157 mouse GSE3384 sample 317

acute myocarditis DOID-3951 mouse GSE35182 sample 802

pulmonary tuberculosis DOID-2957 mouse GSE48027 sample 832

Disease Perturbations from GEO down random 6000 genes

acute myocarditis DOID-3951 mouse GSE35182 sample 801

pulmonary tuberculosis DOID-2957 mouse GSE48027 sample 831

Purpura, Idiopathic Thrombocytopenic C0043117 human GSE574 sample 358

autistic disorder DOID-12849 human GSE6575 sample 1043 NO SIGNIFICANT ENRICHMENT

sickle-cell anemia DOID-10923 human GSE16728 sample 506

amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 679

sickle-cell anemia DOID-10923 human GSE16728 sample 505

Down syndrome DOID-14250 mouse GSE39159 sample 529

acute myocarditis DOID-3951 mouse GSE35182 sample 802

polycystic ovary syndrome DOID-11612 human GSE48301 sample 558

Allen Brain Atlas up random 6000 genes

Anterior hypothalamic nucleus

Suprageniculate nucleus

Parabigeminal nucleus

Dorsal premammillary nucleus

NO SIGNIFICANT ENRICHMENT

Presubiculum

Rhomboid nucleus

Field CA1, stratum radiatum

Olivary pretectal nucleus

Thalamus

Anterodorsal nucleus

ARCHS4 Tissues random 6000 genes

COLONIC MUCOSA

VENTRICLE

PREFRONTAL CORTEX

BRAIN (BULK)

NO SIGNIFICANT ENRICHMENT

SUPERIOR FRONTAL GYRUS

NEURONAL EPITHELIUM

ADIPOSE (BULK TISSUE)

HUMAN ZYGOTE

RENAL CORTEX

MYOBLAST

Jensen TISSUES random 6847 genes

Jensen DISEASES random 6000 genes

NO SIGNIFICANT ENRICHMENT

Hereditary_mucosal_leukokeratosis

Echolalia

Dumping_syndrome

X-linked_endothelial_corneal_dystrophy

IMAGe_syndrome

Clouston_syndrome

Oculodentodigital_dysplasia

Irritant_dermatitis

Disease Perturbations from GEO up random 6000 genes

Brain_glioma

Pachyonychia_congenita

Disease Perturbations from GEO down 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 **SIGNIFICANT ENRICHMENT** 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

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
NO SIGNIFICANT ENRICHMENT

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

Allen Brain Atlas up random 6000 genes

Anterior hypothalamic nucleus

Suprageniculate nucleus

Parabigeminal nucleus

Dorsal premammillary nucleus

Presubiculum

NO SIGNIFICANT ENRICHMENT

Rhomboid nucleus

Field CA1, stratum radiatum

Olivary pretectal nucleus

Thalamus

Anterodorsal nucleus

ARCHS4 Tissues random 6000 genes

GASTRIC TISSUE (BULK)

PREFRONTAL CORTEX

COLON (BULK TISSUE)

LUNG (BULK TISSUE)

NO SIGNIFICANT ENRICHMENT

OMENTUM

NEURONAL EPITHELIUM

LIVER (BULK TISSUE)

ILEUM (BULK)

FETAL BRAIN CORTEX

SUPERIOR FRONTAL GYRUS

Jensen TISSUES random 6847 genes

Jensen DISEASES random 6000 genes

Urinary_schistosomiasis

Hermaphroditism

Zellweger syndrome

Latex_allergy

NO SIGNIFICANT ENRICHMENT

Scimitar_syndrome

Conversion_disorder

Gastric_lymphoma

Juvenile_polyposis_syndrome

Invasive_lobular_carcinoma

Acute_promyelocytic_leukemia

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

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
NO SIGNIFICANT ENRICHMENT

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

Allen Brain Atlas up random 6000 genes

Anterior hypothalamic nucleus

Suprageniculate nucleus

Parabigeminal nucleus

Dorsal premammillary nucleus

Presubiculum

NO SIGNIFICANT ENRICHMENT

Rhomboid nucleus

Field CA1, stratum radiatum

Olivary pretectal nucleus

Thalamus

Anterodorsal nucleus

ARCHS4 Tissues random 6000 genes

SMALL INTESTINE (BULK TISSUE)

VASCULAR SMOOTH MUSCLE

SKIN (BULK TISSUE)

THYROID (BULK TISSUE)

STROMAL CELL

NO SIGNIFICANT ENRICHMENT

RESPIRATORY SMOOTH MUSCLE

RENAL CORTEX

SKELETAL MUSCLE (BULK TISSUE)

REGULATORY T CELLS

PANCREATIC ISLET

Jensen TISSUES random 6847 genes

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 GSE2NG 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

Allen Brain Atlas up random 6000 genes

Presubiculum

Endopiriform nucleus, dorsal part

dorsal endopiriform nucleus

Copula pyramidis

NO SIGNIFICANT ENRICHMENT

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

ARCHS4 Tissues random 6000 genes

NO DATA AVAILABLE

Jensen TISSUES random 6000 genes

NO DATA AVAILABLE

Disease Perturbations from GEO up random 6000 genes

Disease Perturbations from GEO down random 6000 genes

Systemic lupus erythematosus (SLE) C0024141 human GSE12374 sample 123

NO DATA AVAILABLE

NO SIGNIFICANT ENRICHMENT

ARCHS4 Tissues random 6847 genes

TLYMPHOCYTE

OMENTUM

GASTRIC TISSUE (BULK)

PREFRONTAL CORTEX

SPINAL CORD (BULK)

NO SIGNIFICANT ENRICHMENT

SPINAL CORD

LUNG (BULK TISSUE)

CARDIAC MUSCLE FIBER

BREAST (BULK TISSUE)

NEURONAL EPITHELIUM

Jensen TISSUES random 6847 genes

Jensen DISEASES random 6847 genes

NO SIGNIFICANT ENRICHMENT

epididymo-orchitis

Vestibular_neuronitis

Hemophagocytic_lymphohistiocytosis

Kidney_angiomyolipoma

Acrodermatitis_enteropathica

Eclampsia

Choroid_plexus_papilloma

Angiosarcoma

Akinetic_mutism

Venous insufficiency

Disease Perturbations from GEO up random 6000 genes

Nicotine addiction C0028043 human GSE11208 sample 325

Cardiac Hypertrophy C1383860 rat GSE1055 sample 354

Alzheimer's disease DOID-10652 human GSE36980 sample 524

autism spectrum disorder DOID-0060041 human GSE62632 sample 1037
NO SIGNIFICANT ENRICHMENT

multiple sclerosis DOID-2377 human GSE16461 sample 584

polycystic ovary syndrome DOID-11612 human GSE48301 sample 558

Nephroblastoma C0027708 human GSE2712 sample 418

Alzheimer's disease DOID-10652 human GSE36980 sample 520

testis seminoma DOID-5842 human GSE8607 sample 624

Intestinal polyposis UMLS CUI-C1257915 mouse GSE6078 sample 952

Disease Perturbations from GEO down random 6000 genes

Purpura, Idiopathic Thrombocytopenic C0043117 human GSE574 sample 358

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

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 772

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 771
NO SIGNIFICANT ENRICHMENT

Sarcoidosis C0036202 human GSE2657 sample 12

Down syndrome DOID-14250 mouse GSE39159 sample 529

Eczema C0013595 human GSE6012 sample 268

smoldering myeloma DOID-9551 human GSE47552 sample 562

autistic disorder DOID-12849 human GSE6575 sample 1043

dermatomyositis DOID-10223 human GSE48280 sample 705

Allen Brain Atlas up random 6847 genes

Anterior hypothalamic nucleus

Parabigeminal nucleus

Suprageniculate nucleus

Dorsal premammillary nucleus

Presubiculum

NO SIGNIFICANT ENRICHMENT

Rhomboid nucleus

Field CA1, stratum radiatum

Thalamus

lateral posterior nucleus

Anterodorsal nucleus

ARCHS4 Tissues random 6001 genes

GASTRIC TISSUE (BULK)

SKIN (BULK TISSUE)

RENAL CORTEX

KIDNEY (BULK TISSUE)

NO SIGNIFICANT ENRICHMENT

PREFRONTAL CORTEX

PANCREATIC ISLET

SPINAL CORD (BULK)

SMALL INTESTINE (BULK TISSUE)

COLON (BULK TISSUE)

SPINAL CORD

Jensen TISSUES random 6001 genes

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

Disease Perturbations from GEO down random 6000 genes

pancreatic ductal adenocarcinoma DOID-3498 mouse GSE53659 sample 699

Schistosomiases C0036323 mouse GSE19525 sample 439

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 772

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 771
NO SIGNIFICANT ENRICHMENT

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

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
NO SIGNIFICANT ENRICHMENT

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

Allen Brain Atlas up random 6001 genes

Anterior hypothalamic nucleus

Suprageniculate nucleus

Parabigeminal nucleus

Dorsal premammillary nucleus

Field CA3, stratum oriens

NO SIGNIFICANT ENRICHMENT

Anterior tegmental nucleus

Facial motor nucleus

Field CA1

Presubiculum

Anteromedial nucleus

ARCHS4 Tissues random 6000 genes

COLONIC MUCOSA

PREFRONTAL CORTEX

VENTRICLE

ASTROCYTE

NO SIGNIFICANT ENRICHMENT

MYOBLAST

BRAIN (BULK)
FETAL BRAIN CORTEX

NEURONAL EPITHELIUM

ATRIUM

ADIPOSE (BULK TISSUE)

Jensen TISSUES random 6000 genes

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 SIGNIFICAN ENRICHEMNT

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
NO SIGNIFICANT ENRICHMENT

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

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
NO SIGNIFICANT ENRICHMENT

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

Allen Brain Atlas up random 6000 genes

Anterior hypothalamic nucleus

Suprageniculate nucleus

Parabigeminal nucleus

Dorsal premammillary nucleus

Field CA3, stratum oriens

NO SIGNIFICANT ENRICHMENT

Anterior tegmental nucleus

Facial motor nucleus

Field CA1

Presubiculum

Anteromedial nucleus

ARCHS4 Tissues random 6000 genes

GASTRIC TISSUE (BULK)

PREFRONTAL CORTEX

OMENTUM

COLON (BULK TISSUE)

LUNG (BULK TISSUE)

NO SIGNIFICANT ENRICHMENT

NEURONAL EPITHELIUM

LIVER (BULK TISSUE)

FETAL BRAIN CORTEX

RENAL CORTEX

AMNIOTIC FLUID

Jensen TISSUES random 6000 genes

Jensen DISEASES random 6000 genes

Laryngeal squamous_cell_carcinoma

Urinary_schistosomiasis

Ebola hemorrhagic fever

Hermaphroditism

Irritant dermatitis

NO SIGNIFICANT ENRICHMENT

Zellweger_syndrome

Aortic_valve_insufficiency

Scimitar_syndrome

Gastric_lymphoma

Multiple_chemical_sensitivity

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 NO SIGNIFICANT ENRICHMENT

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

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 NO SIGNIFICANT ENRICHMENT

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

Allen Brain Atlas up random 6000 genes

Anterior hypothalamic nucleus

Suprageniculate nucleus

Parabigeminal nucleus

Dorsal premammillary nucleus

Presubiculum

NO SIGNIFICANT ENRICHMENT

Rhomboid nucleus

Field CA1, stratum radiatum

Olivary pretectal nucleus

Thalamus

Anterodorsal nucleus

ARCHS4 Tissues random 4000 genes

OMENTUM

VENTRICLE

SKIN (BULK TISSUE)

WHARTONS JELLY

NO SIGNIFICANT ENRICHMENT

VASCULAR SMOOTH MUSCLE

OVARY (BULK TISSUE)

CEREBELLUM

CHONDROCYTE

THYROID (BULK TISSUE)

SPINAL CORD (BULK)

Jensen TISSUES random 4000 genes

Jensen DISEASES random 4000 genes

Trichotillomania

Exhibitionism

Yaws

Amyotrophic_lateral_sclerosis_type_8

NO SIGNIFICANT ENRICHMENT Nonepidermolytic_palmoplantar_keratoderma

Giardiasis

Intrahepatic_cholestasis

Pulmonary_alveolar_proteinosis

Normal_pressure_hydrocephalus

Amyotrophic_neuralgia

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

Cancer of the Intestine C0346627 mouse GSE3915 sample 90

schizophrenia DOID-5419 human GSE12679 sample 767 NO SIGNIFICANT ENRICHMENT

prostate cancer DOID-10283 human GSE3325 sample 991

nemaline myopathy DOID-3191 mouse GSE3384 sample 971

Primary pulmonary hypoplasia C0456891 mouse GSE1363 sample 20

Nephrolithiasis C0392525 mouse GSE10162 sample 187

Acute Lung Injury C0242488 human GSE10474 sample 168

adrenoleukodystrophy DOID-10588 human GSE34309 sample 864

Disease Perturbations from GEO down random 4000 genes

Duchenne muscular dystrophy (DMD) C0013264 mouse GSE466 sample 328

Huntington's disease DOID-12858 human GSE8762 sample 931

systemic lupus erythematosus DOID-9074 human GSE10325 sample 692

Pauciarticular juvenile arthritis C0157917 human GSE1402 sample 430
NO SIGNIFICANT ENRICHMENT

Infantile neuronal ceroid lipofuscinosis C0268281 mouse GSE6678 sample 141

Retinitis Pigmentosa C0035334 mouse GSE128 sample 33

Asthma, allergic C0155877 mouse GSE2276 sample 348

Turner syndrome DOID-3491 human GSE58435 sample 1059

pancreatic ductal adenocarcinoma DOID-3498 mouse GSE61412 sample 698

Kidney disorder associated with type 2 diabetes mellitus C1720457 mouse GSE2557 sample 8

Allen Brain Atlas up random 4000 genes

Presubiculum

Anterior tegmental nucleus

Hindbrain

Taenia tecta

NO SIGNIFICANT ENRICHMENT

lateral posterior nucleus

Field CA1, stratum oriens

Facial motor nucleus

Basolateral amygdalar nucleus, ventral part

Primary visual area, layer 4

Cortical plate

ARCHS4 Tissues random 4000 genes

GASTRIC TISSUE (BULK)

RENAL CORTEX

SKIN (BULK TISSUE)

SMALL INTESTINE (BULK TISSUE)

NO SIGNIFICANT ENRICHMENT

PANCREATIC ISLET

PREFRONTAL CORTEX

KIDNEY (BULK TISSUE)

SPINAL CORD (BULK)

PLACENTA (BULK)

SPINAL CORD

Jensen TISSUES random 4000 genes

Jensen DISEASES random 4000 genes

Poliomyelitis

Salpingitis

Vaginitis

Parotitis

NO SIGNIFICANT ENRICHMENT

Reticular_dysgenesis

Succinic_semialdehyde_dehydrogenase_deficiency

Complement_deficiency

Borna_disease

Chickenpox

Amyotrophic_lateral_sclerosis_type_8

Disease Perturbations from GEO up random 4000 genes

Breast Cancer C0006142 human GSE3744 sample 24

polycystic ovary syndrome DOID-11612 human GSE48301 sample 558

acute myeloid leukemia DOID-9119 human GSE9476 sample 783

polycystic ovary syndrome DOID-11612 human GSE48301 sample 559
NO SIGNIFICANT ENRICHMENT

breast cancer DOID-1612 human GSE3744 sample 978

nemaline myopathy DOID-3191 mouse GSE3384 sample 976

sporadic breast cancer DOID-8029 human GSE3744 sample 979

autism spectrum disorder DOID-0060041 human GSE62632 sample 1037

Crohn's disease DOID-8778 human GSE1710 sample 992

schizophrenia DOID-5419 human GSE12679 sample 769

Disease Perturbations from GEO down random 4000 genes

acute myocarditis DOID-3951 mouse GSE35182 sample 801

pulmonary tuberculosis DOID-2957 mouse GSE48027 sample 831

Schistosomiases C0036323 mouse GSE19525 sample 439

Purpura, Idiopathic Thrombocytopenic C0043117 human GSE574 sample 358

sickle-cell anemia DOID-10923 human GSE16728 sample 505

amyotrophic lateral sclerosis DOID-332 mouse GSE10953 sample 679

West Nile fever DOID-2366 human GSE30719 sample 874

autistic disorder DOID-12849 human GSE6575 sample 1043

Septic Shock C0036983 human GSE9692 sample 307

Down syndrome DOID-14250 mouse GSE39159 sample 529

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

ARCHS4 Tissues random 4000 genes

COLON (BULK TISSUE)

ILEUM (BULK)

SMALL INTESTINE (BULK TISSUE)

GASTRIC TISSUE (BULK)

OMENTUM

NO SIGNIFICANT ENRICHMENT

NEURONAL EPITHELIUM

SKIN (BULK TISSUE)

CEREBELLUM

PREFRONTAL CORTEX

MYOBLAST

Jensen TISSUES random 4000 genes

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 an emia 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

Allen Brain Atlas up random 4000 genes

Presubiculum

Anterior hypothalamic nucleus

Olivary pretectal nucleus

Parabigeminal nucleus

Suprageniculate nucleus

Dorsal premammillary nucleus

Anterodorsal nucleus

Anterior tegmental nucleus

superficial stratum of DG

Hindbrain

NO SIGNIFICANT ENRICHMENT

ARCHS4 Tissues random 4000 genes

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

CD19+ B CELLS

BLYMPHOCYTE

NO SIGNIFICANT ENRICHMENT

COLON (BULK TISSUE)

PLASMA CELL

GASTRIC TISSUE (BULK)

BLASTOCYST

PERIPHERAL BLOOD

SPLEEN (BULK TISSUE)

Jensen TISSUES random 4000 genes

Jensen DISEASES random 4000 genes

Hereditary_mucosal_leukokeratosis p-value: 0.017; Adjusted p-value: 1.00 p-value: 0.038; Adjusted p-value: 1.00 X-linked endothelial corneal dystrophy Dumping_syndrome Echolalia NO SIGNIFICANT ENRICHMENT Clouston syndrome Oculodentodigital_dysplasia Irritant dermatitis Fabry_disease Granulomatous_amebic_encephalitis

Disease Perturbations from GEO up random 4000 genes

Pachyonychia_congenita

Dental cavity, complex C0399396 human GSE1629 sample 175

actinic keratosis DOID-8866 human GSE2503 sample 628

Actinic keratosis C0022602 human GSE2503 sample 350

lupus erythematosus DOID-8857 human GSE30153 sample 739
NO SIGNIFICANT ENRICHMENT

acute myeloid leukemia DOID-9119 human GSE9476 sample 782

psoriasis DOID-8893 mouse GSE27628 sample 823

ulcerative colitis DOID-8577 human GSE9452 sample 924

type 2 diabetes mellitus DOID-9352 human GSE23343 sample 895

cardiomyopathy DOID-0050700 human GSE9128 sample 781

Epilepsy C0014544 human GSE7486 sample 417

Disease Perturbations from GEO down random 4000 genes

multiple myeloma DOID-9538 human GSE6691 sample 787

Smoldering multiple myeloma C1531608 human GSE5900 sample 404

autistic disorder DOID-12849 human GSE6575 sample 1042

acute myocarditis DOID-3951 mouse GSE35182 sample 801 NO SIGNIFICANT ENRICHMENT

mental retardation DOID-1059 human GSE6575 sample 1044

autistic disorder DOID-12849 human GSE6575 sample 1043

Waldenstrom Macroglobulinemia UMLS CUI-C0024419 human GSE6691 sample 785

Huntington's disease DOID-12858 human GSE8762 sample 930

pulmonary tuberculosis DOID-2957 mouse GSE48027 sample 831

non-systemic juvenile idiopathic arthritis (sJIA) (subgroup-RF-polyarthritis) DOID-848 human GSE21521 sample 5

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

NO SIGNIFICANT ENRICHMENT

Lateral reticular nucleus, parvicellular part

mantle zone of r9Ve

r4 part of magnocellular medial vestibular nucleus

periventricular stratum of r8Tr

superficial stratum of r7Lim

ARCHS4 Tissues random 4000 genes

COLON (BULK TISSUE)

GASTRIC TISSUE (BULK)

PREFRONTAL CORTEX

ILEUM (BULK)

LUNG (BULK TISSUE)

NEURONAL EPITHELIUM

OMENTUM

LIVER (BULK TISSUE)

AMNIOTIC FLUID

RENAL CORTEX

NO SIGNIFICANT ENRICHMENT

Jensen TISSUES random 4000 genes

Jensen DISEASES random 4000 genes

NO SIGNIFICANT ENRICHMENT

Hermaphroditism

Urinary schistosomiasis

Mitochondrial_myopathy

Zellweger_syndrome

Aortic_valve_insufficiency

Latex_allergy

Pulpitis

Dowling-Degos_disease

Walker-Warburg_syndrome

Peroxisomal_disease

Disease Perturbations from GEO up random 4000 genes

Alzheimer's disease DOID-10652 human GSE36980 sample 519

fetal alcohol spectrum disorder DOID-0050696 mouse GSE34305 sample 526

Cardiac Hypertrophy C1383860 rat GSE1055 sample 354

polycystic ovary syndrome DOID-11612 human GSE48301 sample 558
NO SIGNIFICANT ENRICHMENT

ulcerative colitis DOID-8577 human GSE37283 sample 594

Neurogenic Muscular Atrophy C0270948 rat GSE2566 sample 396

facioscapulohumeral muscular dystrophy DOID-11727 human GSE15090 sample 541

Bipolar Disorder C0005586 human GSE5389 sample 302

Crohn's disease DOID-8778 human GSE1710 sample 992

nemaline myopathy DOID-3191 mouse GSE3384 sample 976

Disease Perturbations from GEO down random 4000 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
NO SIGNIFICANT ENRICHMENT

papillary thyroid carcinoma DOID-3969 human GSE54958 sample 652

Down syndrome DOID-14250 mouse GSE39159 sample 529

Nemaline Myopathy C0206157 mouse GSE3384 sample 317

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 771

Viral cardiomyopathy C0264797 human GSE4172 sample 384

juvenile dermatomyositis UMLS CUI-C0263666 human GSE11971 sample 772

Allen Brain Atlas up random 4000 genes

rhombomere 8

retropontine reticular area

medullary hindbrain (medulla)

Visceral area

colliculus superior

NO SIGNIFICANT ENRICHMENT

mantle zone of SC

superficial gray layer of SC

Retrosplenial area, dorsal part, layer 1

Paragigantocellular reticular nucleus, dorsal part

Gustatory areas

ARCHS4 Tissues random 4000 genes

PREFRONTAL CORTEX

SPINAL CORD (BULK)

SUPERIOR FRONTAL GYRUS

OMENTUM

NO SIGNIFICANT ENRICHMENT

SPINAL CORD

LUNG (BULK TISSUE)

NEURONAL EPITHELIUM

FETAL BRAIN CORTEX

ASTROCYTE

BRAIN (BULK)

Jensen TISSUES random 4000 genes

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
NO SIGNIFICANT ENRICHMENT

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

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
NO SIGNIFICANT ENRICHMENT

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

Schistosomiases C0036323 mouse GSE19525 sample 439

Allen Brain Atlas up random 4000 genes

Presubiculum

Anterior hypothalamic nucleus

Olivary pretectal nucleus

Parabigeminal nucleus

Suprageniculate nucleus

Dorsal premammillary nucleus

Anterodorsal nucleus

Anterior tegmental nucleus

superficial stratum of DG

Hindbrain

NO SIGNIFICANT ENRICHMENT

ARCHS4 Tissues random 4000 genes

SMALL INTESTINE (BULK TISSUE)

VASCULAR SMOOTH MUSCLE

SKIN (BULK TISSUE)

THYROID (BULK TISSUE)

STROMAL CELL

NO SIGNIFICANT ENRICHMENT

RESPIRATORY SMOOTH MUSCLE

RENAL CORTEX

SKELETAL MUSCLE (BULK TISSUE)

REGULATORY T CELLS

PANCREATIC ISLET

Jensen TISSUES random 4000 genes

Jensen DISEASES random 4000 genes

NO SIGNIFICANT ENRICHMENT

Enthesopathy

Gyrate_atrophy

Vitelliform macular dystrophy

Opisthorchiasis

Amelogenesis_imperfecta

3-methylglutaconic_aciduria

Mastitis

Cone_dystrophy

Hydrocephalus

Fundus_dystrophy

Disease Perturbations from GEO up random 4000 genes

colorectal adenocarcinoma DOID-0050861 mouse GSE31106 sample 1070

Acute Lung Injury C0242488 human GSE10474 sample 168

neurofibromatosis DOID-8712 mouse GSE1482 sample 666

Type 2 diabetes mellitus C0011860 mouse GSE2899 sample 279
NO SIGNIFICANT ENRICHMENT

Sepsis C0243026 mouse GSE4479 sample 150

Type 2 diabetes mellitus C0011860 rat GSE6428 sample 291

Huntington's disease DOID-12858 mouse GSE3248 sample 728

Huntington's disease DOID-12858 human GSE3248 sample 724

ovarian cancer DOID-2394 human GSE14407 sample 611

Turner Syndrome C0041408 mouse GSE1606 sample 218

Disease Perturbations from GEO down random 4000 genes

Down syndrome DOID-14250 mouse GSE39159 sample 529

Duchenne muscular dystrophy (DMD) C0013264 mouse GSE466 sample 328

Cardiomyopathy, Dilated C0007193 human GSE3585 sample 198

Duchenne muscular dystrophy (DMD) C0013264 mouse GSE1026 sample 72
NO SIGNIFICANT ENRICHMENT

Premature aging C0231341 human GSE10123 sample 55

Alzheimer's disease DOID-10652 human GSE36980 sample 524

Down syndrome DOID-14250 human GSE42956 sample 1061

psoriasis DOID-8893 human GSE53431 sample 813

diabetic retinopathy DOID-8947 mouse GSE12610 sample 911

Tendonopathy 971 human GSE26051 sample 470

Allen Brain Atlas up random 4000 genes

Presubiculum

Endopiriform nucleus, dorsal part

dorsal endopiriform nucleus

Copula pyramidis

bed nucleus of the external capsule

NO SIGNIFICANT ENRICHMENT

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

Jensen TISSUES random 4000 genes

Disease Perturbations from GEO up random 4000 genes

Disease Perturbations from GEO down random 4000 genes

NO DATA AVAILABLE

Jensen TISSUES random 4000 genes

Disease Perturbations from GEO up random 4000 genes

Disease Perturbations from GEO down random 4000 genes

NO DATA AVAILABLE

ARCHS4 Tissues random 4000 genes

TLYMPHOCYTE

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

Jensen DISEASES random 4000 genes

NO SIGNIFICANT ENRICHMENT

Acrodermatitis_enteropathica

Vestibular_neuronitis

Acute_poststreptococcal_glomerulonephritis

Diastrophic_dysplasia

Eclampsia

Biotin_deficiency

Schimke_immuno-osseous_dysplasia

Bowen-Conradi_syndrome

Chikungunya

Western_equine_encephalitis

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
NO SIGNIFICANT ENRICHMENT

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

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 NO SIGNIFICANT ENRICHMENT

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

Allen Brain Atlas up random 4000 genes

Presubiculum

Anterior hypothalamic nucleus

Olivary pretectal nucleus

Parabigeminal nucleus

Suprageniculate nucleus

Dorsal premammillary nucleus

Anterodorsal nucleus

Anterior tegmental nucleus

superficial stratum of DG

Hindbrain

NO SIGNIFICANT ENRICHMENT

ARCHS4 Tissues random 2847 genes

REGULATORY T CELLS

BLOOD PBMC

PERIPHERAL BLOOD

TLYMPHOCYTE

THYMUS (BULK TISSUE)

NATURAL KILLER CELLS

ERYTHROBLAST

BLYMPHOCYTE

SPERM

Jensen TISSUES random 2847 genes

Radicle		p-value: 0.042; Adjusted p-value: 1.00
e2		
Blood_vessel		
Crfk		
HFL-1_cell		
Lens_cortex	NO SIGNIFICANT ENRICHMENT	
SW-1116_cell		
Nthy-ori_3-1_cell		
5637_cell		
Pancreatic_acinar_cell		

Jensen DISEASES random 2847 genes

epididymo-orchitis p-value: 0.041; Adjusted p-value: 1.00

NO SIGNIFICANT ENRICHMENT

Hemophagocytic_lymphohistiocytosis

Hyperthyroxinemia

Subependymal_giant_cell_astrocytoma

chorea-acanthocytosis

Neurofibroma

Adrenocortical_carcinoma

Mulibrey_nanism

ARC_syndrome

Septic_arthritis

Disease Perturbations from GEO up random 2847 genes

Cardiac Hypertrophy C1383860 rat GSE1055 sample 354

polycystic ovary syndrome 11612 human GSE48301 sample 468

systemic lupus erythematosus DOID-9074 human GSE55447 sample 1073

autism spectrum disorder DOID-0060041 human GSE62632 sample 1037
NO SIGNIFICANT ENRICHMENT

Teratospermia UMLS CUI-C0919628 human GSE6968 sample 954

abdominal aortic aneurysm DOID-7693 human GSE47472 sample 1011

papillary thyroid carcinoma DOID-3969 human GSE54958 sample 652

Neurogenic Muscular Atrophy C0270948 rat GSE2566 sample 396

mental retardation DOID-1059 human GSE6575 sample 1047

Crohn's disease DOID-8778 human GSE1710 sample 992

Disease Perturbations from GEO down random 2847 genes

hepatocellular carcinoma DOID-684 human GSE58208 sample 735

Huntington's disease DOID-12858 human GSE8762 sample 930

autistic disorder DOID-12849 human GSE6575 sample 1043

Nemaline Myopathy C0206157 mouse GSE3384 sample 317
NO SIGNIFICANT ENRICHMENT

Down syndrome DOID-14250 mouse GSE39159 sample 529

nemaline myopathy DOID-3191 mouse GSE3384 sample 974

nemaline myopathy DOID-3191 mouse GSE3384 sample 970

Schistosomiases C0036323 mouse GSE19525 sample 439

multiple myeloma DOID-9538 human GSE6691 sample 787

acute myocardial infarction DOID-9408 mouse GSE775 sample 1005

Allen Brain Atlas up random 2847 genes

superficial stratum of r4BL

r8 part of basomedial reticular formation

r4 part of ventral lateral lemniscal nucleus

intermediate stratum of r10Tr

r10 part of dorsal parvicellular reticular format NO SIGNIFICANT ENRICHMENT

Paragigantocellular reticular nucleus

Medial pretectal area

Dentate nucleus

r9 part of dorsal parvicellular reticular formation

Posterolateral visual area, layer 2/3

GO Biological Process 2017b 2847 genes

sensory perception of bitter taste (GO:0050913)

sensory perception of taste (GO:0050909)

regulation of tau-protein kinase activity (GO:1902947)

regulation of protein histidine kinase activity (GO:0032110)

regulation of protein serine/threonine kinase activity (GO:0071900)

detection of chemical stimulus involved in sensory perception of bitter taste (GO:0001580)

positive regulation of mitochondrial translation in esponse to stress (35:50) RICHMENT

regulation of protein localization to prospore membrane (GO:2001231)

regulation of protein localization to plasma membrane (GO:1903076)

positive regulation of toll-like receptor 2 signaling pathway (GO:0034137)

p-value: 0.0048; Adjusted p-value: 1.00

p-value: 0.0089; Adjusted p-value: 1.00

p-value: 0.0123; Adjusted p-value: 1.00

p-value: 0.0123; Adjusted p-value: 1.00

p-value: 0.0220; Adjusted p-value: 1.00

GO Molecular Function 2017 2847 genes

bitter taste receptor activity (GO:0033038)

trace-amine receptor activity (GO:0001594)

taste receptor activity (GO:0008527)

ubiquinol-cytochrome-c reductase activity (GO:0008121)

protein-glutamine gamma-glutamyltransferase activity (GO:0003810)

ubiquitin-ubiquitin ligase activity (GO:0034450)

NO SIGNIFICANT ENRICHMENT

TPR domain binding (GO:0030911)

NEDD8 transferase activity (GO:0019788)

troponin I binding (GO:0031013)

tumor necrosis factor receptor binding (GO:0005164)

p-value: 0.0029; Adjusted p-value: 0.836

p-value: 0.01; Adjusted p-value: 1.00

p-value: 0.013; Adjusted p-value: 1.00

p-value: 0.018; Adjusted p-value: 1.00

Pfam InterPro Domains 2847 genes

Znf_C2H2 p-value: 0.000034; Adjusted p-value: 0.0043
Whey_acidic_protein_4-dis_core

UBX

Syntaxin_N

Znf_Bbox

Trnsglumase_like

Znf_DHHC

Znf_C2CH

Ubiquitin

Emp24_gp25L_p24

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)

NO SIGNIFICANT ENRICHMENT

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 Revalues 0.028; Adjusted p-value: 1.00

ubiquinol-cytochrome-c reductase activity (GO:0008121)

p-value: 0.028; Adjusted p-value: 1.00

uridylyltransferase activity (GO:0070569)NO SIGNIFICANT ENRICHMENT

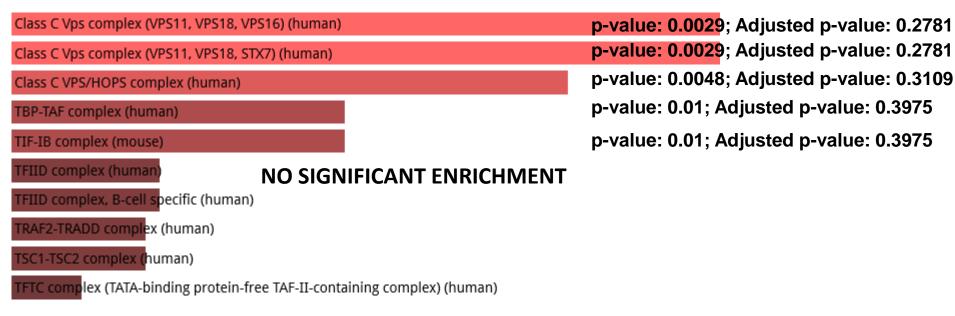
protein-glutamine gamma-glutamyltransferase activity (GO:0003810)

ubiquitin-ubiquitin ligase activity (GO:0034450)

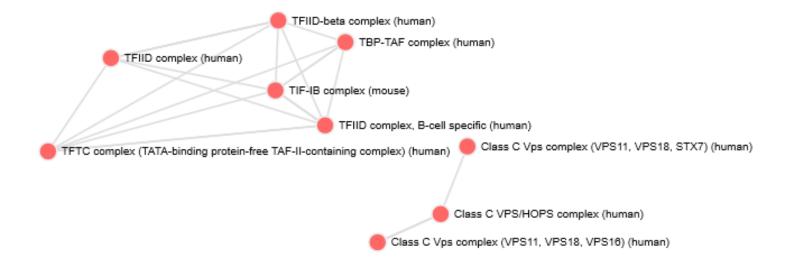
MHC protein binding (GO:0042287)

small protein activating enzyme activity (GO:0008641)

CORUM 2847 genes

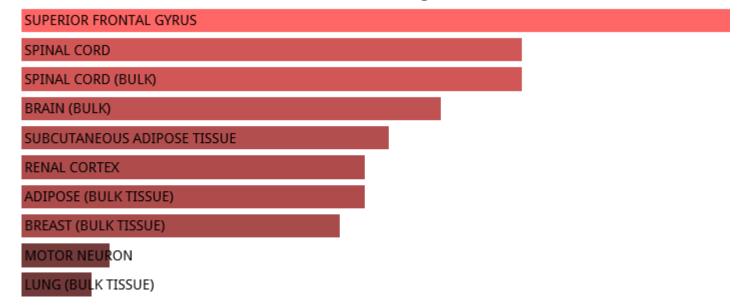


CORUM database of protein complexes identified by mass spectrometry



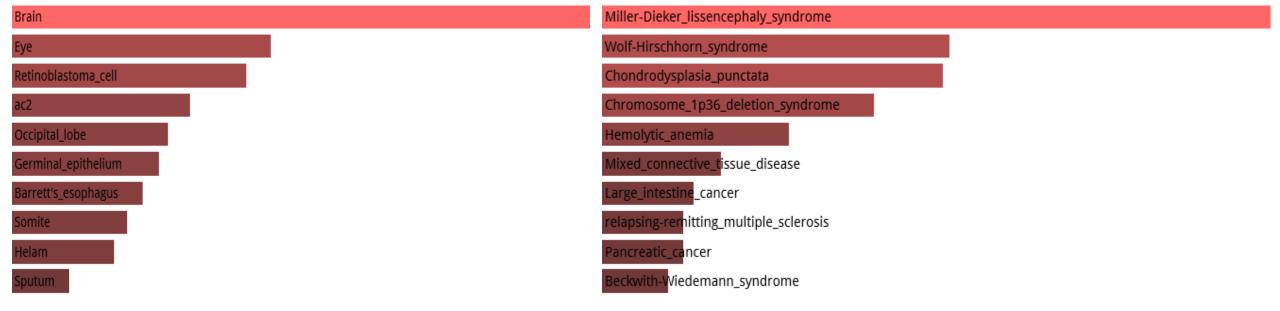
1,279 human-specific STR contractions

ARCHS4 Tissues 973 genes



Jensen TISSUES 973 genes

Jensen DISEASES 973 genes



2,118 fixed human-specific regulatory regions ARCHS4 Tissues 552 genes (non-hESC DHS) PREFRONTAL CORTEX SPINAL CORD SPINAL CORD (BULK) MOTOR NEURON CINGULATE GYRUS BRAIN (BULK) NEURONAL EPITHELIUM CEREBRAL CORTEX SUPERIOR FRONTAL GYRUS DORSAL STRIATUM 410 H3K4me3 sites with human-specific ARCHS4 Tissues 578 genes enrichment in prefrontal cortex neurons SUPERIOR FRONTAL GYRUS NEURONAL EPITHELIUM FETAL BRAIN CEREBRAL CORTEX MOTOR NEURON BRAIN (BULK) CEREBELLUM SPINAL CORD SPINAL CORD (BULK) **FIBROBLAST**