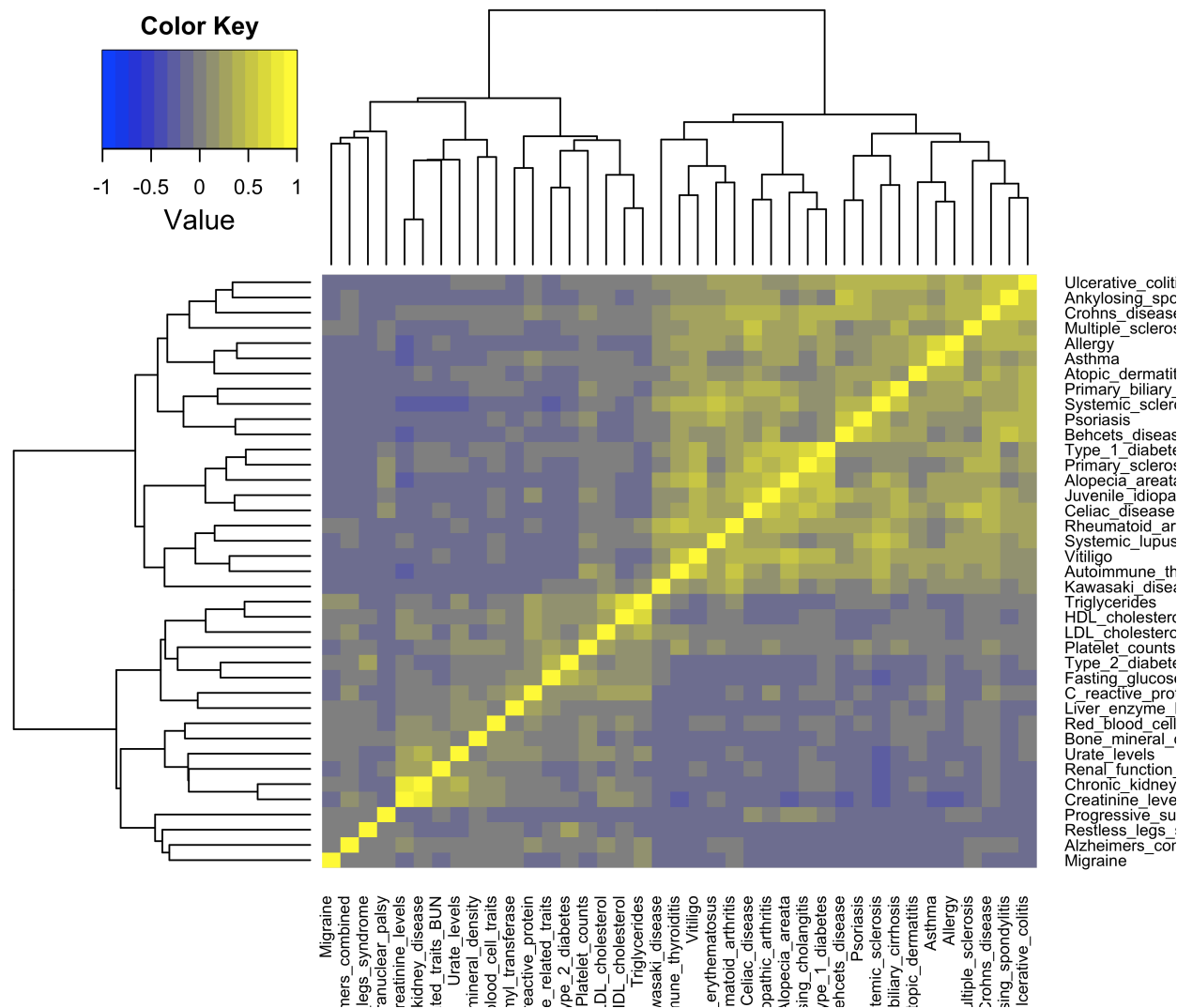


Supplemental Results. Analysis of autoimmune disease- and trait-associated SNP sets

Mikhail Dozmorov, Lukas R. Cara, Cory B. Giles, Bridget T. McInnes, Jonathan D. Wren

We analyzed 39 autoimmune disease- and trait-associated SNP sets, obtained from the Supplemental table 1 (<http://www.nature.com/nature/journal/vaop/ncurrent/extref/nature13835-s1.xls>) of the Farh, K. K.-H., Marson, A., Zhu, J., Kleinewietfeld, M., Housley, W. J., Beik, S., ... Bernstein, B. E. (2014). "Genetic and epigenetic fine mapping of causal autoimmune disease variants" (<http://www.nature.com/nature/journal/vaop/ncurrent/pdf/nature13835.pdf>) Nature. (doi:10.1038/nature13835)doi:10.1038/nature13835 (doi:10.1038/nature13835).

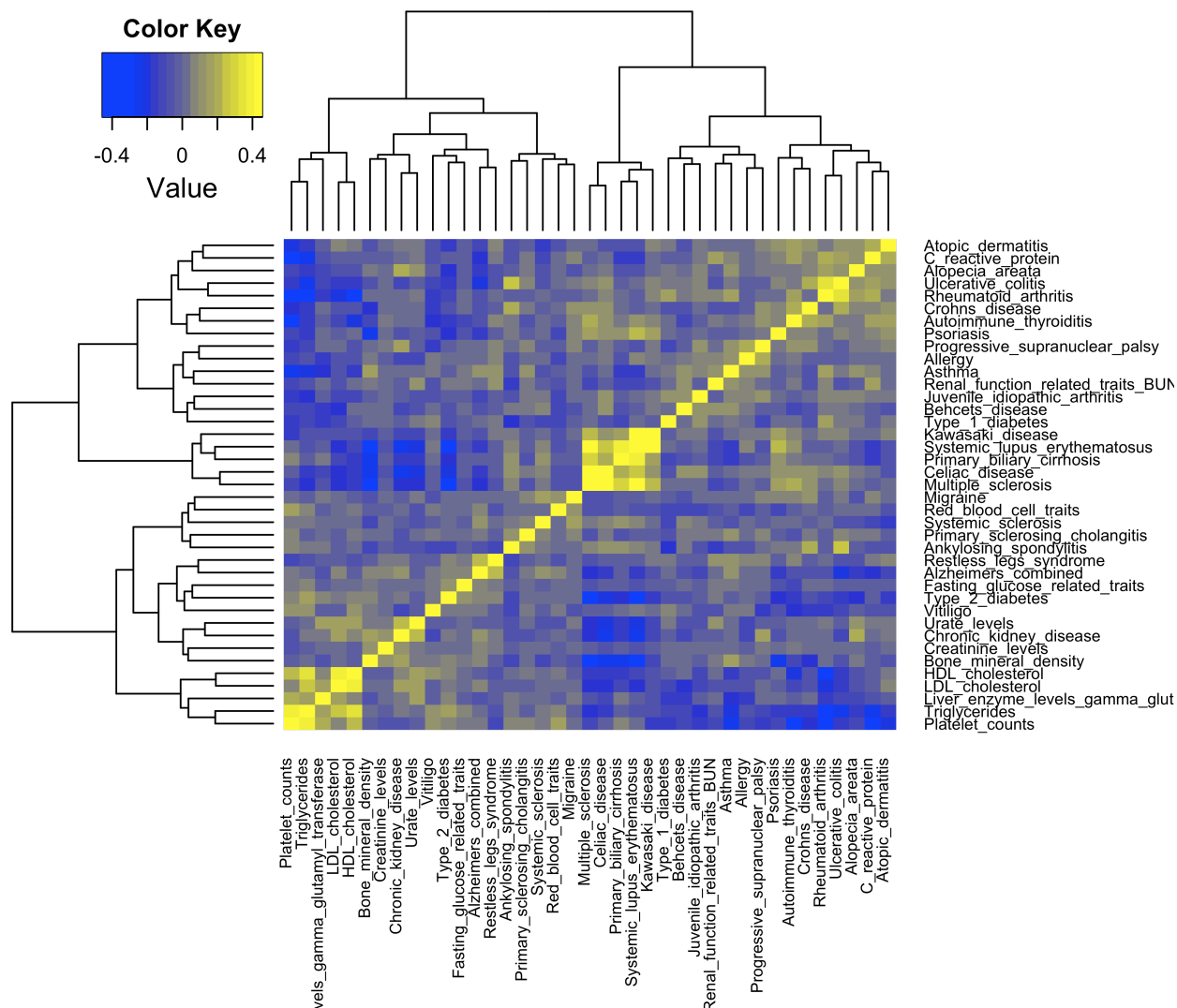
First, we re-created the heatmap of shared genetic features among the autoimmune diseases and traits, that is, counts of genomic elements overlapping between pairs of terms. We will use this heatmap as a reference point to compare with the heatmaps produced by the regulatory similarity analysis.



Analysis of all regulatory datasets

Although we used 4,498 regulatory datasets from the ENCODE project processed with the use with GenomeRunner, some regulatory datasets show not statistically significant enrichments in any of the 39 SNP sets. We removed these datasets as non-informative, and kept the remaining 2,969 regulatory datasets.

We visualized the matrix of pair-wise pearson correlation coefficients among the term-specific regulatory enrichment profiles.



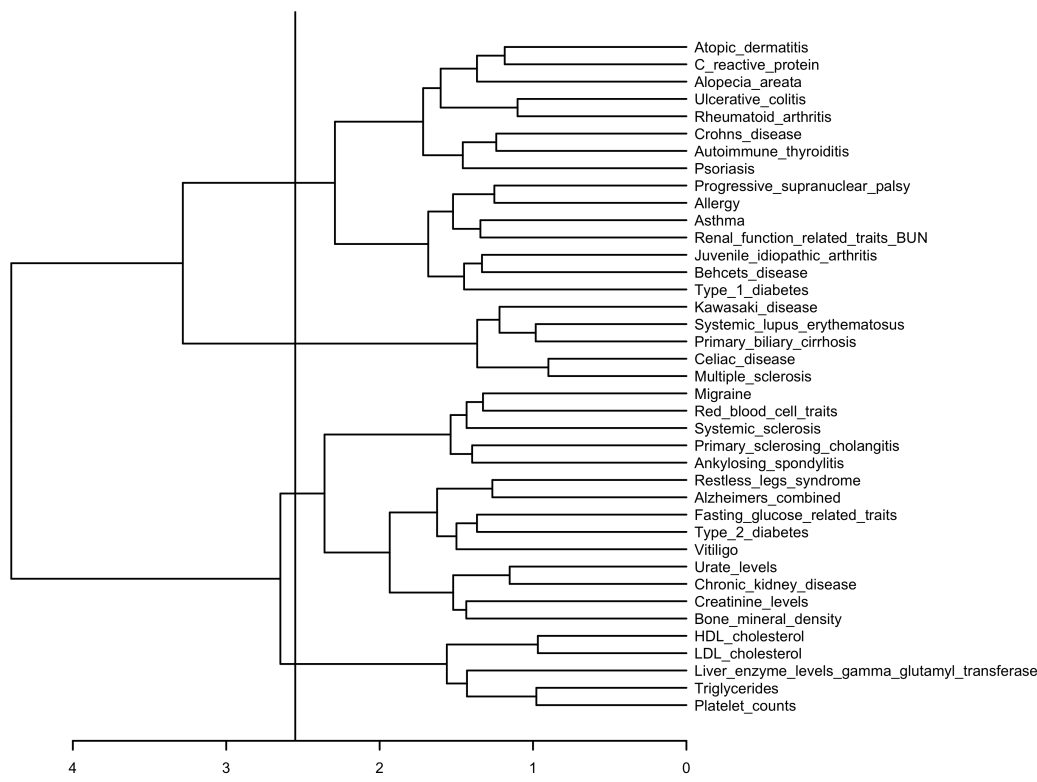
We then compared how regulatory similarity correlates with shared genomic features similarity. pearson correlation coefficient between the two is:

```
## [1] 0.6680046
```

The top 10 pairs of disease-associated SNPs are most similar with each other. The correlation coefficient shows pearson correlation coefficient among the regulatory enrichment profiles for each term-specific SNP set.

##			
##	-----		
##	Disease 1	Disease 2	Corr. coefficient
##	-----		
##	Multiple_sclerosis	Systemic_lupus_erythematosus	0.4557
##			
##	Celiac_disease	Multiple_sclerosis	0.4392
##			
##	Kawasaki_disease	Systemic_lupus_erythematosus	0.4383
##			
##	Primary_biliary_cirrhosis	Systemic_lupus_erythematosus	0.3928
##			
##	HDL_cholesterol	LDL_cholesterol	0.389
##			
##	Platelet_counts	Triglycerides	0.3847
##			
##	Rheumatoid_arthritis	Ulcerative_colitis	0.3566
##			
##	HDL_cholesterol	Triglycerides	0.3503
##			
##	Multiple_sclerosis	Primary_biliary_cirrhosis	0.3322
##			
##	LDL_cholesterol	Triglycerides	0.2817
##	-----		

The regulatory similarity dendrogram can be divided into four separate clusters:



```

## Cluster01 has 5 members
## Platelet_counts
## Triglycerides
## Liver_enzyme_levels_gamma_glutamyl_transferase
## LDL_cholesterol
## HDL_cholesterol
##
## Cluster02 has 14 members
## Bone_mineral_density
## Creatinine_levels
## Chronic_kidney_disease
## Urate_levels
## Vitiligo
## Type_2_diabetes
## Fasting_glucose_related_traits
## Alzheimers_combined
## Restless_legs_syndrome
## Ankylosing_spondylitis
## Primary_sclerosing_cholangitis
## Systemic_sclerosis
## Red_blood_cell_traits
## Migraine
##
## Cluster03 has 5 members
## Multiple_sclerosis
## Celiac_disease
## Primary_biliary_cirrhosis
## Systemic_lupus_erythematosus
## Kawasaki_disease
##
## Cluster04 has 15 members
## Type_1_diabetes
## Behcets_disease
## Juvenile_idiopathic_arthritis
## Renal_function_related_traits_BUN
## Asthma
## Allergy
## Progressive_supranuclear_palsy
## Psoriasis
## Autoimmune_thyroiditis
## Crohns_disease
## Rheumatoid_arthritis
## Ulcerative_colitis
## Alopecia_areata
## C_reactive_protein
## Atopic_dermatitis
##

```

We estimated the differences in regulatory associations of term-specific SNP sets.

The first column shows names of regulatory datasets. The following two columns show the average p-values of the cluster-specific SNP sets-regulatory associations. The smaller a p-value is, the more SNPs in a cluster enriched in corresponding regulatory dataset. A “-” sign indicates that an association is underrepresented (depleted). The “adj.P.Val” column shows whether a difference in the associations between the clusters is statistically significantly different. The last column shows descriptions of the regulatory datasets. The tables were sorted by “adj.P.Val” column; the top 10 or less most significantly different associations are shown.

[1] "c1 vs. c2 , number of degs significant at adj.p.val<0.5: 38"

##

--

##

Row.names

c1

c2

adj.P.Val

V2

--

##

wgEncodeSydhTfbsK562Pol2Ifng30StdPk

0.006469

0.9072

0.001309

K562 Pol2 Standard IFNg 30m

in

##

ChIP-seq Peaks from

##

ENCODE/SYDH

##

##

wgEncodeHaibTfbsHepg2P300V0416101PkRep2

0.00722

0.7239

0.001309

HepG2 p300 v041610.1 ChIP-s

eq

##

Peaks Rep 2 from ENCODE/HAI

B

##

##

wgEncodeOpenChromChipHepg2Pol2Pk

6.913e-07

-0.6861

0.002655

HepG2 Pol2 TFBS ChIP-seq

##

Peaks from

##

ENCODE/OpenChrom-UTA

##

##

wgEncodeHaibTfbsHepg2Pol24h8V0416102PkRep2

8.983e-07

-0.5865

0.007927

HepG2 Pol2-4H8 v041610.2

##

ChIP-seq Peaks Rep 2 from

##

ENCODE/HAIB

##

##

wgEncodeGisRnaPetNhekNucleusPapClustersRep1

8.839e-05

0.9334

0.008137

NHEK nucleus polyA+

##

clone-free RNA PET Cluster

s

##

Rep 1 from ENCODE/GIS

##

##

wgEncodeUwHistoneH7esH3k36me3StdDiffa5dHotspotsRep2

0.005028

0.8859

0.008137

H7-hESC H3K36me3 diffProtA

5

##

d Histone Mod ChIP-seq

##

Hotspots 2 from ENCODE/UW

##

##

wgEncodeBroadHistoneNhekH3k79me2Pk

1.5e-05

-0.5031

0.008137

NHEK H3K79me2 Histone Mods

by

##

ChIP-seq Peaks from

##

ENCODE/Broad

##

##

wgEncodeHaibTfbsHepg2Mybl2sc81192V0422111PkRep1

0.005722

0.9924

0.008363

HepG2 MYBL2 v042211.1


```

##
## -----
##
## [1] "c1 vs. c4 , number of degs significant at adj.p.val<0.5: 85"
##
## -----
----
##
## Row.names c1 c4 adj.P.Val V2
## -----
----
##
## wgEncodeHaibTfbsHepg2P300V0416101PkRep2 0.00722 0.9498 0.0007524 HepG2 p300 v041610.1 ChIP
-seq
##
## Peaks Rep 2 from ENCODE/H
AIB
##
## wgEncodeSydhTfbsK562Pol2Ifng30StdPk 0.006469 0.6275 0.001175 K562 Pol2 Standard IFNg 3
0min
##
## ChIP-seq Peaks from
##
## ENCODE/SYDH
##
## wgEncodeUwHistoneH7esh3k36me3StdDiffa5dHotspotsRep2 0.005028 -0.4649 0.001175 H7-hESC H3K36me3 diffProt
A 5
##
## d Histone Mod ChIP-seq
##
## Hotspots 2 from ENCODE/
UW
##
## wgEncodeGisRnaPetHepg2CytosolPapClustersRep1 6.141e-06 -0.01792 0.001175 HepG2 cytosol polyA+
##
## clone-based RNA PET Clust
ers
##
## Rep 1 from ENCODE/GIS
##
## wgEncodeGisRnaPetNhekNucleusPapClustersRep1 8.839e-05 -0.2856 0.001175 NHEK nucleus polyA+
##
## clone-free RNA PET Clust
ers
##
## Rep 1 from ENCODE/GIS
##
##
## wgEncodeOpenChromChipHepg2Pol2Pk 6.913e-07 -0.6948 0.001175 HepG2 Pol2 TFBS ChIP-se
q
##
## Peaks from
##
## ENCODE/OpenChrom-UTA
##
##
## wgEncodeBroadHistoneHuvecH3k79me2Pk 2.13e-05 -0.106 0.001175 HUVEC H3K79me2 Histone M
ods
##
## by ChIP-seq Peaks from
##
## ENCODE/Broad

```

```

##
##      wgEncodeBroadHistoneNhekH3k79me2Pk      1.5e-05   -0.1112   0.001309   NHEK H3K79me2 Histone Mod
s by
##
##
##
##
##
##      wgEncodeGisRnaPetHepg2NucleusPapClustersRep1      6.638e-06   -0.02357   0.002597   HepG2 nucleus polyA+
##
##
##      clone-based RNA PET Clust
ers
##
##      Rep 1 from ENCODE/GIS
##
##
##      wgEncodeBroadHistoneHuvecH3k27me3StdPk      -4.101e-05   0.1422   0.003049   HUVEC H3K27me3 Histone M
ods
##
##
##      by ChIP-seq Peaks from
##
##
##      ENCODE/Broad
##
## -----
##
##
## [1] "c2 vs. c3 , number of degs significant at adj.p.val<0.5: 159"
##
## -----
##      Row.names      c2      c3      adj.P.Val      V2
## -----
##      wgEncodeOpenChromFaireGm19239Pk      -0.3865   0.0001051   3.218e-10   GM19239 FAIRE Peaks from
##
##      ENCODE/OpenChrom(UNC)
##
##      wgEncodeOpenChromSynthGm18507Pk      -0.566   0.0004778   2.006e-08   GM18507 DNaseI/FAIRE/ChIP
##
##      Synthesis from
##
##      ENCODE/OpenChrom
##
##      wgEncodeHaibTfbsGm12878Nficsc81335V0422111PkRep1      -0.2202   2.709e-06   2.006e-08   GM12878 NFIC v042211.1
##
##      ChIP-seq Peaks Rep 1 from
##
##      ENCODE/HAIB
##
##      wgEncodeHaibTfbsGm12878Nficsc81335V0422111PkRep2      -0.517   0.002245   2.708e-08   GM12878 NFIC v042211.1
##
##      ChIP-seq Peaks Rep 2 from
##
##      ENCODE/HAIB
##
##      wgEncodeBroadHistoneGm12878H3k4me2StdPk      -0.02211   9.346e-10   8.134e-08   GM12878 H3K4me2 Histone Mods
##
##      by ChIP-seq Peaks from
##
##      ENCODE/Broad
##
##      wgEncodeUwDnaseGm12865HotspotsRep2      -0.2974   8.456e-06   1.038e-07   GM12865 DNaseI HS HotSpots
##
##      Rep 2 from ENCODE/UW
##
##      wgEncodeBroadHistoneDnd41H3k09acPk      -0.4638   2.678e-05   1.161e-07   Dnd41 H3K9ac Histone Mods by
##
##      ChIP-seq Peaks from
##
##      ENCODE/Broad
##
##      wgEncodeOpenChromFaireGm18507Pk      -0.4293   4.213e-05   1.561e-07   GM18507 FAIRE Peaks from
##
##      ENCODE/OpenChrom(UNC)

```



```

##
##      wgEncodeBroadHistoneDnd41H3k04me1Pk      -0.1887  3.207e-09  2.345e-07  Dnd41 H3K4me1 Histone Mods by
##                                                    CHIP-seq Peaks from
##                                                    ENCODE/Broad
##
##      wgEncodeSydhTfbsGm12878Pol2s2IggmusPk      -0.24   5.349e-06  3.092e-07  GM12878 Pol2 S2 IgG-mus
##                                                    CHIP-seq Peaks from
##                                                    ENCODE/SYDH
## -----
##
## [1] "c3 vs. c4 , number of degs significant at adj.p.val<0.5: 152"
##
## -----
##      Row.names      c3      c4      adj.P.Val      V2
## -----
##      wgEncodeOpenChromFaireGm19239Pk      0.0001051  0.6658  1.398e-08  GM19239 FAIRE Peaks from
##                                                    ENCODE/OpenChrom(UNC)
##
##      wgEncodeOpenChromSynthGm18507Pk      0.0004778  0.6495  6.122e-07  GM18507 DNaseI/FAIRE/CHIP
##                                                    Synthesis from
##                                                    ENCODE/OpenChrom
##
##      wgEncodeBroadHistoneDnd41H3k09acPk      2.678e-05 -0.8745  8.889e-07  Dnd41 H3K9ac Histone Mods by
##                                                    CHIP-seq Peaks from
##                                                    ENCODE/Broad
##
##      wgEncodeBroadHistoneDnd41H3k04me1Pk      3.207e-09 -0.7861  2.207e-06  Dnd41 H3K4me1 Histone Mods by
##                                                    CHIP-seq Peaks from
##                                                    ENCODE/Broad
##
##      wgEncodeHaibTfbsGm12878Nficsc81335V0422111PkRep1  2.709e-06  0.2551  2.391e-06  GM12878 NFIC v042211.1
##                                                    CHIP-seq Peaks Rep 1 from
##                                                    ENCODE/HAIB
##
##      wgEncodeOpenChromFaireGm18507Pk      4.213e-05  0.6812  2.798e-06  GM18507 FAIRE Peaks from
##                                                    ENCODE/OpenChrom(UNC)
##
##      wgEncodeUwDnaseGm12865HotspotsRep2      8.456e-06  0.5282  2.825e-06  GM12865 DNaseI HS HotSpots
##                                                    Rep 2 from ENCODE/UW
##
##      wgEncodeHaibTfbsGm12878Nficsc81335V0422111PkRep2  0.002245  0.4559  3.584e-06  GM12878 NFIC v042211.1
##                                                    CHIP-seq Peaks Rep 2 from
##                                                    ENCODE/HAIB
##
##      wgEncodeBroadHistoneGm12878H3k4me2StdPk      9.346e-10  0.3976  5.523e-06  GM12878 H3K4me2 Histone Mods
##                                                    by CHIP-seq Peaks from
##                                                    ENCODE/Broad
##
##      wgEncodeSydhTfbsGm12878Chd1a301218aIggmusPk      1.229e-06  0.7272  7.359e-06  GM12878 CHD1 IgG-mus CHIP-seq
##                                                    Peaks from ENCODE/SYDH
## -----

```

```

## [1] "Counts of regulatory elements differentially associated with each group"

```

```
##
## -----
##      &nbsp; c1    c2    c3    c4
## -----
##  **c1**    0    38   172   85
##
##  **c2**    0    0   159    0
##
##  **c3**    0    0    0   152
##
##  **c4**    0    0    0    0
## -----
```

Summary

The differences in regulatory associations are attributed to the clusters 1 and 3.

The cluster 1 (Platelet_counts, Triglycerides, Liver_enzyme_levels_gamma_glutamyl_transferase, LDL_cholesterol, HDL_cholesterol) showed enrichment in PolII, H3K36me3 and H3K79me2 histone modification marks. The signals was derived predominantly from cancer-related cell lines, such as HepG2 (liver hepatocellular carcinoma), K562 (myelogeneous leukemia), A549 (adenocarcinomic human alveolar basal epithelial cells).

The cluster 3 (Multiple_sclerosis, Celiac_disease, Primary_biliary_cirrhosis, Systemic_lupus_erythematosus, Kawasaki_disease) showed strong enrichment in B-cell derived signal. The most predominant cell type was Gm12878 (B lymphoblastoid cell line), CD20+ B cells, although some signal came from Dnd41 cells (T-ALL), T helper cells 1 and 2, and CD14+ monocytes. DNase hypersensitive sites were predominant type of signal enriched in the disease-associated SNPs from this cluster, followed by H3K4 mono-/di- and trimethylation, PolII, H3K79me2, H3K27 trimethylation and acetylation, and the CTCF, NFkB, NFATC1, STAT5A and other transcription factor binding sites

Co-morbidity similarity analysis

We used the data from Hidalgo CA, Blumm N, Barabasi A-L, Christakis NA. PLoS Computational Biology, 5(4):e1000353 (<http://www.ploscompbiol.org/article/info%3Adoi%2F10.1371%2Fjournal.pcbi.1000353>)doi:10.1371/journal.pcbi.1000353 (doi:10.1371/journal.pcbi.1000353), available at (<http://barabasilab.neu.edu/projects/hudine/resource/data/data.html>)<http://barabasilab.neu.edu/projects/hudine/resource/data/data.html> (<http://barabasilab.neu.edu/projects/hudine/resource/data/data.html>). These data provide co-morbidity measurements among pairs of diseases. We map autoimmune disease- and trait names to 3-digits ICD9 codes and evaluate how co-morbidity measurements correlate with regulatory similarity measurements. We used Phi measurement of co-morbidity. The pearson correlation coefficient of Phi and regulatory similarity is:

```
## [1] 0.4011363
```

Iridescent literature similarity

```
## [1] "sharedRels correlation with regulatory similarity"
## [1] 0.3657238
## [1] "obsExp correlation with regulatory similarity"
## [1] 0.5619774
## [1] "minMim correlation with regulatory similarity"
## [1] 0.7311888
## [1] "directStr correlation with regulatory similarity"
## [1] 0.1968273
## [1] "relOverlap correlation with regulatory similarity"
## [1] 0.3851007
## [1] "misn correlation with regulatory similarity"
## [1] 0.7310631
```

UMLS literature similarity

```
## [1] "msh.path correlation with regulatory similarity"
## [1] 0.3092355
## [1] "snomedct.path correlation with regulatory similarity"
## [1] 0.8287126
```

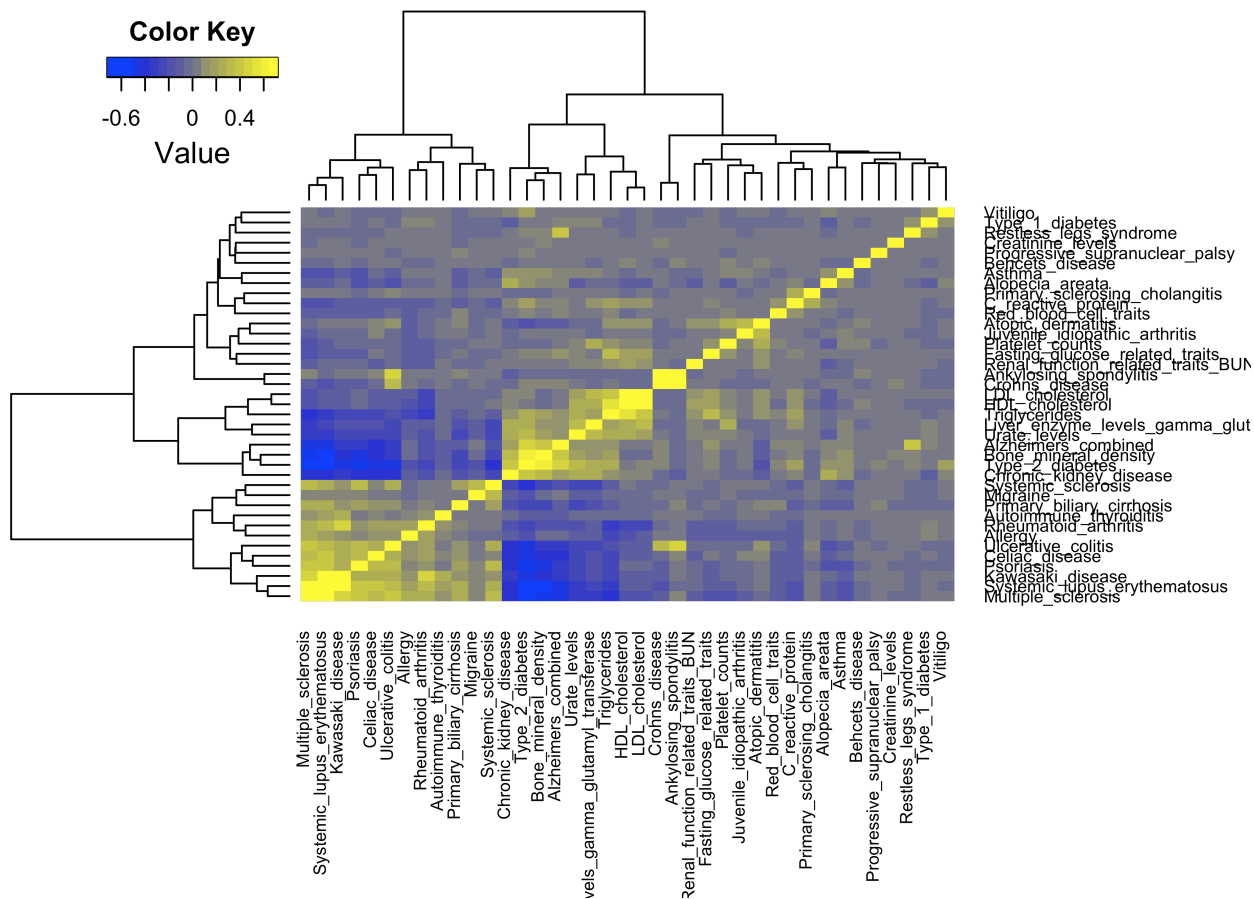
Analysis of TFBSs

We also performed regulatory similarity analysis using subsets of regulatory datasets, such as Transcription Factor Binding Sites or Histone Modification Marks. Here, out of all regulatory datasets, we selected only TFBSs.

```
## [1] 1954    39
```

```
## [1] 1259    39
```

We visualized the matrix of pair-wise pearson correlation coefficients among the term-specific regulatory enrichment profiles.



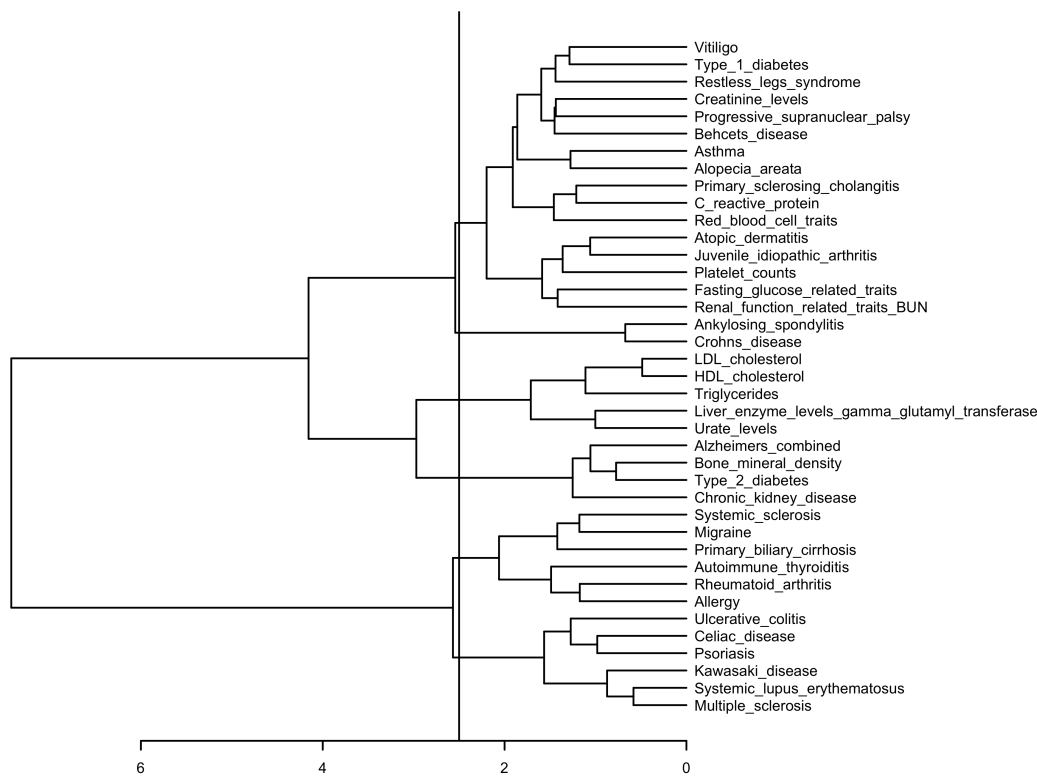
We then compared how regulatory similarity correlates with shared genomic features similarity. pearson correlation coefficient between the two is:

```
## [1] 0.6298022
```

The top 10 pairs of disease-associated SNPs are most similar with each other.

##			
##			
##	Disease 1	Disease 2	Corr. coefficient
##			
##	Kawasaki_disease	Systemic_lupus_erythematosus	0.7189
##			
##	HDL_cholesterol	LDL_cholesterol	0.6883
##			
##	Multiple_sclerosis	Systemic_lupus_erythematosus	0.6683
##			
##	Ankylosing_spondylitis	Crohns_disease	0.6548
##			
##	Bone_mineral_density	Type_2_diabetes	0.584
##			
##	HDL_cholesterol	Triglycerides	0.5278
##			
##	Ankylosing_spondylitis	Ulcerative_colitis	0.5021
##			
##	Kawasaki_disease	Rheumatoid_arthritis	0.4912
##			
##	Alzheimers_combined	Type_2_diabetes	0.4884
##			
##	Kawasaki_disease	Multiple_sclerosis	0.4872
##			

The regulatory similarity dendrogram can be divided into four separate clusters:



```

## Cluster01 has    6 members
## Multiple_sclerosis
## Systemic_lupus_erythematosus
## Kawasaki_disease
## Psoriasis
## Celiac_disease
## Ulcerative_colitis
##
## Cluster02 has    6 members
## Allergy
## Rheumatoid_arthritis
## Autoimmune_thyroiditis
## Primary_biliary_cirrhosis
## Migraine
## Systemic_sclerosis
##
## Cluster03 has    4 members
## Chronic_kidney_disease
## Type_2_diabetes
## Bone_mineral_density
## Alzheimers_combined
##
## Cluster04 has    5 members
## Urate_levels
## Liver_enzyme_levels_gamma_glutamyl_transferase
## Triglycerides
## HDL_cholesterol
## LDL_cholesterol
##
## Cluster05 has    2 members
## Crohns_disease
## Ankylosing_spondylitis
##
## Cluster06 has   16 members
## Renal_function_related_traits_BUN
## Fasting_glucose_related_traits
## Platelet_counts
## Juvenile_idiopathic_arthritis
## Atopic_dermatitis
## Red_blood_cell_traits
## C_reactive_protein
## Primary_sclerosing_cholangitis
## Alopecia_areata
## Asthma
## Behcets_disease
## Progressive_supranuclear_palsy
## Creatinine_levels
## Restless_legs_syndrome
## Type_1_diabetes
## Vitiligo
##

```

We estimated the differences in regulatory associations of term-specific SNP sets.

The first column shows names of regulatory datasets. The following two columns show the average p-values of the cluster-specific SNP sets-regulatory associations. The smaller a p-value is, the more SNPs in a cluster enriched in corresponding regulatory dataset. A “-” sign indicates that an association is underrepresented (depleted). The “adj.P.Val” column shows whether a difference in the

associations between the clusters is statistically significantly different. The last column shows descriptions of the regulatory datasets. The tables were sorted by “adj.P.Val” column; the top 10 or less most significantly different associations are shown.

## [1] "c1 vs. c2 , number of degs significant at adj.p.val<0.5: 62"					
##					
## -----					
##	Row.names	c1	c2	adj.P.Val	V2
## -----					
##	wgEncodeHaibTfbsGm12878Runx3sc101553V0422111PkRep2	8.939e-05	0.561	3.281e-09	GM12878 RUNX3 v042211.1
##					ChIP-seq Peaks Rep 2 from
##					ENCODE/HAIB
##					
##	wgEncodeHaibTfbsGm12878Pmlsc71910V0422111PkRep2	1.027e-05	0.507	3.952e-09	GM12878 PML v042211.1
##					ChIP-seq Peaks Rep 2 from
##					ENCODE/HAIB
##					
##	wgEncodeHaibTfbsGm12878Stat5asc74442V0422111PkRep2	5.806e-05	0.8614	5.249e-09	GM12878 STAT5A v042211.1
##					ChIP-seq Peaks Rep 2 from
##					ENCODE/HAIB
##					
##	wgEncodeHaibTfbsGm12878Pou2f2Pcr1xPkRep2	0.003845	0.9554	1.375e-07	GM12878 POU2F2 PCR1x ChIP-seq
##					Peaks Rep 2 from ENCODE/HAIB
##					
##	wgEncodeHaibTfbsGm12891Pol2Pcr1xPkRep1	0.004814	0.6935	2.882e-07	GM12891 Pol2 PCR1x ChIP-seq
##					Peaks Rep 1 from ENCODE/HAIB
##					
##	wgEncodeHaibTfbsGm12878Stat5asc74442V0422111PkRep1	0.000687	0.906	2.89e-07	GM12878 STAT5A v042211.1
##					ChIP-seq Peaks Rep 1 from
##					ENCODE/HAIB
##					
##	wgEncodeHaibTfbsGm12878Bclaf101388V0416101PkRep2	0.001796	0.4804	5.46e-07	GM12878 BCLAF1 v041610.1
##					ChIP-seq Peaks Rep 2 from
##					ENCODE/HAIB
##					
##	wgEncodeHaibTfbsGm12878Pol2Pcr2xPkRep1	0.0001325	0.6133	5.46e-07	GM12878 Pol2 PCR2x ChIP-seq
##					Peaks Rep 1 from ENCODE/HAIB
##					
##	wgEncodeHaibTfbsGm12878Mta3sc81325V0422111PkRep2	7.449e-07	0.4086	5.46e-07	GM12878 MTA3 v042211.1
##					ChIP-seq Peaks Rep 2 from
##					ENCODE/HAIB
##					
##	wgEncodeHaibTfbsGm12878Atf2sc81188V0422111PkRep1	0.0001343	0.5215	4.325e-06	GM12878 ATF2 v042211.1
##					ChIP-seq Peaks Rep 1 from
##					ENCODE/HAIB
## -----					
##					
## [1] "c1 vs. c3 , number of degs significant at adj.p.val<0.5: 65"					
##					
## -----					

##	Row.names	c1	c3	adj.P.Val	V2
## -----					

##	wgEncodeHaibTfbsGm12878Pmlsc71910V0422111PkRep2	1.027e-05	-0.01199	3.087e-12	GM12878 PML v042211.1
##					ChIP-seq Peaks Rep 2 fr

om						
##					ENCODE/HAIB	
##						
##	wgEncodeHaibTfbsGm12878Runx3sc101553V0422111PkRep2	8.939e-05	-0.07637	1.033e-11	GM12878 RUNX3 v042211.	
1						
##					ChIP-seq Peaks Rep 2 fr	
om						
##					ENCODE/HAIB	
##						
##	wgEncodeHaibTfbsGm12878Stat5asc74442V0422111PkRep2	5.806e-05	-0.02658	1.402e-11	GM12878 STAT5A v042211.	
1						
##					ChIP-seq Peaks Rep 2 fr	
om						
##					ENCODE/HAIB	
##						
##	wgEncodeHaibTfbsGm12878Pol2Pcr2xPkRep1	0.0001325	-0.01003	1.366e-10	GM12878 Pol2 PCR2x ChIP-	
seq						
##					Peaks Rep 1 from ENCODE/H	
AIB						
##						
##	wgEncodeHaibTfbsGm12878Mta3sc81325V0422111PkRep2	7.449e-07	-0.0008688	1.366e-10	GM12878 MTA3 v042211.1	
##					ChIP-seq Peaks Rep 2 fr	
om						
##					ENCODE/HAIB	
##						
##	wgEncodeHaibTfbsGm12878Pou2f2Pcr1xPkRep2	0.003845	-0.09044	1.484e-10	GM12878 POU2F2 PCR1x ChIP	
-seq						
##					Peaks Rep 2 from ENCODE/H	
AIB						
##						
##	wgEncodeHaibTfbsGm12891Pol2Pcr1xPkRep1	0.004814	-0.2148	1.16e-09	GM12891 Pol2 PCR1x ChIP-	
seq						
##					Peaks Rep 1 from ENCODE/H	
AIB						
##						
##	wgEncodeHaibTfbsGm12878Bclaf101388V0416101PkRep2	0.001796	-0.1952	1.91e-09	GM12878 BCLAF1 v041610.	
1						
##					ChIP-seq Peaks Rep 2 fr	
om						
##					ENCODE/HAIB	
##						
##	wgEncodeHaibTfbsGm12878Atf2sc81188V0422111PkRep2	0.0007129	-0.05776	1.951e-09	GM12878 ATF2 v042211.1	
##					ChIP-seq Peaks Rep 2 fr	
om						
##					ENCODE/HAIB	
##						
##	wgEncodeHaibTfbsGm12878Runx3sc101553V0422111PkRep1	0.0001092	-0.01052	4.089e-09	GM12878 RUNX3 v042211.	
1						
##					ChIP-seq Peaks Rep 1 fr	


```

om
##
##
## -----
##
## [1] "c1 vs. c4 , number of degs significant at adj.p.val<0.5: 77"
##
## -----
--
##          Row.names          c1      c4      adj.P.Val          V2
## -----
##
## wgEncodeHaibTfbsGm12878Runx3sc101553V0422111PkRep2  8.939e-05 -0.1192   1.486e-11   GM12878 RUNX3 v042211.1
##
##                                     CHIP-seq Peaks Rep 2 from
##                                     ENCODE/HAIB
##
## wgEncodeHaibTfbsGm12878Pmlsc71910V0422111PkRep2    1.027e-05 -0.1599   5.185e-11   GM12878 PML v042211.1
##
##                                     CHIP-seq Peaks Rep 2 from
##                                     ENCODE/HAIB
##
## wgEncodeHaibTfbsGm12878Stat5asc74442V0422111PkRep2  5.806e-05 -0.2564   3.98e-10   GM12878 STAT5A v042211.1
##
##                                     CHIP-seq Peaks Rep 2 from
##                                     ENCODE/HAIB
##
## wgEncodeHaibTfbsGm12878Nfisc81335V0422111PkRep1    8.776e-06 -0.0344   1.617e-08   GM12878 NFIC v042211.1
##
##                                     CHIP-seq Peaks Rep 1 from
##                                     ENCODE/HAIB
##
## wgEncodeHaibTfbsGm12878Foxmlsc502V0422111PkRep1    0.0002931 -0.09224   1.617e-08   GM12878 FOXM1 v042211.1
##
##                                     CHIP-seq Peaks Rep 1 from
##                                     ENCODE/HAIB
##
## wgEncodeHaibTfbsGm12878Stat5asc74442V0422111PkRep1  0.000687  -0.3028   1.617e-08   GM12878 STAT5A v042211.1
##
##                                     CHIP-seq Peaks Rep 1 from
##                                     ENCODE/HAIB
##
## wgEncodeHaibTfbsGm12878Bclaf101388V0416101PkRep2  0.001796  -0.5678   1.617e-08   GM12878 BCLAF1 v041610.1

```

```

##
ChIP-seq Peaks Rep 2 from

##
ENCODE/HAIB

##
## wgEncodeHaibTfbsGm12878Atf2sc81188V0422111PkRep2 0.0007129 -0.2449 1.898e-08 GM12878 ATF2 v042211.1

##
ChIP-seq Peaks Rep 2 from

##
ENCODE/HAIB

##
## wgEncodeHaibTfbsGm12891Pol2Pcr1xPkRep1 0.004814 -0.6803 1.898e-08 GM12891 Pol2 PCR1x ChIP-seq
q
## Peaks Rep 1 from ENCODE/HAIB
B
##
## wgEncodeHaibTfbsGm12878Pou2f2Pcr1xPkRep2 0.003845 -0.6217 2.253e-08 GM12878 POU2F2 PCR1x ChIP-seq
eq
## Peaks Rep 2 from ENCODE/HAIB
B
## -----
--
##
## [1] "c1 vs. c6 , number of degs significant at adj.p.val<0.5: 66"
##
## -----
-
## Row.names c1 c6 adj.P.Val V2
## -----
-
## wgEncodeHaibTfbsGm12878Runx3sc101553V0422111PkRep2 8.939e-05 -0.9353 2.59e-12 GM12878 RUNX3 v042211.1

##
ChIP-seq Peaks Rep 2 from

##
ENCODE/HAIB

##
## wgEncodeHaibTfbsGm12878Pmlsc71910V0422111PkRep2 1.027e-05 -0.9168 3.592e-12 GM12878 PML v042211.1

##
ChIP-seq Peaks Rep 2 from

##
ENCODE/HAIB

##
## wgEncodeHaibTfbsGm12878Stat5asc74442V0422111PkRep2 5.806e-05 0.7588 5.061e-11 GM12878 STAT5A v042211.1

##
ChIP-seq Peaks Rep 2 from

##
ENCODE/HAIB

##
## wgEncodeHaibTfbsGm12878Bclaf101388V0416101PkRep2 0.001796 -0.9275 4.125e-10 GM12878 BCLAF1 v041610.1

##
ChIP-seq Peaks Rep 2 from

```


[illegible]

```

## wgEncodeHaibTfbsHepg2Foxal1sc6553V0416101PkRep1 -0.9877 0.006962 0.0199 HepG2 FOXA1 6553 v041610.1
## ChIP-seq Peaks Rep 1 from
## ENCODE/HAIB
##
## wgEncodeHaibTfbsHepg2Pol24h8V0416102PkRep2 -0.1786 1.017e-06 0.02195 HepG2 Pol2-4H8 v041610.2
## ChIP-seq Peaks Rep 2 from
## ENCODE/HAIB
##
## wgEncodeHaibTfbsHepg2Tead4sc101184V0422111PkRep1 0.3458 0.001973 0.02195 HepG2 TEAD4 v042211.1
## ChIP-seq Peaks Rep 1 from
## ENCODE/HAIB
##
## wgEncodeHaibTfbsHepg2Pol2Pcr2xPkRep2 -0.5218 0.001153 0.02202 HepG2 Pol2 PCR2x ChIP-seq
## Peaks Rep 2 from ENCODE/HAIB
##
## wgEncodeSydhTfbsHepg2Pol2IggrabPk -0.4535 0.000279 0.02202 HEPG2 Pol2 IgG-rab ChIP-seq
## Peaks from ENCODE/SYDH
##
## wgEncodeHaibTfbsHepg2Pol24h8V0416102PkRep1 -0.6623 0.0004534 0.02298 HepG2 Pol2-4H8 v041610.2
## ChIP-seq Peaks Rep 1 from
## ENCODE/HAIB
##
## wgEncodeHaibTfbsHepg2Mbd4sc271530V0422111PkRep1 -0.986 0.009568 0.07367 HepG2 MBD4 v042211.1 ChIP-seq
## Peaks Rep 1 from ENCODE/HAIB
##
## wgEncodeHaibTfbsHepg2Hey1V0416101PkRep2 -0.6034 0.003643 0.007673
## -----
##
## [1] "c3 vs. c6 , number of degs significant at adj.p.val<0.5: 1"
##
## -----
## Row.names c3 c6 adj.P.Val V2
## -----
## wgEncodeHaibTfbsGm12878Mta3sc81325V0422111PkRep2 -0.0008688 -0.6721 0.04644 GM12878 MTA3 v042211.1
## ChIP-seq Peaks Rep 2 from
## ENCODE/HAIB
## -----
## [1] "c4 vs. c6 , number of degs significant at adj.p.val<0.5: 11"
##
## -----
## Row.names c4 c6 adj.P.Val V2
## -----
## wgEncodeHaibTfbsHepg2Mybl2sc81192V0422111PkRep2 0.0007234 0.9488 5.268e-05 HepG2 MYBL2 v042211.1
## ChIP-seq Peaks Rep 2 from
## ENCODE/HAIB
##
## wgEncodeHaibTfbsHepg2Tead4sc101184V0422111PkRep1 0.001973 0.9985 0.0001274 HepG2 TEAD4 v042211.1
## ChIP-seq Peaks Rep 1 from
## ENCODE/HAIB
##
## wgEncodeHaibTfbsHepg2Mybl2sc81192V0422111PkRep1 0.003229 0.9916 0.0002865 HepG2 MYBL2 v042211.1
## ChIP-seq Peaks Rep 1 from
## ENCODE/HAIB
##
## wgEncodeHaibTfbsHepg2Foxal1sc6553V0416101PkRep1 0.006962 0.8901 0.001191 HepG2 FOXA1 6553 v041610.1
## ChIP-seq Peaks Rep 1 from

```

##						ENCODE/HAIB
##						
##	wgEncodeHaibTfbsHepg2Pol24h8V0416102PkRep1	0.0004534	0.6691	0.003514		HepG2 Pol2-4H8 v041610.2
##						ChIP-seq Peaks Rep 1 from
##						ENCODE/HAIB
##						
##	wgEncodeHaibTfbsHepg2Pol2Pcr2xPkRep2	0.001153	0.6707	0.005071		HepG2 Pol2 PCR2x ChIP-seq
##						Peaks Rep 2 from ENCODE/HAIB
##						
##	wgEncodeHaibTfbsHepg2Pol24h8V0416102PkRep2	1.017e-06	0.4296	0.005513		HepG2 Pol2-4H8 v041610.2
##						ChIP-seq Peaks Rep 2 from
##						ENCODE/HAIB
##						
##	wgEncodeHaibTfbsHepg2Mbd4sc271530V0422111PkRep1	0.009568	0.9875	0.005573		HepG2 MBD4 v042211.1 ChIP-seq
##						Peaks Rep 1 from ENCODE/HAIB
##						
##	wgEncodeSydhTfbsHepg2Pol2IggrabPk	0.000279	0.5467	0.005573		HEPG2 Pol2 IgG-rab ChIP-seq
##						Peaks from ENCODE/SYDH
##						
##	wgEncodeHaibTfbsHepg2Hey1V0416101PkRep1	0.000285	0.7581	0.001305		
##	-----					

[1] "Counts of regulatory elements differentially associated with each group"

##						
##	-----					
##	 	c1	c2	c3	c4	c5
##						
##	**c1**	0	62	65	77	0
##						
##	**c2**	0	0	3	11	0
##						
##	**c3**	0	0	0	11	0
##						
##	**c4**	0	0	0	0	11
##						
##	**c5**	0	0	0	0	0
##						
##	**c6**	0	0	0	0	0
##	-----					

Summary

Now, the same set of diseases is different from the others.

Co-morbidity similarity analysis

We used the data from Hidalgo CA, Blumm N, Barabasi A-L, Christakis NA. PLoS Computational Biology, 5(4):e1000353 (<http://www.ploscompbiol.org/article/info%3Adoi%2F10.1371%2Fjournal.pcbi.1000353>)doi:10.1371/journal.pcbi.1000353 (doi:10.1371/journal.pcbi.1000353), available at (<http://barabasilab.neu.edu/projects/hudine/resource/data/data.html>)<http://barabasilab.neu.edu/projects/hudine/resource/data/data.html> (<http://barabasilab.neu.edu/projects/hudine/resource/data/data.html>). These data provide co-morbidity measurements among pairs of diseases. We map autoimmune disease- and trait names to 3-digits ICD9 codes and evaluate how co-morbidity measurements correlate with regulatory similarity measurements. We used Phi measurement of co-morbidity. The pearson correlation coefficient of Phi and regulatory similarity is:

```
## [1] 0.3759245
```

Iridescent literature similarity

```
## [1] "sharedRels correlation with regulatory similarity"
## [1] 0.3484246
## [1] "obsExp correlation with regulatory similarity"
## [1] 0.5045555
## [1] "minMim correlation with regulatory similarity"
## [1] 0.6470513
## [1] "directStr correlation with regulatory similarity"
## [1] 0.2835723
## [1] "relOverlap correlation with regulatory similarity"
## [1] 0.3602434
## [1] "mism correlation with regulatory similarity"
## [1] 0.672326
```

UMLS literature similarity

```
## [1] "msh.path correlation with regulatory similarity"
## [1] 0.2614632
## [1] "snomedct.path correlation with regulatory similarity"
## [1] 0.7064502
```

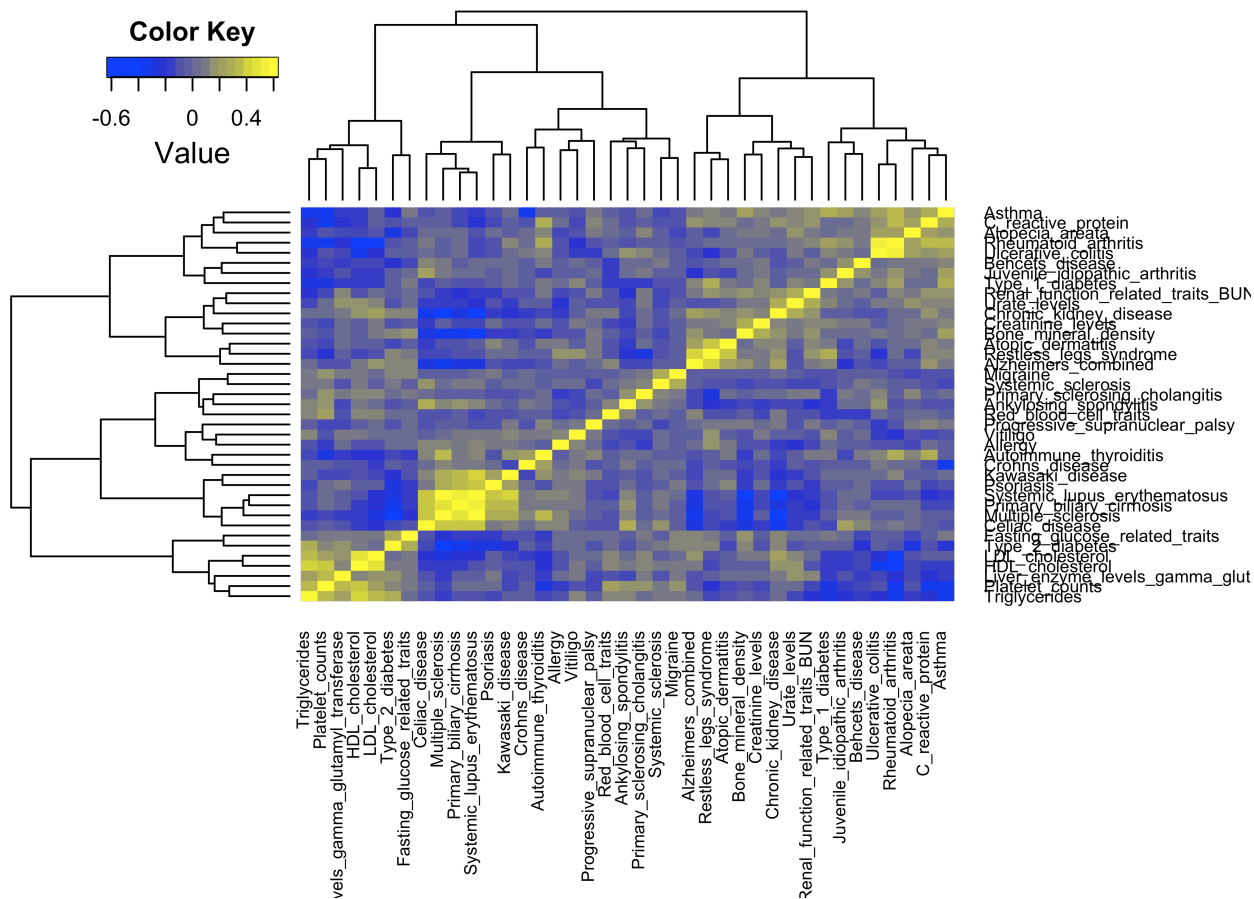
Analysis of histone marks

We also performed regulatory similarity analysis using subsets of regulatory datasets, such as Transcription Factor Binding Sites or Histone Modification Marks. Here, out of all regulatory datasets, we selected only Histone marks.

```
## [1] 721 39
```

```
## [1] 610 39
```

We visualized the matrix of pair-wise pearson correlation coefficients among the term-specific regulatory enrichment profiles.



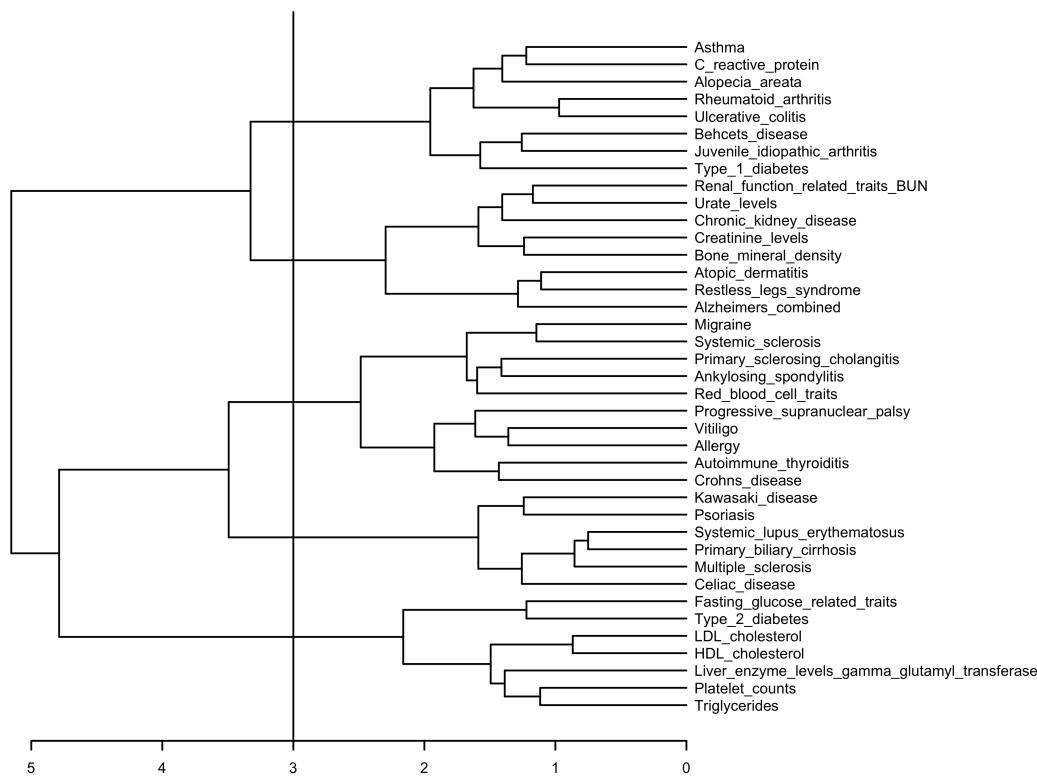
We then compared how regulatory similarity correlates with shared genomic features similarity. pearson correlation coefficient between the two is:

```
## [1] 0.6299304
```

The top 10 pairs of disease-associated SNPs are most similar with each other.

##	Disease 1	Disease 2	Corr. coefficient
##	Multiple_sclerosis	Systemic_lupus_erythematosus	0.6312
##	Primary_biliary_cirrhosis	Systemic_lupus_erythematosus	0.5528
##	Rheumatoid_arthritis	Ulcerative_colitis	0.5467
##	Multiple_sclerosis	Primary_biliary_cirrhosis	0.4959
##	HDL_cholesterol	LDL_cholesterol	0.4798
##	Kawasaki_disease	Systemic_lupus_erythematosus	0.4668
##	Alzheimers_combined	Restless_legs_syndrome	0.4515
##	HDL_cholesterol	Triglycerides	0.4354
##	Psoriasis	Systemic_lupus_erythematosus	0.407
##	Celiac_disease	Multiple_sclerosis	0.4017

The regulatory similarity dendrogram can be divided into four separate clusters:



```

## Cluster01 has    7 members
## Triglycerides
## Platelet_counts
## Liver_enzyme_levels_gamma_glutamyl_transferase
## HDL_cholesterol
## LDL_cholesterol
## Type_2_diabetes
## Fasting_glucose_related_traits
##
## Cluster02 has    6 members
## Celiac_disease
## Multiple_sclerosis
## Primary_biliary_cirrhosis
## Systemic_lupus_erythematosus
## Psoriasis
## Kawasaki_disease
##
## Cluster03 has   10 members
## Crohns_disease
## Autoimmune_thyroiditis
## Allergy
## Vitiligo
## Progressive_supranuclear_palsy
## Red_blood_cell_traits
## Ankylosing_spondylitis
## Primary_sclerosing_cholangitis
## Systemic_sclerosis
## Migraine
##
## Cluster04 has    8 members
## Alzheimers_combined
## Restless_legs_syndrome
## Atopic_dermatitis
## Bone_mineral_density
## Creatinine_levels
## Chronic_kidney_disease
## Urate_levels
## Renal_function_related_traits_BUN
##
## Cluster05 has    8 members
## Type_1_diabetes
## Juvenile_idiopathic_arthritis
## Behcets_disease
## Ulcerative_colitis
## Rheumatoid_arthritis
## Alopecia_areata
## C_reactive_protein
## Asthma
##

```

We estimated the differences in regulatory associations of term-specific SNP sets.

The first column shows names of regulatory datasets. The following two columns show the average p-values of the cluster-specific SNP sets-regulatory associations. The smaller a p-value is, the more SNPs in a cluster enriched in corresponding regulatory dataset. A “-” sign indicates that an association is underrepresented (depleted). The “adj.P.Val” column shows whether a difference in the associations between the clusters is statistically significantly different. The last column shows descriptions of the regulatory datasets. The tables were sorted by “adj.P.Val” column; the top 10 or less most significantly different associations are shown.

[illegible]

```

##
##   wgEncodeBroadHistoneGm12878H3k04me3StdPkV2      -0.02222  4.708e-08  3.425e-05  GM12878 H3K4me3 Histone Mods
##
##                                                    by ChIP-seq Peaks from
##
##                                                    ENCODE/Broad
##
##
##   wgEncodeBroadHistoneGm12878H3k04me1StdPkV2      -5.857e-05 6.263e-15  3.615e-05  GM12878 H3K4me1 Histone Mods
##
##                                                    by ChIP-seq Peaks from
##
##                                                    ENCODE/Broad
##
##
##   wgEncodeUwHistoneGm12864H3k04me3StdHotspotsRep2  -0.523   0.003202   3.615e-05  GM12864 H3K4me3 Histone Mod
##
##                                                    ChIP-seq Hotspots 2 from
##
##                                                    ENCODE/UW
## -----
-
##
## [1] "c1 vs. c3 , number of degs significant at adj.p.val<0.5: 6"
##
## -----
##           Row.names                c1      c3      adj.P.Val      V2
## -----
##   wgEncodeBroadHistoneA549H3k79me2Dex100nmPk  0.0001024  -0.9021   0.02348  A549 DEX 100 nM H3K79me2
##                                                    Histone Mods by ChIP-seq
##                                                    Peaks from ENCODE/Broad
##
##   wgEncodeBroadHistoneA549H3k79me2Etoh02Pk    0.0002004   0.8357   0.08259  A549 EtOH 0.02% H3K79me2
##                                                    Histone Mods by ChIP-seq
##                                                    Peaks from ENCODE/Broad
##
##   wgEncodeBroadHistoneNhlfH3k79me2Pk          4.751e-05   0.6981   0.08259  NHLF H3K79me2 Histone Mods by
##                                                    ChIP-seq Peaks from
##                                                    ENCODE/Broad
##
##   wgEncodeBroadHistoneHuvecH3k79me2Pk         0.0001697   0.7884   0.08259  HUVEC H3K79me2 Histone Mods
##                                                    by ChIP-seq Peaks from
##                                                    ENCODE/Broad
##
##   wgEncodeBroadHistoneNhekH3k79me2Pk          0.0001432   0.7962   0.08259  NHEK H3K79me2 Histone Mods by
##                                                    ChIP-seq Peaks from
##                                                    ENCODE/Broad
##
##   wgEncodeBroadHistoneHuvecH3k27me3StdPk      -0.0002469  -0.8449   0.08259  HUVEC H3K27me3 Histone Mods
##                                                    by ChIP-seq Peaks from
##                                                    ENCODE/Broad
## -----
##
## [1] "c1 vs. c4 , number of degs significant at adj.p.val<0.5: 34"
##
## -----

```

##	Row.names	c1	c4	adj.P.Val	V2
##	-----	-----	-----	-----	-----
##	wgEncodeBroadHistoneA549H3k79me2Dex100nmPk	0.0001024	-0.1158	0.002307	A549 DEX 100 nM H3K79me2 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
##					
##					
##					
##	wgEncodeBroadHistoneOsteoH4k20me1Pk	0.007273	-0.09496	0.00362	Osteoblasts H4K20me1 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
##					
##					
##					
##	wgEncodeBroadHistoneNhaH4k20me1Pk	0.002923	-0.1011	0.004216	NH-A H4K20me1 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
##					
##					
##					
##	wgEncodeBroadHistoneNhekH3k79me2Pk	0.0001432	-0.07455	0.006329	NHEK H3K79me2 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
##					
##					
##					
##	wgEncodeBroadHistoneA549H3k79me2Etoh02Pk	0.0002004	-0.2685	0.009682	A549 EtOH 0.02% H3K79me2 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
##					
##					
##					
##	wgEncodeBroadHistoneNhdfadH3k36me3StdPk	0.04191	-0.005215	0.0131	NHDF-Ad H3K36me3 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
##					
##					
##					
##	wgEncodeBroadHistoneHuvecH3k79me2Pk	0.0001697	-0.3014	0.0131	HUVEC H3K79me2 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
##					
##					
##					
##	wgEncodeBroadHistoneNhekH3k36me3StdPk	0.002936	-0.02459	0.01416	NHEK H3K36me3 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
##					
##					
##					
##	wgEncodeBroadHistoneHmecH3k36me3StdPk	0.001914	-0.0002847	0.01711	HMEC H3K36me3 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
##					
##					
##					
##	wgEncodeBroadHistoneNhlfH3k79me2Pk	4.751e-05	-0.6816	0.02063	NHLF H3K79me2 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
##					
##					
##	-----	-----	-----	-----	-----
##					
##	[1] "c1 vs. c5 , number of degs significant at adj.p.val<0.5: 57"				
##					
##	-----	-----	-----	-----	-----
##	Row.names	c1	c5	adj.P.Val	V2
##	-----	-----	-----	-----	-----
##	wgEncodeBroadHistoneA549H3k79me2Dex100nmPk	0.0001024	-0.007415	4.128e-05	A549 DEX 100 nM H3K79me2 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
##					
##					
##					
##	wgEncodeBroadHistoneA549H3k79me2Etoh02Pk	0.0002004	-0.0104	0.0003534	A549 EtOH 0.02% H3K79me2 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
##					
##					
##					
##	wgEncodeBroadHistoneNhaH3k79me2Pk	0.0009854	-0.002468	0.0004093	NH-A H3K79me2 Histone Mods by

##					ChIP-seq Peaks from
##					ENCODE/Broad
##	wgEncodeBroadHistoneHuvecH3k79me2Pk	0.0001697	-0.0123	0.0004093	HUVEC H3K79me2 Histone Mods
##					by ChIP-seq Peaks from
##					ENCODE/Broad
##	wgEncodeBroadHistoneOsteoH3k79me2Pk	0.002415	-0.004605	0.0004484	Osteoblasts H3K79me2 Histone
##					Mods by ChIP-seq Peaks from
##					ENCODE/Broad
##	wgEncodeBroadHistoneNhlfH3k79me2Pk	4.751e-05	-0.01706	0.0006228	NHLF H3K79me2 Histone Mods by
##					ChIP-seq Peaks from
##					ENCODE/Broad
##	wgEncodeBroadHistoneHuvecH3k27me3StdPk	-0.0002469	0.01744	0.0006936	HUVEC H3K27me3 Histone Mods
##					by ChIP-seq Peaks from
##					ENCODE/Broad
##	wgEncodeBroadHistoneHsmmH3k79me2StdPk	0.0006338	-0.0148	0.0006936	HSMM H3K79me2 Histone Mods by
##					ChIP-seq Peaks from
##					ENCODE/Broad
##	wgEncodeBroadHistoneNhekH3k79me2Pk	0.0001432	-0.02136	0.0006936	NHEK H3K79me2 Histone Mods by
##					ChIP-seq Peaks from
##					ENCODE/Broad
##	wgEncodeBroadHistoneHsmmH3k27me3StdPk	-0.001048	1.409e-09	0.0008928	HSMM H3K27me3 Histone Mods by
##					ChIP-seq Peaks from
##					ENCODE/Broad
##	-----				
##	[1] "c2 vs. c3 , number of degs significant at adj.p.val<0.5: 37"				
##	-----				
##	Row.names	c2	c3	adj.P.Val	V2
##	-----				
##	wgEncodeUwHistoneGm12875H3k04me3StdHotspotsRep1	0.0005212	-0.8239	9.755e-07	GM12875 H3K4me3 Histone Mod
##					ChIP-seq Hotspots 1 from
##					ENCODE/UW
##	wgEncodeBroadHistoneGm12878H3k9acStdPk	3.849e-12	0.5541	1.714e-05	GM12878 H3K9ac Histone Mods
##					by ChIP-seq Peaks from
##					ENCODE/Broad
##	wgEncodeUwHistoneGm12865H3k04me3StdHotspotsRep2	5.452e-07	-0.8528	2.26e-05	GM12865 H3K4me3 Histone Mod
##					ChIP-seq Hotspots 2 from
##					ENCODE/UW
##	wgEncodeBroadHistoneGm12878H3k4me2StdPk	8.782e-09	-0.6556	2.26e-05	GM12878 H3K4me2 Histone Mods
##					by ChIP-seq Peaks from
##					ENCODE/Broad
##	wgEncodeUwHistoneGm12865H3k04me3StdHotspotsRep1	2.881e-06	-0.9392	0.0001006	GM12865 H3K4me3 Histone Mod
##					ChIP-seq Hotspots 1 from
##					ENCODE/UW
##					

```

## wgEncodeUwHistoneGm12864H3k04me3StdHotspotsRep2 0.003202 -0.9901 0.0001157 GM12864 H3K4me3 Histone Mod
##
##                               CHIP-seq Hotspots 2 from
##                               ENCODE/UW
##
## wgEncodeBroadHistoneGm12878H3k79me2StdPk 1.867e-08 -0.8139 0.0001487 GM12878 H3K79me2 Histone Mods
##
##                               by CHIP-seq Peaks from
##                               ENCODE/Broad
##
## wgEncodeBroadHistoneDnd41H3k09acPk 0.0001527 0.7596 0.0001965 Dnd41 H3K9ac Histone Mods by
##
##                               CHIP-seq Peaks from
##                               ENCODE/Broad
##
## wgEncodeBroadHistoneDnd41H3k04me1Pk 9.105e-08 0.7344 0.0003011 Dnd41 H3K4me1 Histone Mods by
##
##                               CHIP-seq Peaks from
##                               ENCODE/Broad
##
## wgEncodeBroadHistoneGm12878H3k04me3StdPkV2 4.708e-08 0.7527 0.0003011 GM12878 H3K4me3 Histone Mods
##
##                               by CHIP-seq Peaks from
##                               ENCODE/Broad
## -----
##
## [1] "c2 vs. c4 , number of degs significant at adj.p.val<0.5: 56"
##
## -----
-
##                               Row.names                               c2                               c4                               adj.P.Val                               V2
## -----
-
## wgEncodeUwHistoneGm12875H3k04me3StdHotspotsRep1 0.0005212 -0.3919 3.243e-07 GM12875 H3K4me3 Histone Mod
##
##                               CHIP-seq Hotspots 1 from
##                               ENCODE/UW
##
## wgEncodeBroadHistoneGm12878H3k4me2StdPk 8.782e-09 -0.02811 3.964e-06 GM12878 H3K4me2 Histone Mods
##
##                               by CHIP-seq Peaks from
##                               ENCODE/Broad
##
## wgEncodeBroadHistoneGm12878H3k9acStdPk 3.849e-12 -0.1383 3.964e-06 GM12878 H3K9ac Histone Mods
##
##                               by CHIP-seq Peaks from
##                               ENCODE/Broad
##
## wgEncodeBroadHistoneGm12878H3k79me2StdPk 1.867e-08 -0.005656 7.731e-06 GM12878 H3K79me2 Histone Mod
s
##                               by CHIP-seq Peaks from
##                               ENCODE/Broad
##

```

```

## wgEncodeUwHistoneGm12865H3k04me3StdHotspotsRep2 5.452e-07 -0.2947 1.082e-05 GM12865 H3K4me3 Histone Mod
##
## ChIP-seq Hotspots 2 from
##
## ENCODE/UW
##
##
## wgEncodeBroadHistoneGm12878H3k04me3StdPkV2 4.708e-08 -0.01621 2.189e-05 GM12878 H3K4me3 Histone Mods
##
## by ChIP-seq Peaks from
##
## ENCODE/Broad
##
##
## wgEncodeBroadHistoneGm12878H3k9me3StdPk 2.129e-10 -1.413e-05 2.824e-05 GM12878 H3K9me3 Histone Mods
##
## by ChIP-seq Peaks from
##
## ENCODE/Broad
##
##
## wgEncodeUwHistoneGm12864H3k04me3StdHotspotsRep2 0.003202 -0.5261 2.824e-05 GM12864 H3K4me3 Histone Mod
##
## ChIP-seq Hotspots 2 from
##
## ENCODE/UW
##
##
## wgEncodeUwHistoneGm12865H3k04me3StdHotspotsRep1 2.881e-06 -0.3212 2.824e-05 GM12865 H3K4me3 Histone Mod
##
## ChIP-seq Hotspots 1 from
##
## ENCODE/UW
##
##
## wgEncodeBroadHistoneDnd41H3k09acPk 0.0001527 -0.3194 2.824e-05 Dnd41 H3K9ac Histone Mods by
##
## ChIP-seq Peaks from
##
## ENCODE/Broad
##
## -----
##
##
## [1] "c2 vs. c5 , number of degs significant at adj.p.val<0.5: 64"
##
## -----
##
## Row.names c2 c5 adj.P.Val V2
## -----
## wgEncodeUwHistoneGm12875H3k04me3StdHotspotsRep1 0.0005212 0.7538 1.095e-05 GM12875 H3K4me3 Histone Mod
## ChIP-seq Hotspots 1 from
## ENCODE/UW
##
##
## wgEncodeBroadHistoneGm12878H3k9acStdPk 3.849e-12 0.1729 0.0001041 GM12878 H3K9ac Histone Mods
## by ChIP-seq Peaks from
## ENCODE/Broad
##
##

```



```
## wgEncodeUwHistoneGm12865H3k04me3StdHotspotsRep2 5.452e-07 0.5657 0.0001705 GM12865 H3K4me3 Histone Mod
##
##
##
##
## wgEncodeBroadHistoneGm12878H3k4me2StdPk 8.782e-09 0.4824 0.0001705 GM12878 H3K4me2 Histone Mods
##
##
##
##
## wgEncodeBroadHistoneDnd41H3k09acPk 0.0001527 -0.7471 0.0002681 Dnd41 H3K9ac Histone Mods by
##
##
##
##
## wgEncodeBroadHistoneGm12878H3k79me2StdPk 1.867e-08 -0.4384 0.0002681 GM12878 H3K79me2 Histone Mods
##
##
##
##
## wgEncodeUwHistoneGm12865H3k04me3StdHotspotsRep1 2.881e-06 0.7265 0.0002681 GM12865 H3K4me3 Histone Mod
##
##
##
##
## wgEncodeUwHistoneGm12864H3k04me3StdHotspotsRep2 0.003202 0.8722 0.0002936 GM12864 H3K4me3 Histone Mod
##
##
##
##
## wgEncodeBroadHistoneDnd41H3k04me1Pk 9.105e-08 -0.5579 0.0003657 Dnd41 H3K4me1 Histone Mods by
##
##
##
##
## wgEncodeBroadHistoneCd20H3k04me2Pk 6.1e-09 -0.3652 0.0005895 CD20+ H3K4me2 Histone Mods by
##
##
##
##
```

[1] "c3 vs. c5 , number of degs significant at adj.p.val<0.5: 3"

```
##
## -----
##
## Row.names c3 c5 adj.P.Val V2
## -----
##
## wgEncodeBroadHistoneHsmmH3k27me3StdPk -0.6298 1.409e-09 0.06695 HSMM H3K27me3 Histone Mods by
##
##
##
##
## wgEncodeBroadHistoneDnd41Ezh239875Pk -0.253 0.0006476 0.06695 Dnd41 EZH2 (39875) Histone
##
##
##
##
##
##
## wgEncodeBroadHistoneA549H3k36me3Dex100nmPk 0.08468 -0.0008403 0.08249 A549 DEX 100 nM H3K36me3
##
##
##
##
##
##
```

[1] "c4 vs. c5 , number of degs significant at adj.p.val<0.5: 2"

```
##
## -----
##
## Row.names c4 c5 adj.P.Val V2
## -----
##
## wgEncodeBroadHistoneHsmmH3k27me3StdPk -0.02171 1.409e-09 0.0213 HSMM H3K27me3 Histone Mods by
```

```
##
##
##
## wgEncodeBroadHistoneNhekEzh239875Pk -0.1117 0.009518 0.0213 NHEK EZH2 (39875) Histone
##
## Mods by ChIP-seq Peaks from
##
## ENCODE/Broad
## -----
```

```
## [1] "Counts of regulatory elements differentially associated with each group"
```

```
##
## -----
## &nbsp; c1 c2 c3 c4 c5
## -----
## **c1** 0 57 6 34 57
##
## **c2** 0 0 37 56 64
##
## **c3** 0 0 0 0 3
##
## **c4** 0 0 0 0 2
##
## **c5** 0 0 0 0 0
## -----
```

Summary

Co-morbidity similarity analysis

We used the data from Hidalgo CA, Blumm N, Barabasi A-L, Christakis NA. PLoS Computational Biology, 5(4):e1000353 (<http://www.ploscompbiol.org/article/info%3Adoi%2F10.1371%2Fjournal.pcbi.1000353>)doi:10.1371/journal.pcbi.1000353 (doi:10.1371/journal.pcbi.1000353), available at (<http://barabasilab.neu.edu/projects/hudine/resource/data/data.html>)<http://barabasilab.neu.edu/projects/hudine/resource/data/data.html> (<http://barabasilab.neu.edu/projects/hudine/resource/data/data.html>). These data provide co-morbidity measurements among pairs of diseases. We map autoimmune disease- and trait names to 3-digits ICD9 codes and evaluate how co-morbidity measurements correlate with regulatory similarity measurements. We used Phi measurement of co-morbidity. The pearson correlation coefficient of Phi and regulatory similarity is:

```
## [1] 0.3799314
```

Iridescent literature similarity

```
## [1] "sharedRels correlation with regulatory similarity"
## [1] 0.3395089
## [1] "obsExp correlation with regulatory similarity"
## [1] 0.5154583
## [1] "minMim correlation with regulatory similarity"
## [1] 0.6535677
## [1] "directStr correlation with regulatory similarity"
## [1] 0.1748984
## [1] "relOverlap correlation with regulatory similarity"
## [1] 0.3594159
## [1] "misn correlation with regulatory similarity"
## [1] 0.6776696
```

UMLS literature similarity

```
## [1] "msh.path correlation with regulatory similarity"
## [1] 0.2743516
## [1] "snomedct.path correlation with regulatory similarity"
## [1] 0.7275952
```

Misc

Distribution of maxMin correlation coefficients

