Genome Biology in 2017

Mehmet Tevfik DORAK, MD PhD

2nd Practical Bioinformatics Course
Istanbul, 17/18 April 2017

YOUR FUTURE
STARTS WITH HOPE

Liverpool Hope University EST. 1844

Schedule

Day 1
Genome Biology in 2017
Bioinformatics Tools in Epigenetics
Pathogenicity Assessment of DNA Sequence Mutations
Introduction to Galaxy
Massive Data Sources
Practical

Gene Expression Regulation

- **Chromatin modifications**
- **Transcriptional regulation (TF-mediated)**
- **Post-transcriptional (ncRNA-mediated)**
- **Translational (RNA decay; ribosome occupancy)**

eQTL: Expression quantitative trait loci (variants associated with expression levels)

Types of eQTLs:

dsQTL	: DNase I sensitivity quantitative trait loci
enhSNP	: enhancer SNP
eQTN	: expression quantitative trait nucleotide (causative eQTL)
esQTL	: expression-specific QTLs (greater correlation with mRNA than protein abundance)
hQTL	: histone-modification quantitative trait loci
haQTL	: histone acetylation level quantitative trait loci
meQTL	: methylation level quantitative trait loci
miR-eQTL	: miRNA expression level quantitative trait loci
rQTL	: ribosome occupancy quantitative trait loci
rbSNP	: RNA-binding protein binding site SNP
pQTL	: protein abundance quantitative trait loci
psQTL	: protein-specific QTLs (greater correlation with protein than mRNA abundance)
RdQTL	: RNA decay quantitative trait loci
sQTL	: Splicing quantitative trait loci
QTL^{epi}	: Epigenetic quantitative trait loci (correlations with DMR)

Gene Expression Regulation

- Chromatin looping
- Transcription factor binding site alteration
- DNase I hypersensitivity (DHS) region modification
- Histone modifications
- DNA methylation
- ncRNA sequence or binding site alteration
- RNA splicing

Ultimately >>> **eQTL effect**

Gene Expression Regulation

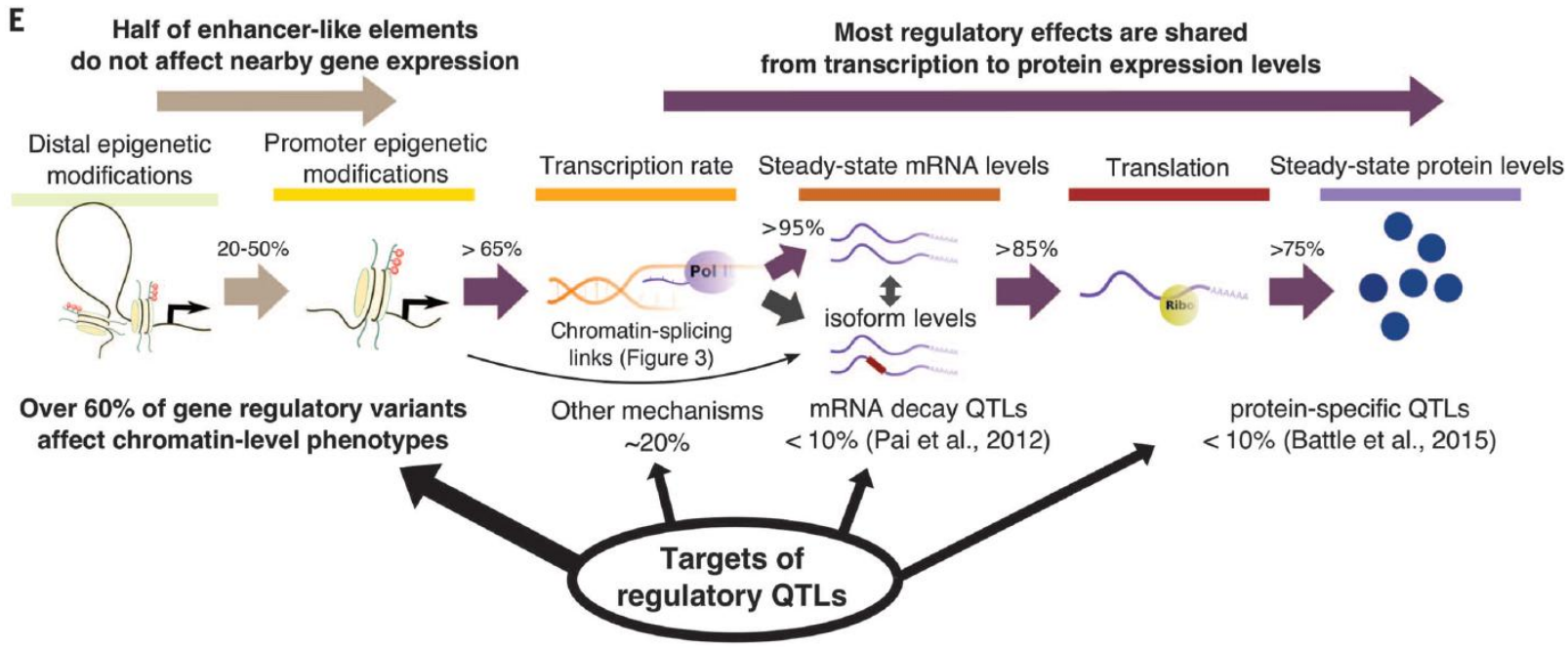


Fig. 2. Percolation of genetic effects through the gene regulatory cascade. (**A**) Correlation of effect sizes across different measurements from eQTLs identified in the GEUVADIS YRI sample (*6*). Txn rate, transcription rate. (**B**) QTL sharing across the regulatory cascade. Each panel shows the estimated fraction of QTLs identified at one stage that are preserved at the next stage of regulation. The four bars in each panel correspond to the *P*-value threshold for ascertaining QTLs in each assay, using the linear regression *t* statistics. Bars represent 80% confidence intervals on π_1, the fraction of true positives (*16*). The enhancer→TSS panel considers the effect of H3K27ac QTLs on the nearest TSS. (**C**) The fraction of expression QTLs that also affect chromatin-level phenotypes, as estimated by two models, and for matched control SNPs. About 35% of gene eQTLs do not appear to affect chromatin traits. QTLs for H3K4me1 and H3K4me3 are from (*8*). (**D**) Functional context of eQTL SNPs that are not associated with chromatin changes ("unexplained") versus those eQTLs that are also chromatin QTLs. 5′ untranslated regions were excluded from the "gene exons" annotation. Five annotations with bootstrap $P > 0.05$ are not shown. (**E**) Summary of the effects of regulatory QTLs and of their sharing through the regulatory cascade.

RNA splicing is a primary link between genetic variation and disease

Yang I. Li,[1] Bryce van de Geijn,[2] Anil Raj,[1] David A. Knowles,[3,4] Allegra A. Petti,[5] David Golan,[1] Yoav Gilad,[2]* Jonathan K. Pritchard[1,6,7]*

Gene Expression Regulation

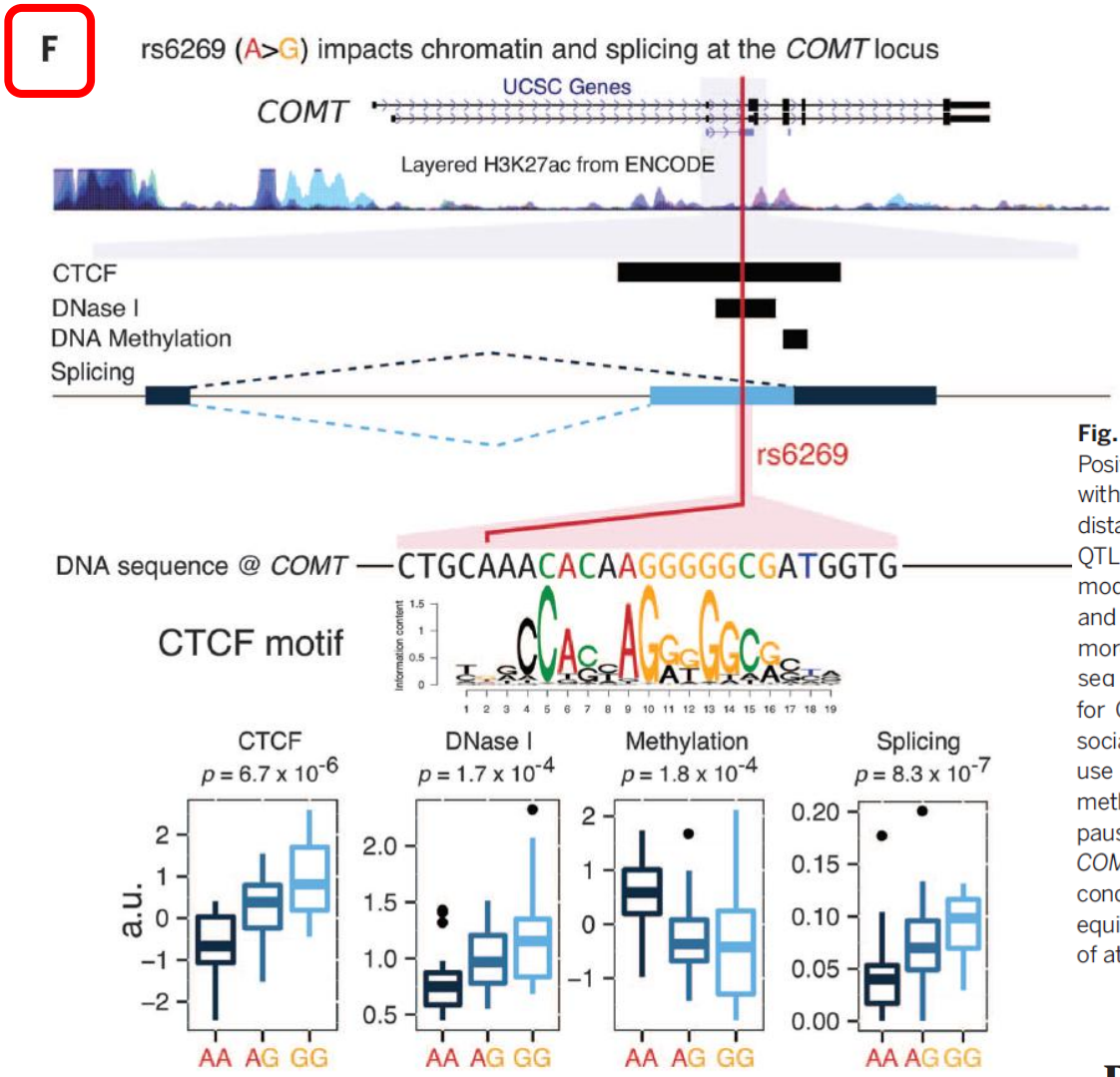


Fig. 3. Properties of sQTLs. Most sQTLs act independently from eQTLs: Positional distributions of (**A**) eQTLs and (**B**) sQTLs at 5% FDR are consistent with our mechanistic understanding of gene transcription and splicing. (**C**) The distance between the best eQTL and best sQTL for genes with both types of QTL is typically large, suggesting distinct causal variants. (**D**) A hierarchical model reveals distinct genomic features that are most relevant for eQTLs and sQTLs, respectively. (**E**) QTLs for CTCF binding, and H3K27ac levels are more likely to be sQTLs than matched SNPs within CTCF and H3K27ac ChIP-seq peaks, respectively. (**F**) Example of an sQTL (rs6269) that is also a QTL for CTCF, DNaseI sensitivity, and DNA methylation. The allele that is associated with increased CTCF occupancy is also associated with increased use of an alternative upstream splice site for an exon of the catechol-*O*-methyltransferase gene, *COMT*, which is consistent with the model that PolII pausing at CTCF binding sites can promote upstream exon inclusion (*21*). *COMT*, which regulates dopamine, has possible roles in neuropsychiatric conditions (*25*). In Europeans, the sQTL is in nearly complete linkage disequilibrium with a missense variant, rs4680, which has been the main focus of attention to date.

RNA splicing is a primary link between genetic variation and disease

Yang I. Li,[1] Bryce van de Geijn,[2] Anil Raj,[1] David A. Knowles,[3,4] Allegra A. Petti,[5] David Golan,[1] Yoav Gilad,[2]* Jonathan K. Pritchard[1,6,7]*

Gene Expression Regulation

	A	B	C	D	E	F	G	H	I
1	cluster	pvalue	intron_coord	snp_pos	eQTL	ctcfQTL	acQTL	dsQTL	meQTL
2	clu5430	2.21E-04	chr19:1011611:1011740	rs10405583	2.19E-03	1.00E+00	1.00E+00	2.49E-03	3.21E-03
3	clu1743	6.99E-08	chr11:68665480:68668003	rs544370	2.98E-21	1.00E+00	2.06E-07	1.00E+00	1.36E-06
4	clu4984	3.83E-05	chr14:24572464:24573051	rs144856253	7.36E-01	1.00E+00	1.23E-03	4.13E-03	1.00E+00
5	clu677	4.29E-05	chr12:57867959:57868235	rs4760273	4.96E-01	1.00E+00	1.00E+00	2.13E-03	1.43E-03
6	clu11546	3.45E-07	chr2:113955211:113955327	rs148543122	1.98E-01	1.00E+00	1.95E-03	1.00E+00	2.35E-03
7	clu2617	5.28E-07	chr10:127473829:127483449	rs10901449	1.54E-02	1.00E+00	6.02E-04	1.00E+00	8.11E-04
8	clu9550	2.04E-04	chr5:169188629:169191038	rs7721990	3.62E-03	1.00E+00	3.83E-04	2.16E-03	1.00E+00
9	clu5714	2.17E-05	chr19:10675710:10676582	rs3745244	1.26E-01	1.00E+00	7.21E-04	4.20E-03	1.00E+00
10	clu7346	4.36E-04	chr20:44424082:44429713	rs6032531	1.35E-01	2.18E-12	6.47E-04	1.00E+00	1.00E+00
11	clu14580	4.84E-04	chr8:145065075:145065368	rs7462703	7.96E-01	1.00E+00	1.00E+00	2.71E-03	4.10E-03
12	clu13108	4.84E-04	chr1:155238150:155238499	rs1057941	4.59E-01	1.00E+00	2.42E-06	3.02E-03	1.00E+00
13	clu3566	3.29E-24	chr17:74553939:74557370	rs4647885	1.60E-03	1.00E+00	2.31E-03	1.00E+00	4.91E-03
14	clu4373	8.10E-20	chr16:89165171:89167070	rs8060043	1.64E-07	1.00E+00	1.64E-04	1.00E+00	1.99E-04
15	clu7010	4.03E-15	chr22:50316091:50316260	rs6520064	3.17E-01	1.77E-05	1.00E-03	1.00E+00	1.00E+00
16	clu7713	2.74E-06	chr7:6049129:6057445	rs34097030	2.26E-02	1.00E+00	1.00E+00	6.80E-04	6.49E-04
17	clu5946	9.91E-06	chr19:19743075:19744612	rs873870	8.67E-01	1.00E+00	1.00E+00	1.81E-03	3.97E-03
18	clu6361	1.66E-16	chr19:54704756:54705028	rs3852889	5.76E-01	1.00E+00	8.71E-04	1.00E+00	2.61E-03
19	clu5321	1.29E-16	chr14:92583986:92587540	rs2896197	8.47E-03	1.00E+00	1.26E-04	1.00E+00	4.90E-03
20	clu7661	3.24E-05	chr7:643253:647502	rs111526379	1.75E-02	1.00E+00	1.75E-07	1.00E+00	1.58E-03
21	clu2223	5.34E-04	chr10:51592619:51606988	rs2012677	4.39E-01	1.00E+00	3.14E-06	3.43E-03	1.00E+00
22	clu12882	5.70E-08	chr1:111739875:111740328	rs599134	1.89E-04	4.96E-09	7.55E-10	2.18E-14	1.00E+00
23	clu5215	1.09E-04	chr14:71064494:71067333	rs12589444	6.93E-01	1.00E+00	6.32E-04	1.66E-03	1.00E+00
24	clu6966	3.02E-04	chr22:42995799:42998776	rs6002854	1.17E-01	1.00E+00	9.26E-05	4.23E-03	1.00E+00

RNA splicing is a primary link between genetic variation and disease

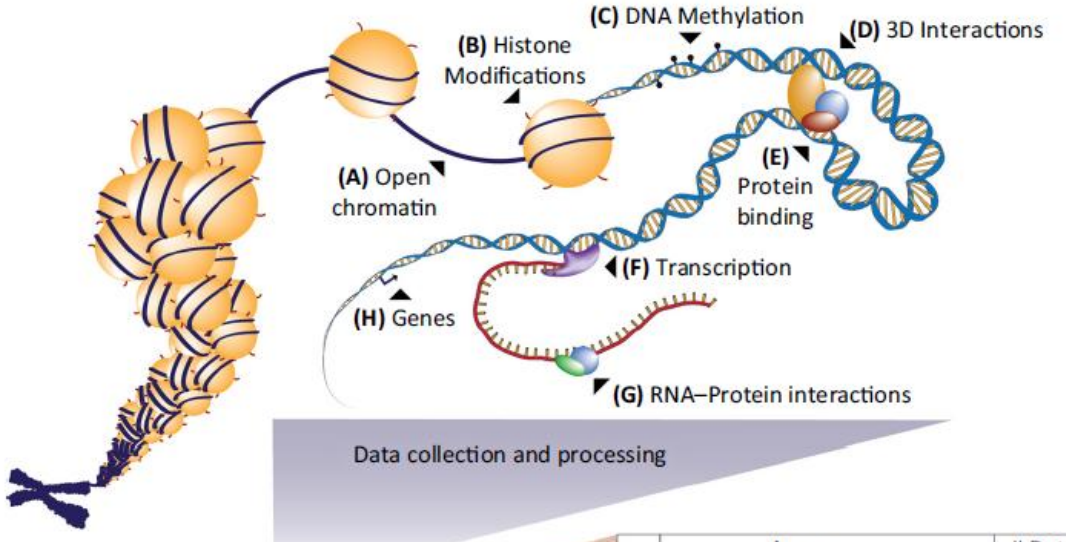
Yang I. Li,[1] Bryce van de Geijn,[2] Anil Raj,[1] David A. Knowles,[3,4] Allegra A. Petti,[5] David Golan,[1] Yoav Gilad,[2*] Jonathan K. Pritchard[1,6,7*]

Gene Expression Regulation

Review

Deciphering ENCODE

Adam G. Diehl[1] and Alan P. Boyle[1,2,*]



	Assays	# Datasets
A	DNase-seq	691
	FAIRE-seq	
B	ChIP-seq	2641
C	WGBS	490
	RRBS	
	meDIP-seq	
D	5C	28
	ChIA-PET	
E	ChIP-seq	1230
F	RNA-seq	
G	icLIP/ecLIP	231
	RIP-seq/RIP-ChIP	
H	Computational prediction	--*
	RT-PCR	

* Data in GENCODE

Figure 1. An Illustrative Example of the Data Types Available through the Encyclopedia of DNA Elements (ENCODE) Project Portals. These include measurements of (A) Open chromatin using DNase-seq and FAIRE-seq, (B) ChIP-seq for histone modifications, (C) DNA methylation, (D) 3D interactions, (E) ChIP-seq for transcription factors and other chromatin-associated proteins, (F) transcriptional output including RNA-seq, CAGE, and RNA-PET, (G) RNA–protein interactions, and (H) gene body predictions through computational and manual annotation as part of GENCODE.

Trends in Genetics

Gene Expression Regulation

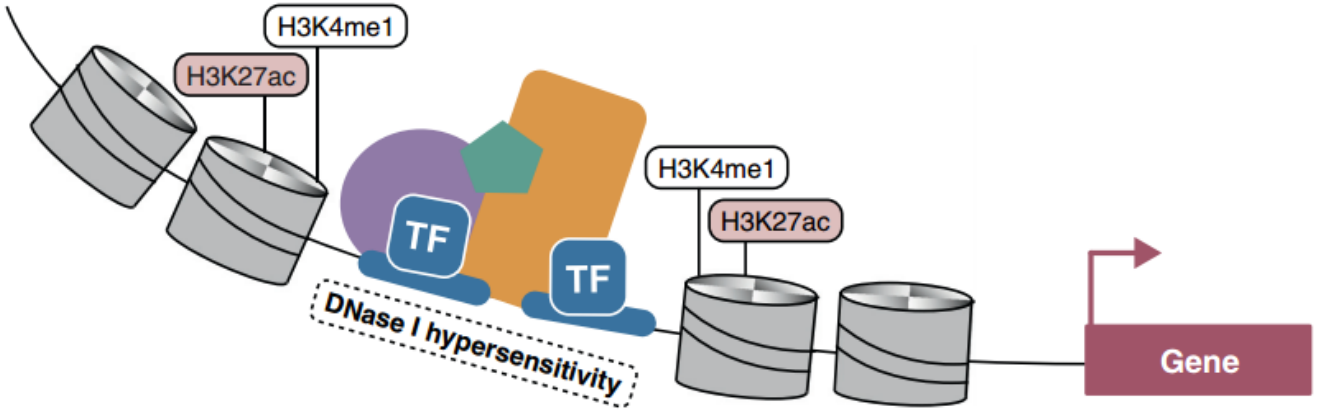


Figure 1 Model of enhancer function. Transcriptional enhancer elements are noncoding stretches of DNA that regulate gene expression levels, most often in *cis*. Active enhancer elements are located in open chromatin sensitive to DNase I digestion and flanked by histones marked with H3K4me1 and H3K27ac. Enhancers are often bound by a number of transcription factors (TF), such as p300 (blue). Mediator and cohesin are part of a complex (orange, green and purple) that mediates physical contacts between enhancers and their target promoters.

Ritchie and Flicek *Genome Medicine* 2014, **6**:87
http://genomemedicine.com/content/6/10/87

Genome Medicine

REVIEW

Computational approaches to interpreting genomic sequence variation

Graham RS Ritchie[1,2] and Paul Flicek[1,2*]

Gene Expression Regulation

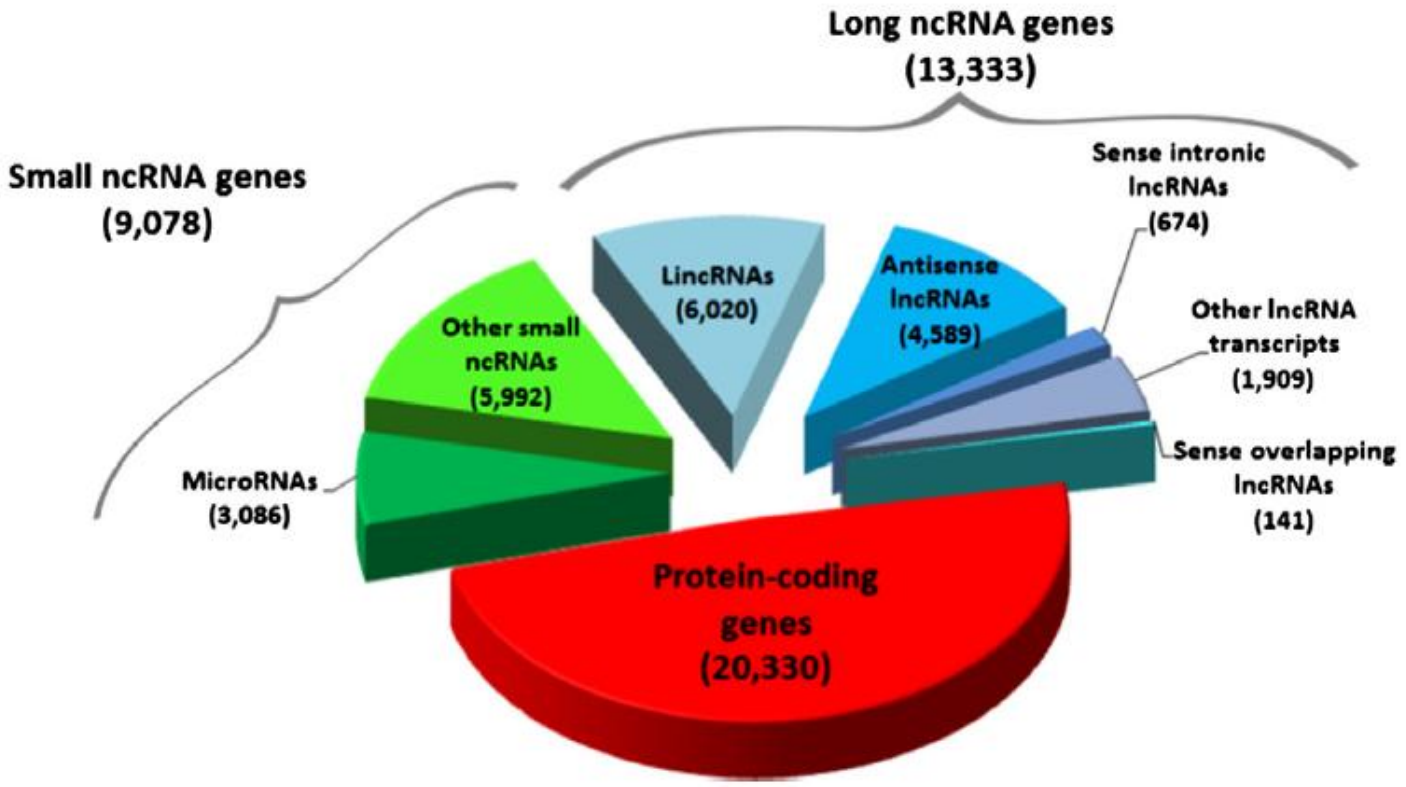


Fig. 1. Abundance of regulatory ncRNA species versus protein coding genes in the human genome. The numbers are based on Gencode V17 (http://www.gencodegenes.org/releases/17.html).

Contents lists available at ScienceDirect
Biochimica et Biophysica Acta
journal homepage: www.elsevier.com/locate/bbadis
Review
Genetic variation in the non-coding genome: Involvement of micro-RNAs and long non-coding RNAs in disease ☆
Barbara Hrdlickova [1], Rodrigo Coutinho de Almeida [1], Zuzanna Borek, Sebo Withoff *

Gene Expression Regulation: DHS

We identified 8,902 locations at which the DNase-seq read depth correlated significantly with genotype at a nearby single nucleotide polymorphism or insertion/deletion (false discovery rate = 10%). We call such variants 'DNase I sensitivity quantitative trait loci' (dsQTLs). We found that dsQTLs are strongly enriched within inferred transcription factor binding sites and are frequently associated with allele-specific changes in transcription factor binding. A substantial fraction (16%) of dsQTLs are also associated with variation in the expression levels of nearby genes (that is, these loci are also classified as eQTLs). Conversely, we estimate that as many as 55% of eQTL single nucleotide polymorphisms are also dsQTLs. Our observations indicate that dsQTLs are highly abundant in the human genome and are likely to be important contributors to phenotypic variation.

DNase I sensitivity QTLs are a major determinant of human expression variation

Jacob F. Degner[1,2]*, Athma A. Pai[1]*, Roger Pique-Regi[1]*, Jean-Baptiste Veyrieras[1,3], Daniel J. Gaffney[1,4], Joseph K. Pickrell[1], Sherryl De Leon[4], Katelyn Michelini[4], Noah Lewellen[4], Gregory E. Crawford[5,6], Matthew Stephens[1,7], Yoav Gilad[1] & Jonathan K. Pritchard[1,4]

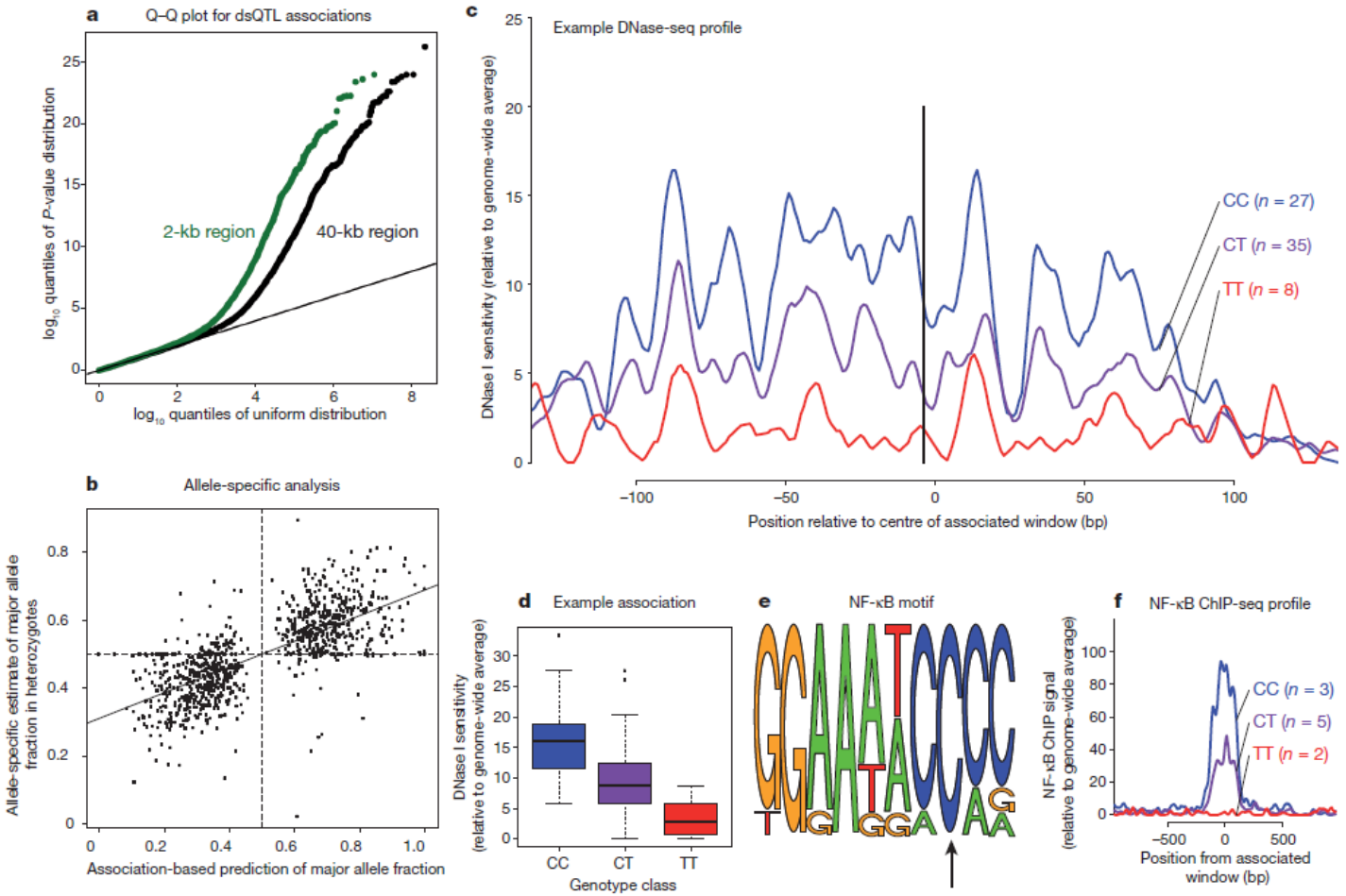


Figure 1 | Genome-wide identification of dsQTLs and a typical example. **a**, Q–Q plots for all tests of association between DNase I cut rates in 100-bp windows, and variants within 2-kb (green) and 40-kb (black) regions centred on the target DHS windows. **b**, Allele-specific analysis of dsQTLs in heterozygotes. Plotted are the predicted (*x* axis) and observed (*y* axis) fractions of reads carrying the major allele based on the genotype means. **c**, Example of a dsQTL (rs4953223). The black line indicates the position of the associated SNP. **d**, Box plot showing that rs4953223 is strongly associated with local chromatin accessibility ($P = 3 \times 10^{-13}$). **e**, The T allele, which is associated with low DNase I sensitivity, disrupts the binding motif of a previously identified NF-κB-binding site at this location[14]. **f**, NF-κB ChIP-seq data from ten individuals[7] indicates a strong effect of this SNP on NF-κB binding.

DNase I sensitivity QTLs are a major determinant of human expression variation

Jacob F. Degner[1,2]*, Athma A. Pai[1]*, Roger Pique-Regi[1]*, Jean-Baptiste Veyrieras[1,3], Daniel J. Gaffney[1,4], Joseph K. Pickrell[1], Sherryl De Leon[4], Katelyn Michelini[4], Noah Lewellen[4], Gregory E. Crawford[5,6], Matthew Stephens[1,7], Yoav Gilad[1] & Jonathan K. Pritchard[1,4]

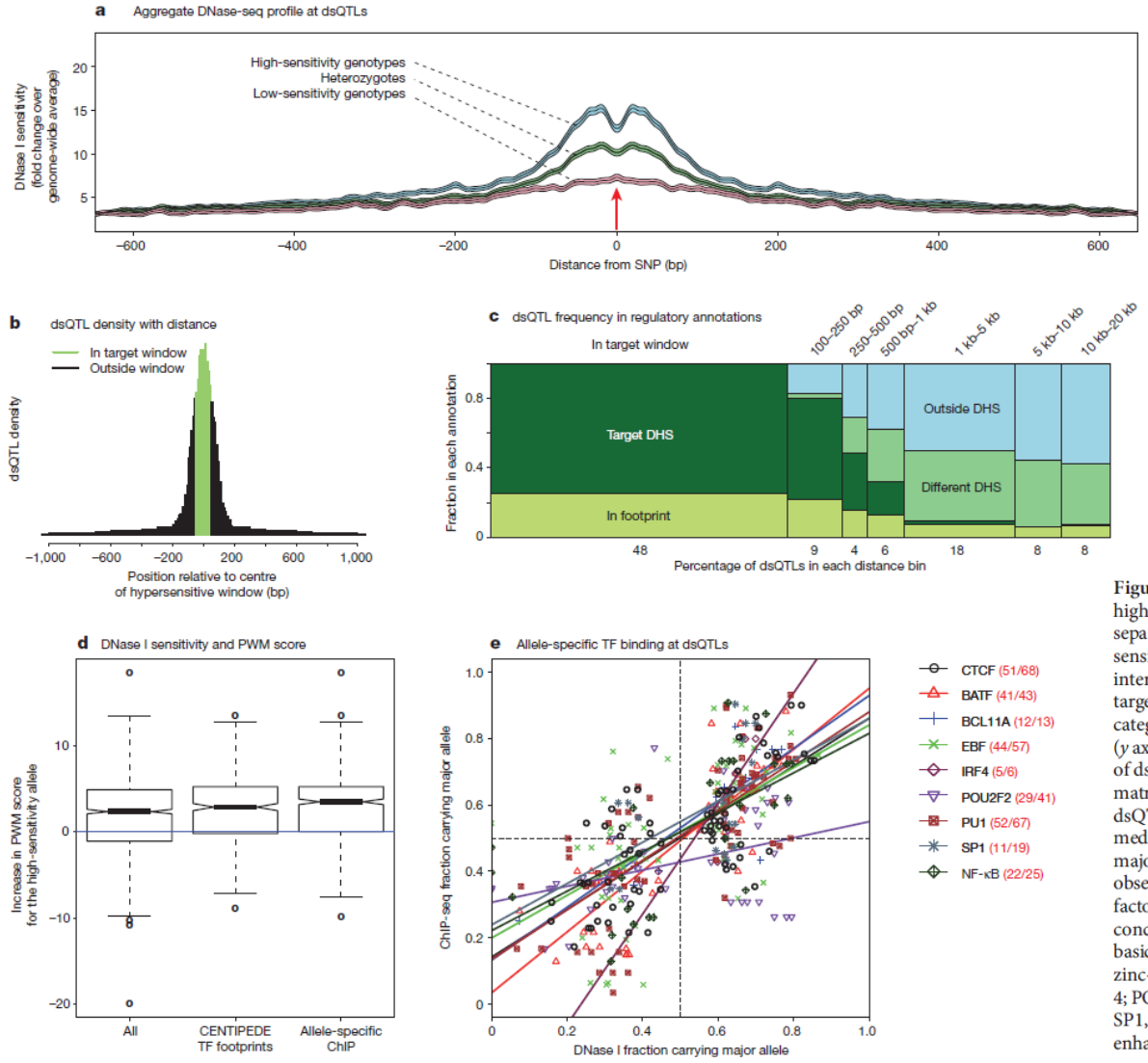


Figure 2 | Properties of dsQTLs. **a**, Aggregated plot of DNase I sensitivity for high-confidence dsQTLs that lie within the target DHS. Individuals were separated into the high-sensitivity (blue), heterozygote (green), and low-sensitivity (red) classes. The shading indicates the bootstrap 95% confidence intervals. **b**, The peak density of dsQTLs is very tightly focused around the target DHS window. **c**, Total fraction of *cis*-dsQTLs that fall into different categories of distance from the target window (*x* axis) and different annotations (*y* axis). The total area of each rectangle is proportional to the estimated number of dsQTLs in that category. **d**, Box plot showing distribution of position weight matrix (PWM) score differences between high-sensitivity and low-sensitivity dsQTL alleles, respectively. Notches indicate 95% confidence intervals for median. **e**, The *x* axis shows the fraction of sequence reads predicted to carry the major allele based on the DNase I genotype means; the *y* axis shows the observed fraction in ChIP-seq data. The lines show the regression fits for each factor separately; the numbers in the key show the fraction of sites that are in a concordant direction for each factor. CTCF, CCCTC binding factor; BATF, basic leucine zipper transcription factor; BCL11A, B-cell CLL/lymphoma 11A zinc-finger protein; EBF, early B-cell factor 1; IRF4, interferon regulatory factor 4; POU2F2, POU class 2 homeobox 2; PU1, proviral integration oncogene spi1; SP1, Sp1 transcription factor; NF-κB, nuclear factor of κ light polypeptide gene enhancer in B-cells 1.

DNase I sensitivity QTLs are a major determinant of human expression variation

Jacob F. Degner[1,2]*, Athma A. Pai[1]*, Roger Pique-Regi[1]*, Jean-Baptiste Veyrieras[1,3], Daniel J. Gaffney[1,4], Joseph K. Pickrell[1], Sherryl De Leon[4], Katelyn Michelini[4], Noah Lewellen[4], Gregory E. Crawford[5,6], Matthew Stephens[1,7], Yoav Gilad[1] & Jonathan K. Pritchard[1,4]

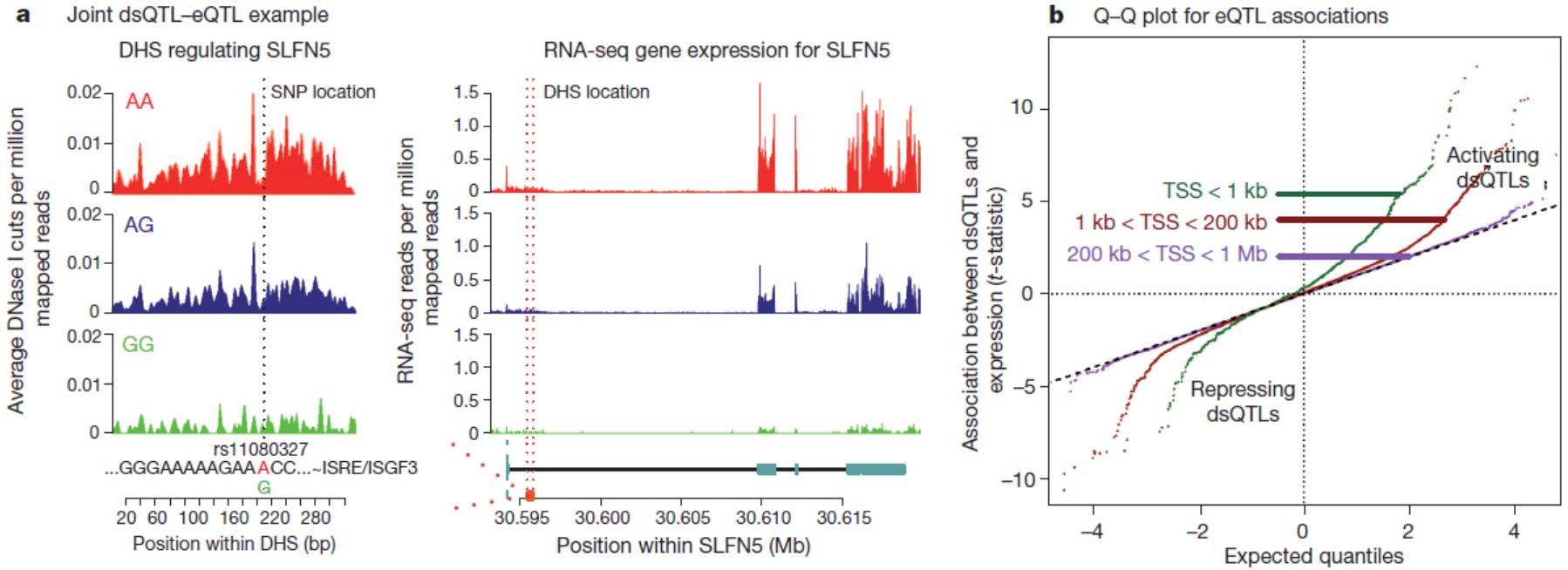


Figure 3 | Relationship between dsQTLs and eQTLs. **a**, Example of a dsQTL SNP that is also an eQTL for the gene *SLFN5*. The SNP disrupts an interferon-sensitive response element, thereby changing local chromatin accessibility within the first intron of *SLFN5*. Expression of *SLFN5* has been shown to be inducible by interferon α in melanoma cell lines. DNase-seq (left) and RNA-seq (right) measurements from DNase-seq and RNA-seq are plotted, stratified by genotype at the putative causal SNP. **b**, Q–Q plot of the *t*-statistic for association with gene expression changes (eQTL) of dsQTL SNPs. The sign of the eQTL *t*-statistic is with respect to the genotype that increases DNase sensitivity.

DNase I sensitivity QTLs are a major determinant of human expression variation

Jacob F. Degner[1,2]*, Athma A. Pai[1]*, Roger Pique-Regi[1]*, Jean-Baptiste Veyrieras[1,3], Daniel J. Gaffney[1,4], Joseph K. Pickrell[1], Sherryl De Leon[4], Katelyn Michelini[4], Noah Lewellen[4], Gregory E. Crawford[5,6], Matthew Stephens[1,7], Yoav Gilad[1] & Jonathan K. Pritchard[1,4]

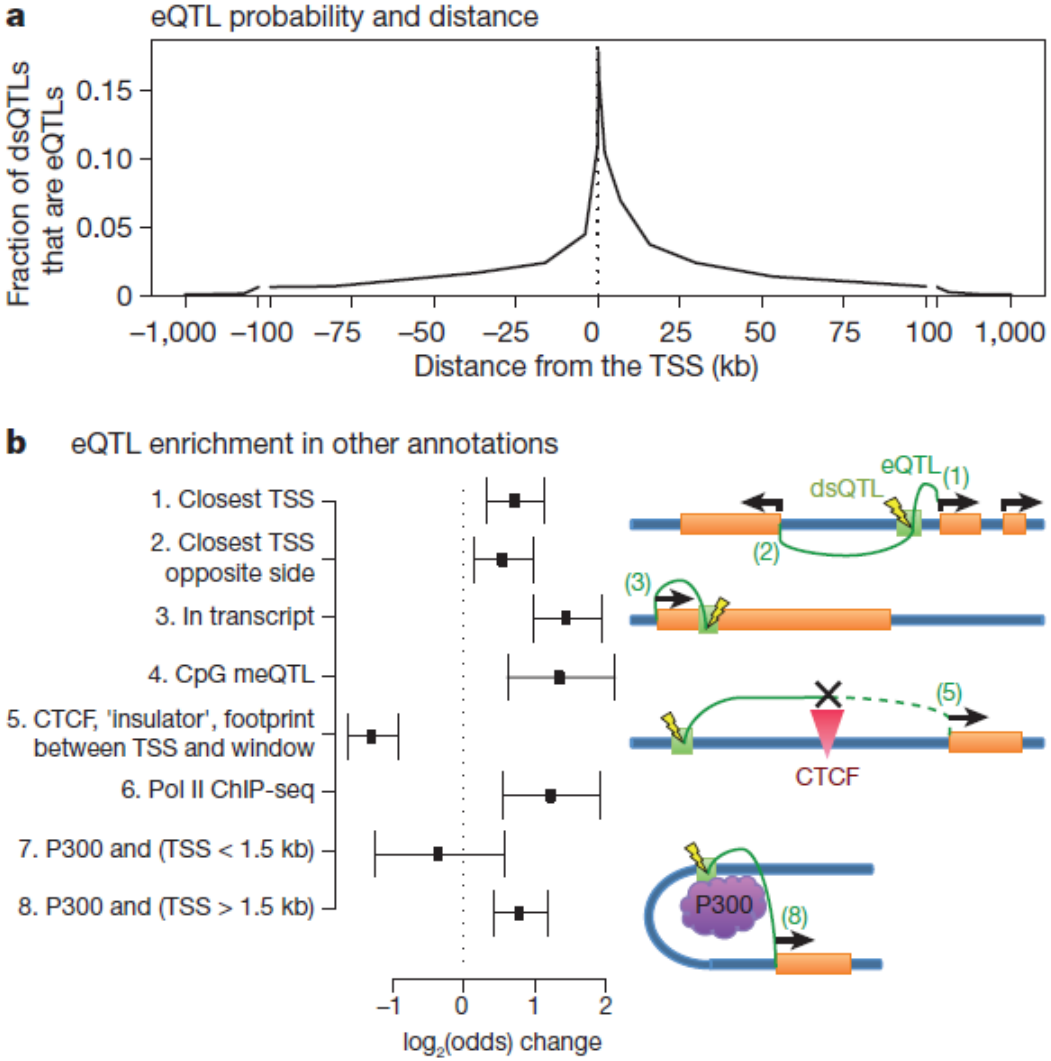
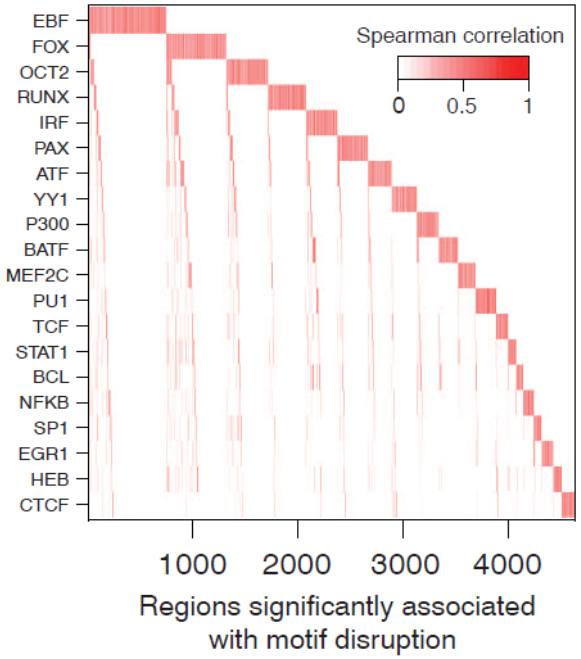


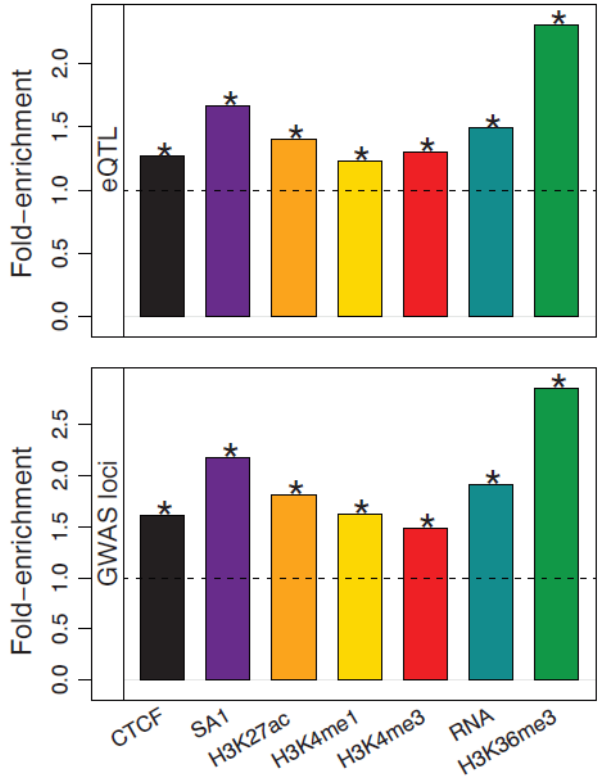
Figure 4 | Relationship between dsQTLs and eQTLs. **a**, Most joint dsQTL–eQTLs lie close to the gene TSS. **b**, Effect of various factors on the log odds that a given dsQTL is also an eQTL, while controlling for the strong distance relationship observed in **a**. In annotations (1) and (2) we do not consider the direction of transcription. In annotations (6–8) ChIP-seq is measured on the dsQTL window. In annotations (4) and (6), 'meQTL' refers to a dsQTL that is also associated with methylation levels of a nearby CpG site[27] and 'Pol II' refers to the presence of an RNA polymerase II ChIP-seq peak overlapping the DHS associated with the dsQTL[23]. One of the most significant annotations in delineating the regulatory regions is defined by the presence of the CTCF insulator element, which decreases 2.4-fold the probability that a dsQTL is an eQTL. Error bars represent 95% confidence intervals.

DNase I sensitivity QTLs are a major determinant of human expression variation

Jacob F. Degner[1,2]*, Athma A. Pai[1]*, Roger Pique-Regi[1]*, Jean-Baptiste Veyrieras[1,3], Daniel J. Gaffney[1,4], Joseph K. Pickrell[1], Sherryl De Leon[4], Katelyn Michelini[4], Noah Lewellen[4], Gregory E. Crawford[5,6], Matthew Stephens[1,7], Yoav Gilad[1] & Jonathan K. Pritchard[1,4]

Gene Expression Regulation





B

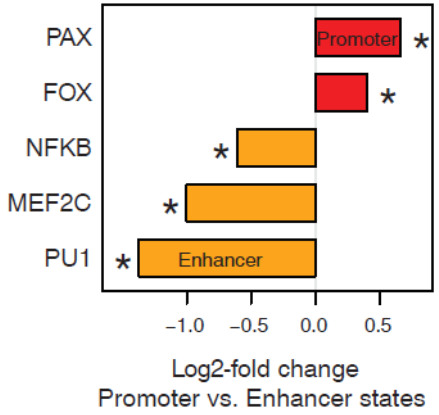


Fig. 4. Mechanism and functional consequences of chromatin variation. (**A**) Correlation coefficients of TF motif disruption scores and H3K27ac signal across individuals. Motifs are sorted based on the number of associated peaks; peaks are sorted based on their associated motifs. (**B**) Log2 fold-enrichment of motifs in promoter (red) versus enhancer (orange) states. Only significant enrichments (Fisher's exact test $P < 0.05$) are shown. (**C**) eQTLs and GWAS hits in variable regions. Stars indicate $P < 0.05$.

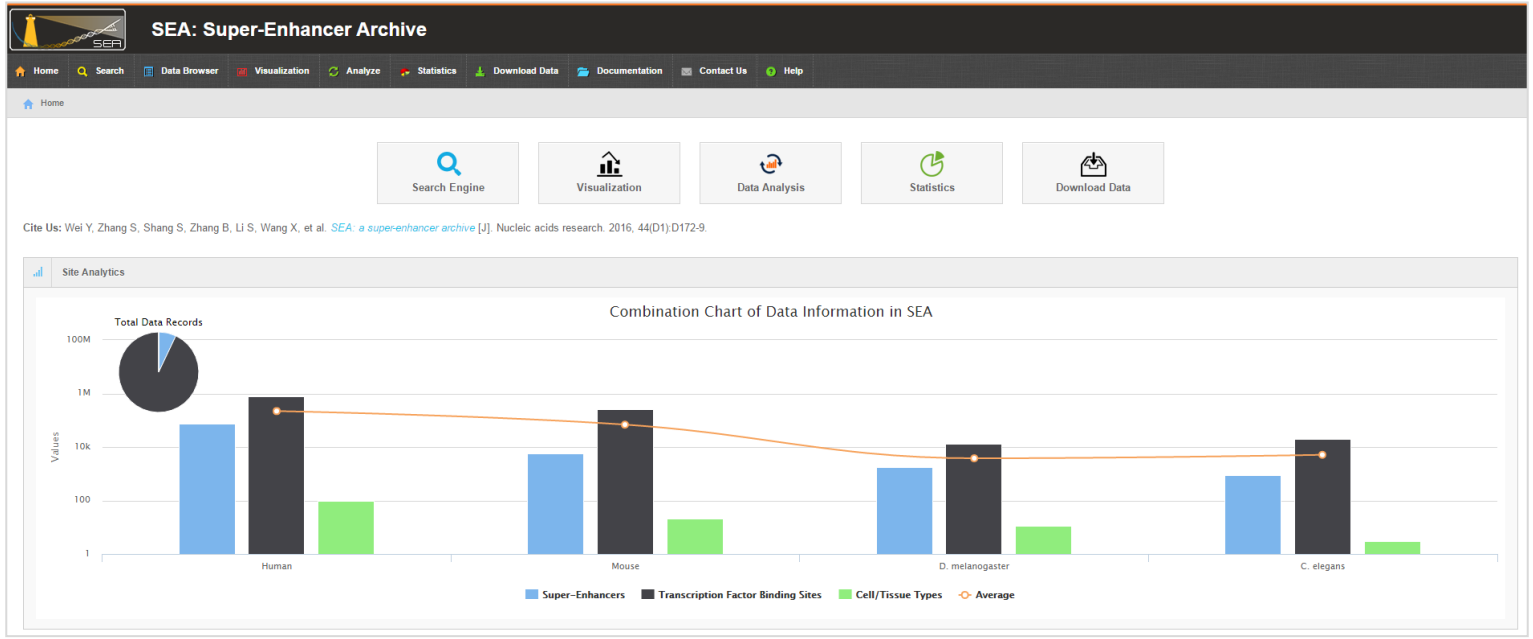
Extensive Variation in Chromatin States Across Humans

Maya Kasowski,[1,2*] Sofia Kyriazopoulou-Panagiotopoulou,[3*] Fabian Grubert,[1*] Judith B. Zaugg,[1*] Anshul Kundaje,[1,3,4,5*] Yuling Liu,[8] Alan P. Boyle,[1] Qiangfeng Cliff Zhang,[1] Fouad Zakharia,[1] Damek V. Spacek,[1] Jingjing Li,[1] Dan Xie,[1] Anthony Olarerin-George,[6] Lars M. Steinmetz,[1,7] John B. Hogenesch,[6] Manolis Kellis,[4,5] Serafim Batzoglou,[3] Michael Snyder[1†]

Gene Expression Regulation: Super Enhancers

Super-Enhancers in the Control of Cell Identity and Disease

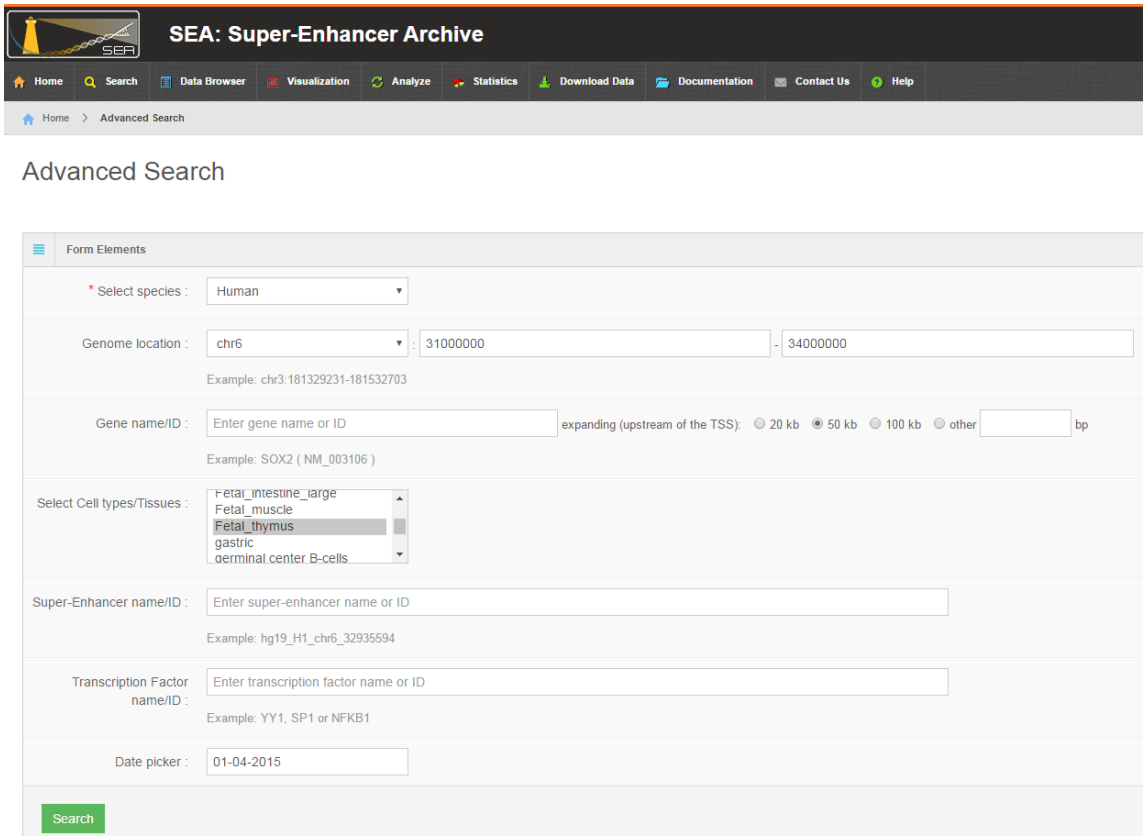
Denes Hnisz,[1,3] Brian J. Abraham,[1,3] Tong Ihn Lee,[1,3] Ashley Lau,[1,2] Violaine Saint-André,[1] Alla A. Sigova,[1] Heather A. Hoke,[1,2] and Richard A. Young[1,2,*]



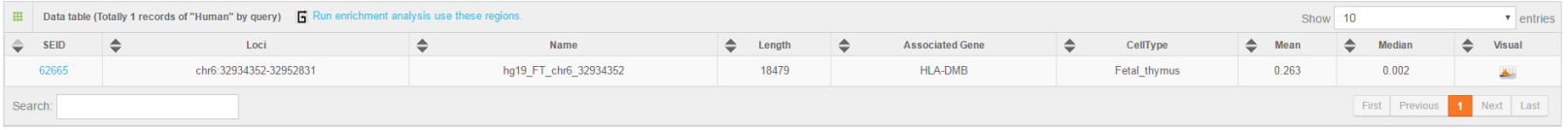
SEA: a super-enhancer archive

Yanjun Wei[1,†], Shumei Zhang[1,†], Shipeng Shang[1,†], Bin Zhang[1], Song Li[1], Xinyu Wang[1], Fang Wang[1], Jianzhong Su[1], Qiong Wu[2], Hongbo Liu[1,*] and Yan Zhang[1,*]

Gene Expression Regulation: Super Enhancers



Search Results

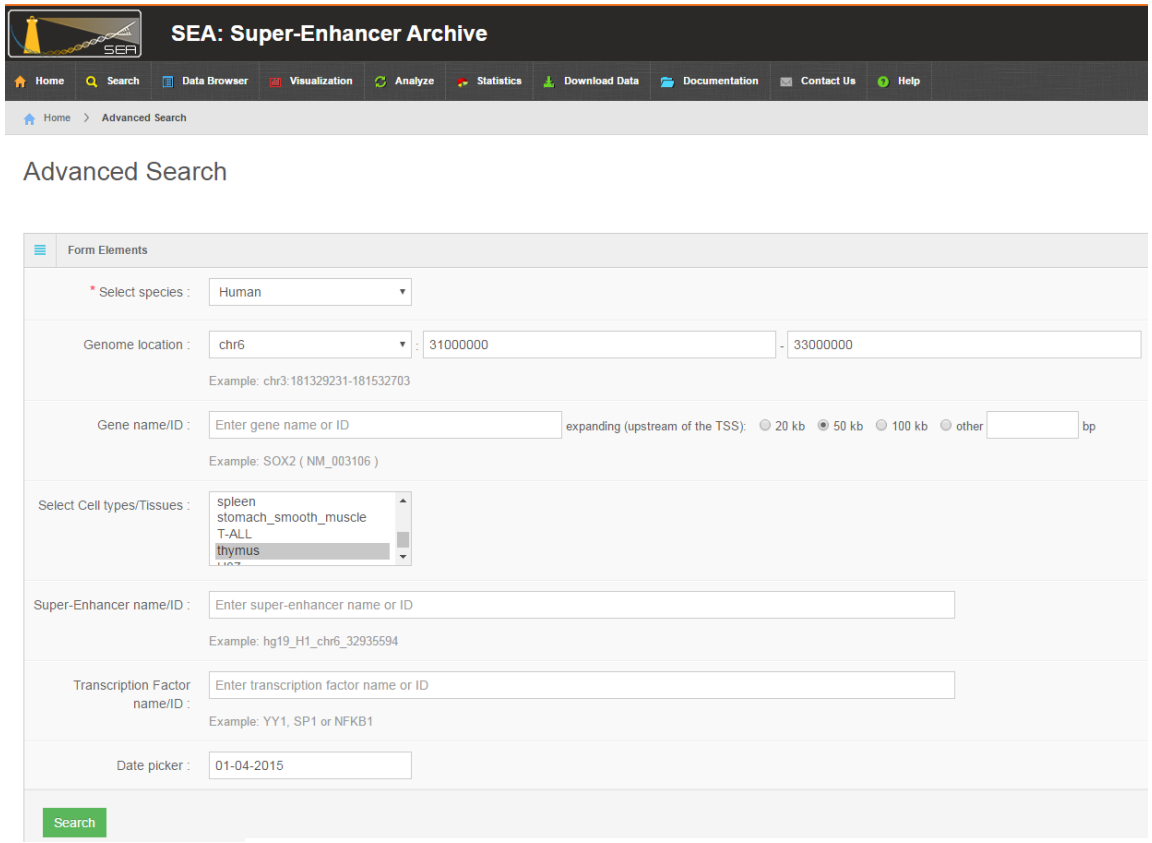


SEA: a super-enhancer archive

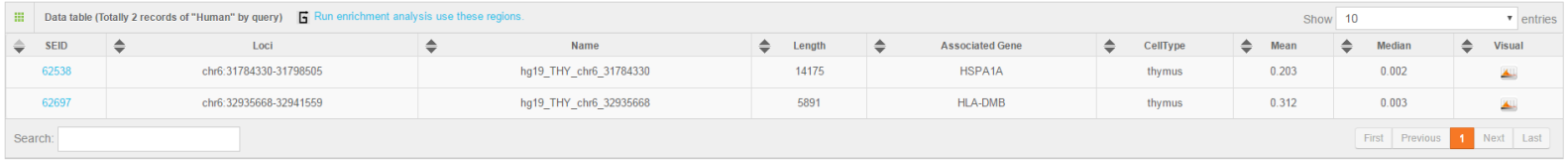
Yanjun Wei[1,†], Shumei Zhang[1,†], Shipeng Shang[1,†], Bin Zhang[1], Song Li[1], Xinyu Wang[1], Fang Wang[1], Jianzhong Su[1], Qiong Wu[2], Hongbo Liu[1,*] and Yan Zhang[1,*]

Liverpool Hope University

Gene Expression Regulation: Super Enhancers



Search Results



SEA: a super-enhancer archive

Yanjun Wei[1,†], Shumei Zhang[1,†], Shipeng Shang[1,†], Bin Zhang[1], Song Li[1], Xinyu Wang[1], Fang Wang[1], Jianzhong Su[1], Qiong Wu[2], Hongbo Liu[1,*] and Yan Zhang[1,*]

Liverpool Hope University EST. 1844

GeneHancer in GeneCards

Genomics for HFE Gene

Products: Regulatory Element

Regulatory Elements for HFE Gene

Enhancers for HFE Gene IMPROVED!

Filter: (9 results) See all 9 »

GeneHancer Identifier	Enhancer Score	Enhancer Sources	Gene-Enhancer Score	TSS distance (kb)	Number of Genes Away	Size (kb)	Transcription Factor Binding Sites within enhancer	Gene Targets for Enhancer
GH06F026060	0.9★	Ensembl, ENCODE	11.5	-27.0	1	0.6	2 TFs CEBPB ZBTB33	11 genes HIST1H4C
GH06F026091	0.6	Ensembl	11.1★	+5.1	0	1.2	5 TFs RAD51 IKZF1	6 genes ENSG0000027246
GH06F026102	0.7	ENCODE	11.1	+17.0	3	3.6	307 TFs MLX CREB3L1	10 genes ENSG0000027246
GH06F026120	1★	FANTOM5, ENCODE	9.9	+36.7	17	7.8	306 TFs MLX CREB3L1	35 genes ABT1
GH06F026042	0.2	ENCODE	7.3	-42.5	9	4.7	142 TFs HDGF PKNOX1	8 genes ENSG0000027246

★ - Elite enhancer/Elite enhancer-gene association

Download Table

Download GeneHancer data dump

Enhancers around HFE on UCSC Golden Path with GeneCards custom track

GeneHancer: genome-wide integration of enhancers and target genes in GeneCards

Simon Fishilevich[1,†], Ron Nudel[1,†], Noa Rappaport[1], Rotem Hadar[1], Inbar Plaschkes[1], Tsippi Iny Stein[1], Naomi Rosen[1], Asher Kohn[2], Michal Twik[1], Marilyn Safran[1], Doron Lancet[1,*] and Dana Cohen[1,*]

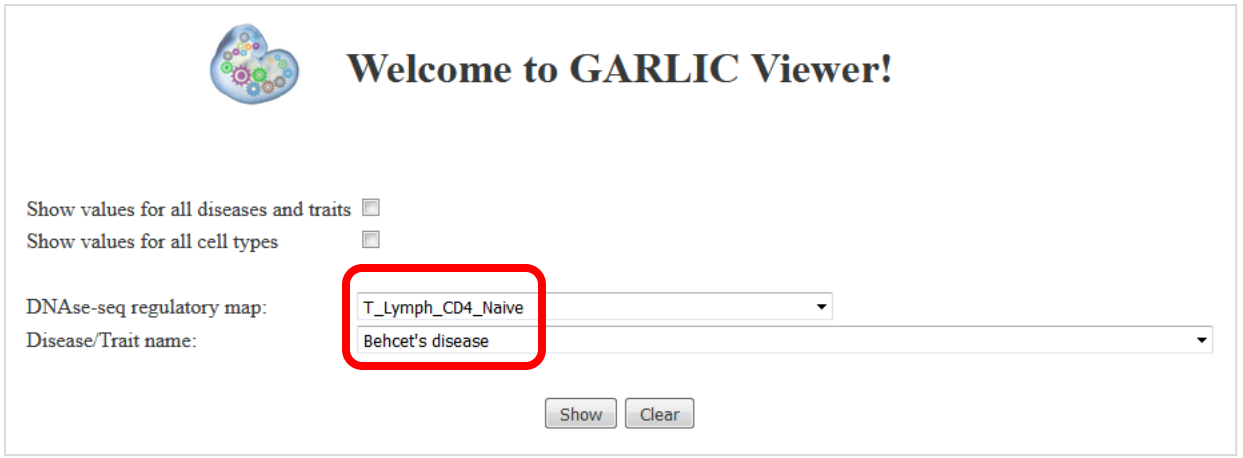
Liverpool Hope University

Gene Expression Regulation: GARLIC

GARLIC: a bioinformatic toolkit for aetiologically connecting diseases and cell type-specific regulatory maps

Miloš Nikolić; Argyris Papantonis; Alvaro Rada-Iglesias

Gene Expression Regulation: GARLIC



GARLIC: a bioinformatic toolkit for aetiologically connecting diseases and cell type-specific regulatory maps

Miloš Nikolić; Argyris Papantonis; Alvaro Rada-Iglesias

3D Genome

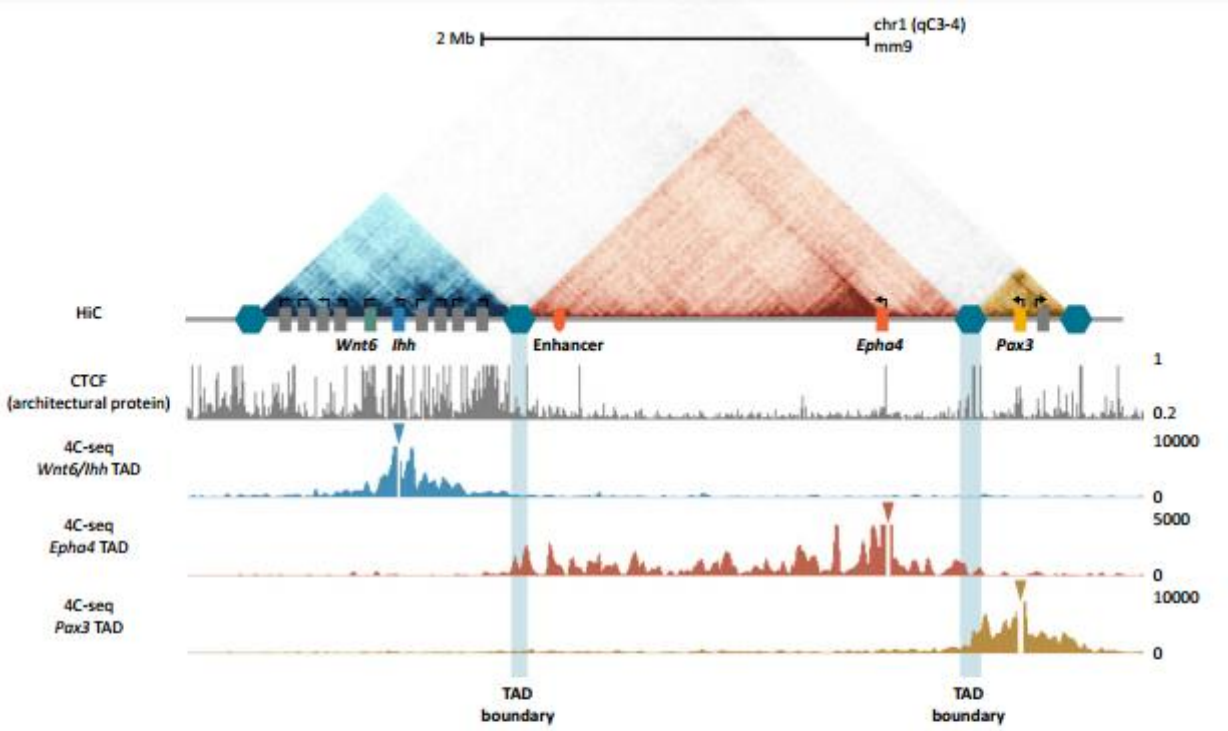


Figure 1. Higher-Order Chromatin Folding and Topologically Associating Domain (TAD) Structure at the *Eph receptor A4* (*Epha4*) Locus. Hi-C interactions are shown in a heat map in which each dot reflects two interaction pairs. The resulting interaction profile shows the formation of triangles (schematically enhanced in color) that represent individual TADs. There is a high degree of interaction within each TAD but little contact between TADs. Abrupt changes in the directionality of contacts demarcate boundary regions (blue hexagon). Of note is the very large TAD containing one (*Epha4*) gene, whereas the flanking TADs are much smaller (right) or contain many genes (left). Below, the binding profile of the CCCTC-binding factor (CTCF) transcription factor is shown. Note the scarcity of binding sites in the *Epha4* TAD and the enrichment at the boundaries. CTCF is also associated with gene promoters. The 4C-seq profiles of the viewpoints *Indian hedgehog* (*Ihh*), *Epha4*, and *Paired box3* (*Pax3*) are depicted below. Note that the interaction profiles are restricted to the respective TADs. Data from [23,30].

Breaking TADs: How Alterations of Chromatin Domains Result in Disease

Darío G. Lupiáñez,[1,2,3,@] Malte Spielmann,[1,2,3] and Stefan Mundlos[1,2,3,*]

3D Genome

The eukaryotic nucleus is a complex 3D environment in which genome function depends not only on the linear arrangement of regulatory sequence elements, but also on their spatial organization for effective control of gene expression.

Analysis of the role of chromatin 3D organization in gene expression is progressing rapidly, largely due to the development of chromosome conformation capture methods such as Hi-C.

Sequences within these Topologically Associated Domains (TADs) interact more frequently with sites inside than outside the domain. TADs with a median size of 880 kb have been found in mammals.

Breaking TADs: How Alterations of Chromatin Domains Result in Disease

Darío G. Lupiáñez,[1,2,3,@] Malte Spielmann,[1,2,3] and Stefan Mundlos[1,2,3,*]

3D Genome

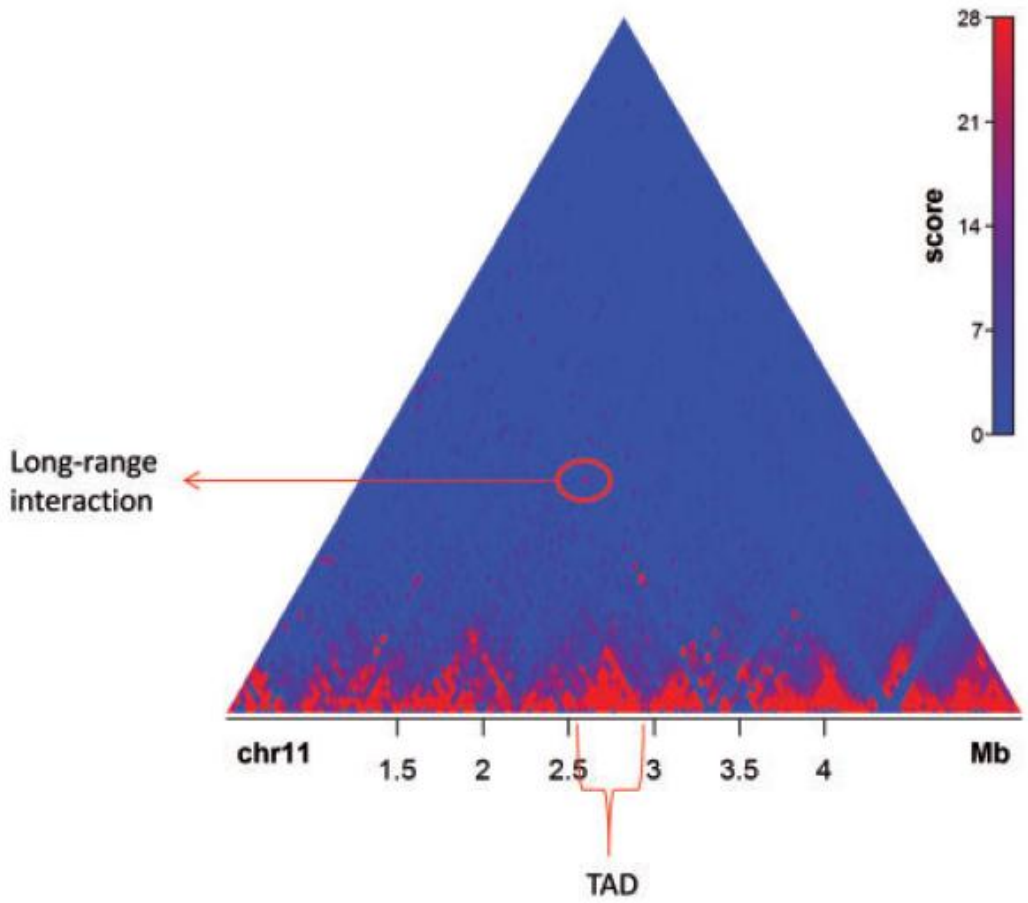


Figure 2. An example of a Hi-C contact map. Hi-C contact map of a segment of mouse chromosome 11, generated using Sushi [90] from Dixon *et al.* [85] data. A TAD and a long-range interaction between two loci are annotated. A colour version of this figure is available online at BIB online: https://academic.oup.com/bib.

OXFORD

Briefings in Bioinformatics, 17(6), 2016, 980–995

doi: 10.1093/bib/bbv097
Advance Access Publication Date: 19 November 2015
Software Review

In the loop: promoter–enhancer interactions and bioinformatics

Antonio Mora, Geir Kjetil Sandve, Odd Stokke Gabrielsen and Ragnhild Eskeland

3D Genome: CCSI

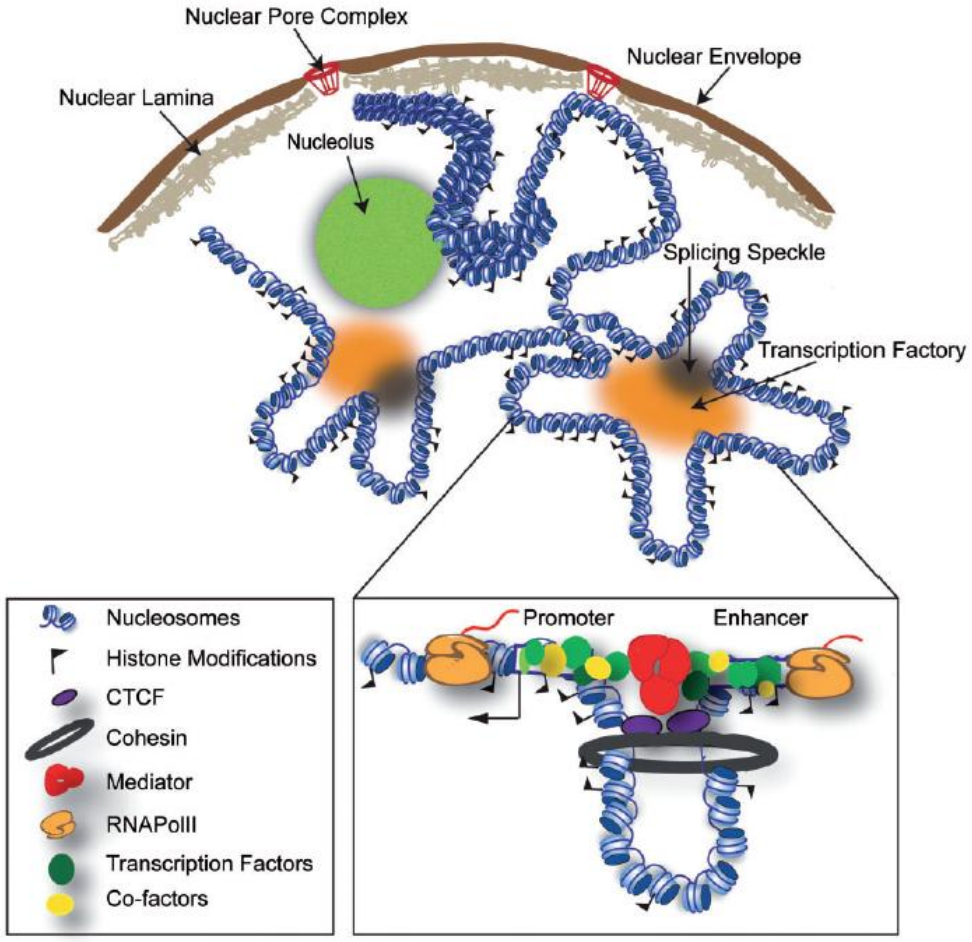


Figure 1. Models of chromatin organization. A diagram of different models of chromatin organization in the nuclear space. Interphase chromatin that interacts with the nuclear lamina (grey), nucleolus (green), nuclear pores (red), transcription factories (orange) and splicing speckles (black) are depicted here. Generally, lamin- and nucleolar-associated domains are transcriptionally repressed and have a more condensed chromatin, whereas chromatin that loops to the nuclear pore, transcription factories and splicing speckles are transcriptionally active and therefore have a more open chromatin structure (here, depicted as 10 nm chromatin fibre). Enhancers can activate gene expression over a distance and contain binding sites for TFs that recruit co-factors (activators or repressors). A promoter–enhancer looping mechanism mediated by cohesin (brown), CTCF (purple) and the mediator complex (red) that brings the enhancer into close proximity to its target promoter are presented in the enlarged box. The enhancer and promoter are marked with white boxes, and the transcription start site of the transcribed target gene is annotated with an arrow. TFs (green) and co-factors (yellow) bind the enhancer and are brought close to the basal transcription machinery at the promoter. RNAPolII (orange) transcribes pre-mRNA from the target gene and eRNA from the enhancer. Some of these models may co-exist for different PEIs; however, there are also other models that we could not show. A colour version of this figure is available online at BIB online: https://academic.oup.com/bib.

OXFORD

Briefings in Bioinformatics, 17(6), 2016, 980–995

doi: 10.1093/bib/bbv097
Advance Access Publication Date: 19 November 2015
Software Review

In the loop: promoter–enhancer interactions and bioinformatics

Antonio Mora, Geir Kjetil Sandve, Odd Stokke Gabrielsen and Ragnhild Eskeland

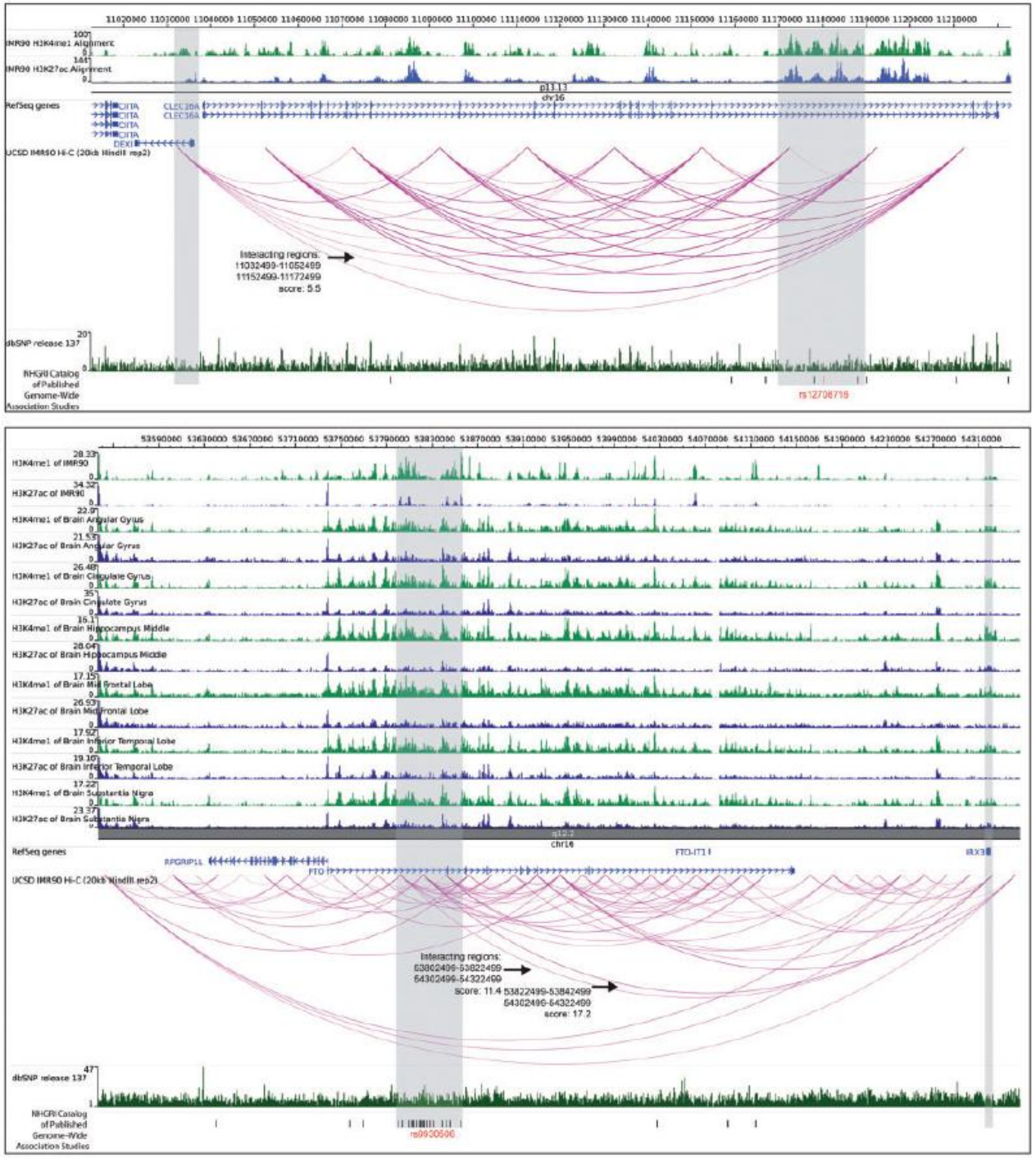


Figure 3. Long-range interactions functionally connect disease-associated SNPs with disease candidate genes. (**A**) Physical proximity between *DEXI* gene locus and autoimmune disease-associated SNPs in the intron of *CLEC16A*. Hi-C data from human foetal lung (IMR-90) cells (from the Ren lab, [85]) show interactions between *CLEC16A* intron 19 and the *DEXI* locus. The enhancer marks in IMR-90 cells for H3K4me1 and H3K27ac are shown in green and blue, respectively, and the filter threshold for the Hi-C data was set to 5. SNPs in the region are in black, and the eQTL SNP rs12708716 is marked in red. The arc (pink) for interacting regions (grey) is highlighted with an arrow. (**B**) Long-range interactions links obesity-associated variants in *FTO* with the *IRX3* locus. Hi-C data in human foetal lung (IMR-90) cells show interactions between the first intron of *FTO* with *IRX3*. The tracks for H3K4me1 and H3K27ac are shown in green and blue from IMR-90 cells and different human brain tissues from the NIH Roadmap Epigenomics Mapping Consortium. The filter threshold for the Hi-C data was set to 10. SNPs in the region are in black, and the BMI-associated SNP rs9930506 is marked in red. Arcs (pink) for interacting regions (grey) are highlighted with arrows. These public data sets are available and visualized with the WashU EpiGenome Browser (http://epigenomegateway.wustl.edu/browser/). dbSNP release 137 is shown in dark green, and the The National Human Genome Research Institute (NHGRI) Catalogues of GWAS are visualized in UCSC browser (http://genome-euro.ucsc.edu) [166]. A colour version of this figure is available online at BIB online: https://academic.oup.com/bib

OXFORD

Briefings in Bioinformatics, 17(6), 2016, 980–995

doi: 10.1093/bib/bbv097
Advance Access Publication Date: 19 November 2015
Software Review

In the loop: promoter–enhancer interactions and bioinformatics

Antonio Mora, Geir Kjetil Sandve, Odd Stokke Gabrielsen and Ragnhild Eskeland

3D Genome: CCSI

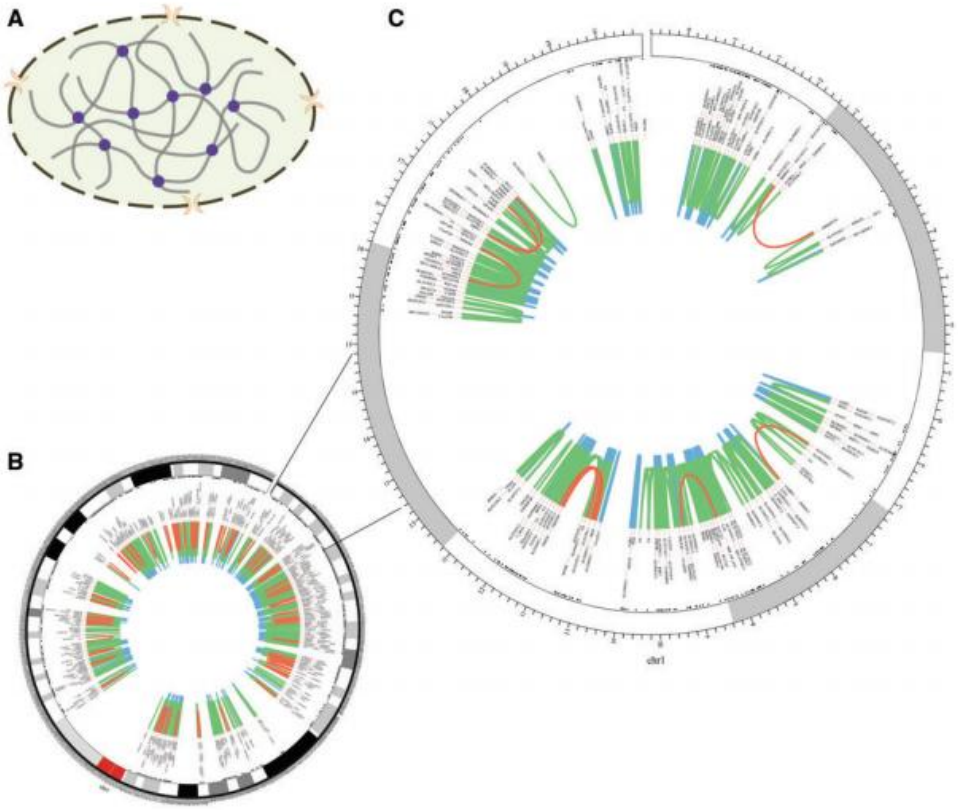


Figure 2. Chromosomal 3D structure and promoter–promoter interactions of Chr1 in IMR90 cell line based on a set of Hi-C data. **(A)** Chromosomal 3D structure. The dashed circle with two orange crescents that stand for nuclear pore complex is the nucleus membrane. The thick grey lines are chromatin and the purple circles stand for proteins that link chromatin together. **(B)** Promoter–promoter interactions of Chr1. **(C)** Promoter–promoter interactions of Chr1:1-20000000, zooming into the interactions. The red lines stand for long-range interactions (distance between interaction pair > 500 kb), while the blue lines for short-range (distance < 50 kb) and the green lines for middle-range (distance spanning 50–500 kb). The black texts are the gene names of corresponding loci.

DATABASE

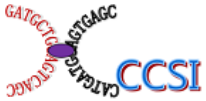
Database, 2016, 1–7
doi: 10.1093/database/bav124
Original article

Original article

CCSI: a database providing chromatin–chromatin spatial interaction information

Xiaowei Xie[1], Wenbin Ma[1], Zhou Songyang[1], Zhenhua Luo[1,2], Junfeng Huang[1], Zhiming Dai[3,*] and Yuanyan Xiong[1,4,*]

3D Genome: CCSI



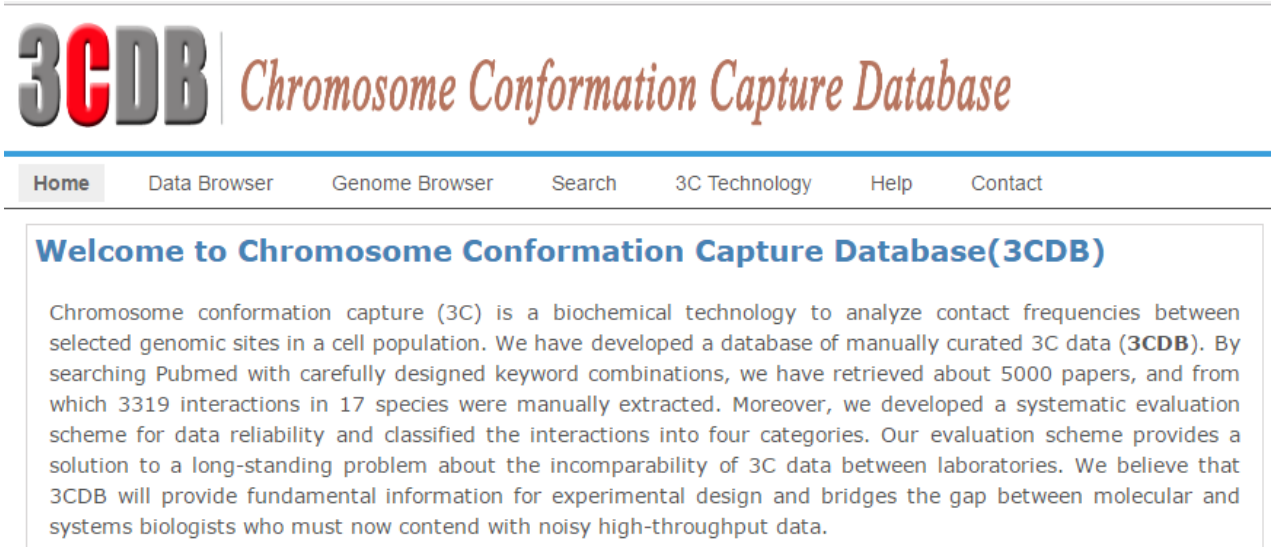
Chromatin Chromatin Space Interaction

Home | Method | Search | Download | Help | Update

Welcome to CCSI database

Here, **CCSI (Chromatin Chromatin Space Interaction)** database presents 3,017,962 chromatin interaction pairs with annotation of genes, enhancers and SNPs in many cell lines of human, mouse and yeast. These data were obtained by means of **3C, 4C, 5C, ChIA-PET and Hi-C** technology in a cell's natural state, nearly all of which detected the three-dimensional architecture of chromosome by coupling ligation in close spatial proximity followed by high-throughput sequencing. So transcriptional regulatory mechanism in disease pathogenesis associated with spatial interactions among genes, enhancers and SNPs could be explored on the base of it.

3D Genome: 3CDB



3CDB Chromosome Conformation Capture Database

MYC Search

General Search,e.g. CD34 or CD34 ci0000001

Home | Data Browser | Genome Browser | Search | 3C Technology | Help | Contact

Search Result

	Name	Anchor					
		Locus1	Fragment location	Primer sequence	Strand	Locus2	Fragment location
1	hs0000001	MYC,colon cancer risk locus	NA	NA	NA	MYC	NA
2	hs0000002	MYC,breast cancer risk locus	NA	NA	NA	MYC	NA
3	hs0000006	MYC,breast cancer risk locus	NA	NA	NA	MYC	NA
4	hs0000651	MYC,enhancer,segment E	chr8:128406277-128416808	AATTTCCCATCCACATTCCACAAGCAACTGT	+	MYC,promoter	chr8:128745463-128751378
5	hs0000530	MYC,promoter	chr8:128745990-128756984	AGCAGCAGATACCGCCCCTCCT	-	MYC,enhancer,segment A	chr8:128226105-128226859
6	hs0000531	MYC,promoter	chr8:128745990-128756984	AGCAGCAGATACCGCCCCTCCT	-	MYC,enhancer,segment C	chr8:128235179-128236036

DATABASE
Database, 2016, 1–7
doi: 10.1093/database/baw044
Original article

Original article

3CDB: a manually curated database of chromosome conformation capture data

Xiaoxiao Yun[1,2,†], Lili Xia[1,2,†], Bixia Tang[1,2], Hui Zhang[1,2], Feifei Li[1] and Zhihua Zhang[1,*]

[1]CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, 100101, China and [2]University of Chinese Academy of Sciences, Beijing, 100049, China

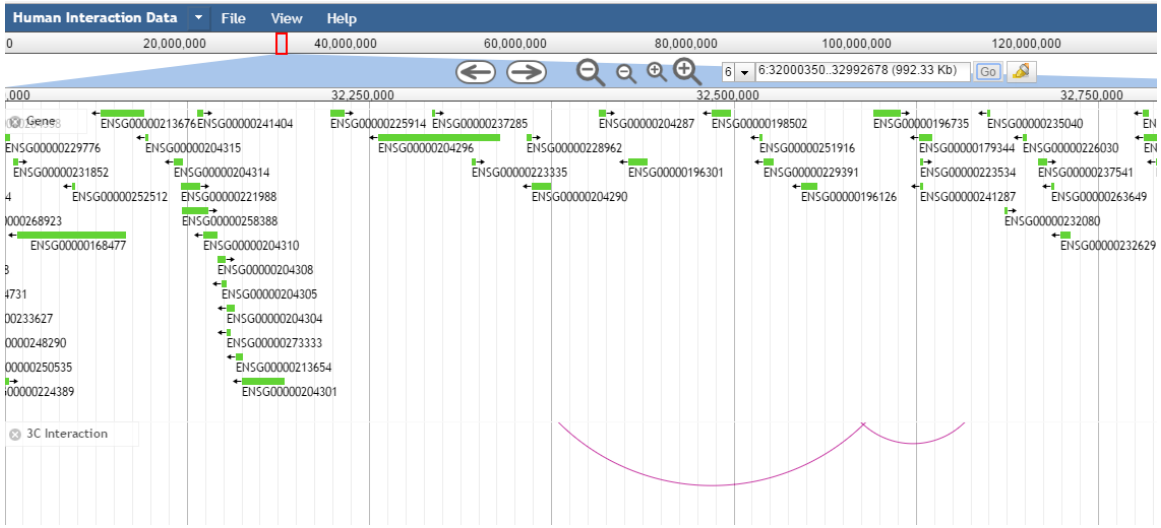
3D Genome: 3CDB

3CDB Chromosome Conformation Capture Database

Home | Data Browser | Genome Browser | Search | 3C Technology | Help | Contact

Welcome to Chromosome Conformation Capture Database(3CDB)

Chromosome conformation capture (3C) is a biochemical technology to analyze contact frequencies between selected genomic sites in a cell population. We have developed a database of manually curated 3C data (**3CDB**). By searching Pubmed with carefully designed keyword combinations, we have retrieved about 5000 papers, and from which 3319 interactions in 17 species were manually extracted. Moreover, we developed a systematic evaluation scheme for data reliability and classified the interactions into four categories. Our evaluation scheme provides a solution to a long-standing problem about the incomparability of 3C data between laboratories. We believe that 3CDB will provide fundamental information for experimental design and bridges the gap between molecular and systems biologists who must now contend with noisy high-throughput data.



DATABASE

Database, 2016, 1–7
doi: 10.1093/database/baw044
Original article

Original article

3CDB: a manually curated database of chromosome conformation capture data

Xiaoxiao Yun[1,2,†], Lili Xia[1,2,†], Bixia Tang[1,2], Hui Zhang[1,2], Feifei Li[1] and Zhihua Zhang[1,*]

[1]CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, 100101, China and [2]University of Chinese Academy of Sciences, Beijing, 100049, China

3D Genome: rVarBase

rVarBase:an updated database for regulatory features of human variants

Chromatin states, regulatory elements and target genes

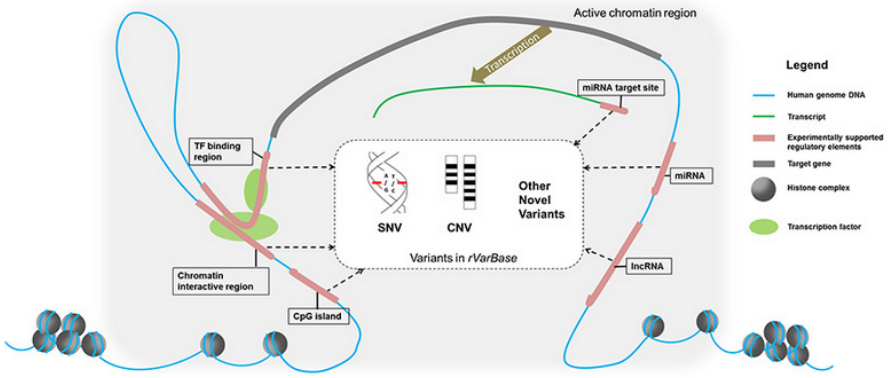
The 2.0 version of rSNPBase

Home | Search | Document | Download | Feedback | About Us

About rVarbase

rVarBase annotates variant's regulatory feature in three fields: chromatin state of the region surrounding variant, regulatory elements overlapped with variant, and variant's potential target genes. It also provides optioned extended annotation for variants, including: LD-proxies of known SNP, SNP/CNV that is overlapped with or located in queried variant, traits (disease and expression quantitative trait) associated with variant. rVarBase is an updated version of the database **rSNPBase**, it is consistent with the old version in utilizing experimentally supported regulatory elements from ENCODE and other data resources to make relevant annotation (such as involved regulatory manner and potential target gene).rVarBase is different from the old version in several new features:

- New variant types
- New dimension of annotation
- New regulatory manner
- More detailed annotation on variant overlapped TFBS
- More extended data
- New search manner



Citation: Guo, L., Du, Y., Qu, S., & Wang, J. (2015). rVarBase: an updated database for regulatory features of human variants. Nucleic acids research, gkv1107 pubmed

Quick Search:

Search here... Search

Input of quick search could be:

- dbSNP/dbVar ID: **rs12345 /nsv7879**
- a single nucleotide as 0-based coordinates: **chr1:12345**
- chromosomal regions: **chr1:12345-34567**

what's new

- New variant types
- New dimension of annotation
- New regulatory manner
- More detailed annotation on variant overlapped TFBS
- More extended data
- New search manner

Quick links

rSNPBase a database for curated regulatory SNPs
Experimental evidences, multiple types of regulation, & rSNP and its LD-proxies

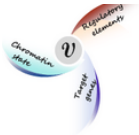
ENCODE: Encyclopedia of DNA Elements

ROADMAP epigenomics PROJECT

rVarBase: an updated database for regulatory features of human variants

Liyuan Guo; Yang Du; Susu Qu; Jing Wang

3D Genome: rVarBase



*r*VarBase:an updated database for regulatory features of human variants

Chromatin states, regulatory elements and target genes

The 2.0 version of rSNPBase

Home | Search | Document | Download | Feedback | About Us

Variant report

Variant	rs4813720
Chromosome Location	chr20:4797409-4797410
allele	A/G
Outlinks	Ensembl UCSC

Chromatin state | **Related regulatory elements** | Target genes | Other information

TF binding region (count:0) | CpG islands (count:0) | **Chromatin interactive region (count:1)** | LncRNA region (count:0) | Mature miRNA region (count: 0) | miRNA target sites (count:0)

(count:1 , 50 per page) page: 1

No.	Distal block	Cell Line	Cell type	Cell Stage
1	chr20:4794840..4796651-chr20:4796825..4799356,2	K562	blood:	

rVarBase: an updated database for regulatory features of human variants

Liyuan Guo; Yang Du; Susu Qu; Jing Wang

3D Genome: CTCFBSDB

CTCFBSDB 2.0: A database for CTCF binding sites and genome organization

Table of Contents

1. Background
2. Sources of binding sites
3. Database access and content
4. In silico CTCFBS prediction tool
5. Recent updates
6. Download
7. References
8. Contact us

1. Background

CCCTC-binding factor (CTCF) is a versatile transcription regulator that is evolutionarily conserved from fruit fly to human. CTCF binds to different DNA sequences through combinatorial use of 11-zinc fingers, and shows distinct functions (transcription activation/repression and chromatin insulation) depending on the biological context [1,2]. Insulators, with the functions of enhancer-blocking and domain-bordering, are critical regulatory elements for gene expression control [3,4]. They represent a class of diverged DNA sequences capable of shielding genes against inappropriate cis-regulatory signals from their genomic neighborhood. Recent studies also linked insulators to epigenetics, such as imprinting [5,6] and X-chromosome inactivation [7]. In eukaryotic genomes, maintenance of distinct chromatin domains is critical for transcription control, and CTCF has been identified as playing a crucial role in the global organization of chromatic architecture [2]. Evidence for this CTCF function has been strengthened by Hi-C experiments that have shown that interacting genomic regions commonly contain CTCF binding sites and that the boundaries of genomic topological domains are enriched for CTCF binding sites [8,9,10]. To analyze this important type of DNA regulatory element, we created a CTCF binding site database (CTCFBSDB), a comprehensive collection of experimentally determined and computationally predicted CTCF binding sites (CTCFBS) from the literature. The database is designed to facilitate the studies on insulators and their roles in demarcating functional genomic domains.

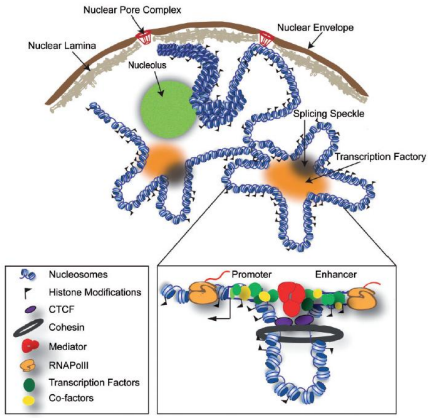


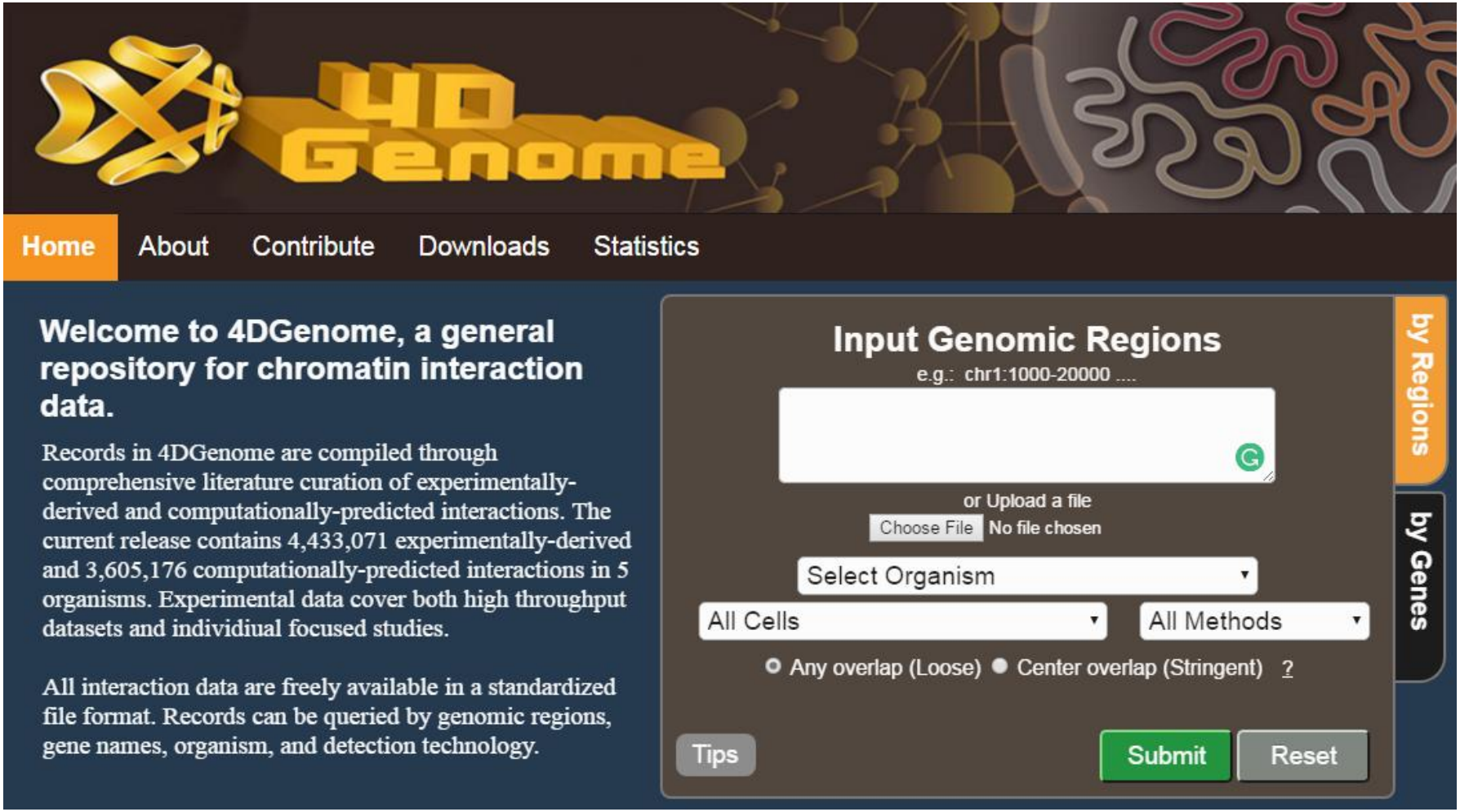
Figure 1. Models of chromatin organization. A diagram of different models of chromatin organization in the nuclear space. Interphase chromatin that interacts with the nuclear lamina (grey), nucleolus (green), nuclear pores (red), transcription factories (orange) and splicing speckles (black) are depicted here. Generally, lamin- and nucleolar-associated domains are transcriptionally repressed and have a more condensed chromatin, whereas chromatin that loops to the nuclear pore, transcription factories and splicing speckles are transcriptionally active and therefore have a more open chromatin structure [here, depicted as 10 nm chromatin fibre]. Enhancers can activate gene expression over a distance and contain binding sites for TFs that recruit co-factors (activators or repressors). A promoter–enhancer looping mechanism mediated by cohesin (brown), CTCF (purple) and the mediator complex (red) that brings the enhancer into close proximity to its target promoter are presented in the enlarged box. The enhancer and promoter are marked with white boxes, and the transcription start site of the transcribed target gene is annotated with an arrow. TFs (green) and co-factors (yellow) bind the enhancer and are brought close to the basal transcription machinery at the promoter. RNAPolII (orange) transcribes pre-mRNA from the target gene and eRNA from the enhancer. Some of these models may co-exist for different PEIs; however, there are also other models that we could not show. A colour version of this figure is available online at BIB online: https://academic.oup.com/bib.

D188–D194 *Nucleic Acids Research, 2013, Vol. 41, Database issue* *Published online 27 November 2012*
doi:10.1093/nar/gks1165

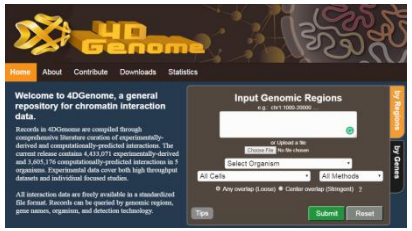
CTCFBSDB 2.0: a database for CTCF-binding sites and genome organization

Jesse D. Ziebarth[1,2], Anindya Bhattacharya[1,2] and Yan Cui[1,2,*]

4D Genome



4D Genome



Download Tables

Full Dataset [884M]

By Organism

Drosophila melanogaster (dm3) [5.4M]

Homo sapiens (hg19) [375M]

Mus musculus (mm9) [505M]

Plasmodium falciparum (3D7) [37M]

Saccharomyces cerevisiae (sacCer3) [26M]

1	InteractorA Chr	Start_hg19	End_hg19	InteractorB Chr	Start_hg19	End_hg19	Agene	Bgene	Cell/Tissue	Detection Method	Confidence Score 1	Confidence Score 2	Contact Frequen	PMID
422	chr6	10315975	10329941	chr20	38750056	38752056	NA	MAFB,ENSG00000204103	Islet	4C	1.00E-10	1.00E-10	NA	24413736
423	chr6	157503620	157517838	chr15	60243703	60245703	NA	CDC24A	Islet	4C	1.00E-10	1.00E-10	NA	24413736
424	chr6	161565011	161582986	chr15	60243703	60245703	AGPAT4,ENSG00000026652	CDC24A	Islet	4C	1.00E-10	1.00E-10	NA	24413736
425	chr6	72459549	72460254	chr15	60243703	60245703	NA	CDC24A	Islet	4C	1.00E-10	1.00E-10	NA	24413736
426	chr6	86476726	86488417	chr15	60243703	60245703	NA	CDC24A	Islet	4C	1.00E-10	1.00E-10	NA	24413736
427	chr6	87749120	87765233	chr15	60243703	60245703	NA	CDC24A	Islet	4C	1.00E-10	1.00E-10	NA	24413736
109004	chr6	113338404	113342157	chr9	100019521	100023609	NA	NA	MCF7	ChIA-PET	9.94E-03	9.94E-03	2	19890323
109063	chr6	116573341	116574450	chr8	98346549	98350824	NA	NA	MCF7	ChIA-PET	1.29E-02	1.29E-02	2	19890323
109450	chr6	129655352	129659267	chr7	11739501	11741816	NA	NA	MCF7	ChIA-PET	6.86E-03	6.86E-03	2	19890323
109483	chr6	13188204	13190556	chr7	119897914	119899333	NA	NA	MCF7	ChIA-PET	1.69E-03	1.69E-03	2	19890323
109555	chr6	133675788	133677357	chrX	113301744	113305058	NA	NA	MCF7	ChIA-PET	2.25E-03	2.25E-03	2	19890323
110072	chr6	140366524	140369650	chr8	87119559	87124371	NA	NA	MCF7	ChIA-PET	7.97E-03	7.97E-03	2	19890323
110217	chr6	144461139	144464271	chr7	25996058	25997037	NA	NA	MCF7	ChIA-PET	9.84E-03	9.84E-03	2	19890323
110218	chr6	144461139	144464271	chr7	71651988	71653553	NA	NA	MCF7	ChIA-PET	1.69E-03	1.69E-03	2	19890323
110414	chr6	149441521	149447514	chr8	95804116	95805941	NA	NA	MCF7	ChIA-PET	1.46E-02	1.46E-02	2	19890323
111610	chr6	170636828	170642065	chr9	116356261	116362118	NA	RGS3,ENSG00000138835	MCF7	ChIA-PET	1.48E-02	1.48E-02	2	19890323
113870	chr6	33971691	33975907	chrX	81952949	81955903	NA	NA	MCF7	ChIA-PET	2.85E-01	2.85E-01	2	19890323
114100	chr6	35281954	35285440	chrX	125368031	125372077	DEF6,ENSG00000023892	NA	MCF7	ChIA-PET	5.20E-03	5.20E-03	2	19890323
114653	chr6	40571645	40576759	chr8	110919605	110924033	NA	NA	MCF7	ChIA-PET	1.42E-02	1.42E-02	2	19890323
115297	chr6	44075370	44079223	chr8	113967345	113970934	NA	NA	MCF7	ChIA-PET	9.94E-03	9.94E-03	2	19890323
115892	chr6	65292602	65298442	chrX	24000947	24006427	NA	NA	MCF7	ChIA-PET	1.26E-02	1.26E-02	2	19890323
115931	chr6	67912721	67914213	chr7	2577564	2579485	NA	NA	MCF7	ChIA-PET	5.75E-03	5.75E-03	2	19890323
115999	chr6	69744408	69747906	chr8	93602364	93609113	NA	NA	MCF7	ChIA-PET	9.94E-03	9.94E-03	2	19890323
116078	chr6	72333637	72336657	chr8	12647790	12650485	NA	NA	MCF7	ChIA-PET	6.08E-04	6.08E-04	2	19890323
116106	chr6	73768705	73775935	chrX	115325824	115333437	NA	NA	MCF7	ChIA-PET	2.24E-02	2.24E-02	2	19890323
116107	chr6	73768705	73775935	chrX	30100645	30102858	NA	NA	MCF7	ChIA-PET	9.94E-03	9.94E-03	2	19890323
116358	chr6	82780446	82787083	chr9	5436074	5439569	NA	PLGRKT,ENSG00000107020	MCF7	ChIA-PET	8.84E-03	8.84E-03	2	19890323
116631	chr6	91110675	91115107	chr7	71592320	71593324	NA	NA	MCF7	ChIA-PET	6.86E-03	6.86E-03	2	19890323
116652	chr6	92449274	92451582	chr8	143446104	143450546	NA	NA	MCF7	ChIA-PET	9.94E-03	9.94E-03	2	19890323

Nature. 2009 Nov 5;462(7269):58-64. doi: 10.1038/nature08497.

An oestrogen-receptor-alpha-bound human chromatin interactome.

Fullwood MJ[1], Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KD, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, Karuturi RK, Herve T, Bourque G, Stunnenberg HG, Ruan X, Cacheux-Rataboul V, Sung WK, Liu ET, Wei CL, Cheung E, Ruan Y.

⊕ Author information

Abstract

Genomes are organized into high-level three-dimensional structures, and DNA elements separated by long genomic distances can in principle interact functionally. Many transcription factors bind to regulatory DNA elements distant from gene promoters. Although distal binding sites have been shown to regulate transcription by long-range chromatin interactions at a few loci, chromatin interactions and their impact on transcription regulation have not been investigated in a genome-wide manner. Here we describe the development of a new strategy, chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) for the de novo detection of global chromatin interactions, with which we have comprehensively mapped the chromatin interaction network bound by oestrogen receptor alpha (ER-alpha) in the human genome. We found that most high-confidence remote ER-alpha-binding sites are anchored at gene promoters through long-range chromatin interactions, suggesting that ER-alpha functions by extensive chromatin looping to bring genes together for coordinated transcriptional regulation. We propose that chromatin interactions constitute a primary mechanism for regulating transcription in mammalian genomes.

Liverpool Hope University EST. 1844

Key Points

- Genetic variation affects chromatin states and gene expression regulation much more than protein structure and function

- DNAse I hypersensitivity regions are enriched for GWAS hits

- There are many ways an eQTL correlates with gene expression levels, including effects on chromatin states, transcription factor / miRNA binding site alteration, RNA splicing and DNA methylation.

- The effects are generally cell type-specific

- Most of these correlations are already catalogued in open access databases.

… Looking forward

Day 1
Genome Biology in 2017
Bioinformatics Tools in Epigenetics
Pathogenicity Assessment of DNA Sequence Mutations
Introduction to Galaxy
Massive Data Sources
Practical

YOUR FUTURE
STARTS WITH HOPE

Liverpool Hope University EST. 1844

YOUR FUTURE
STARTS WITH HOPE

Liverpool Hope University EST. 1844