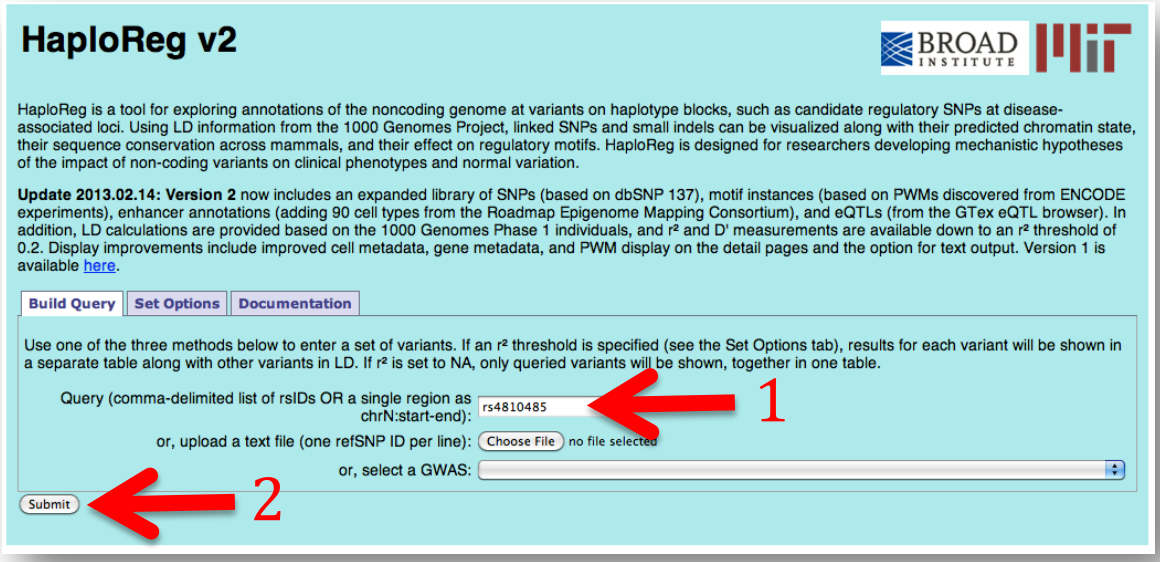
Using HaploReg and RegulomeDB to mine ENCODE data:

(Updated 17 May 2013, Mike Pazin)

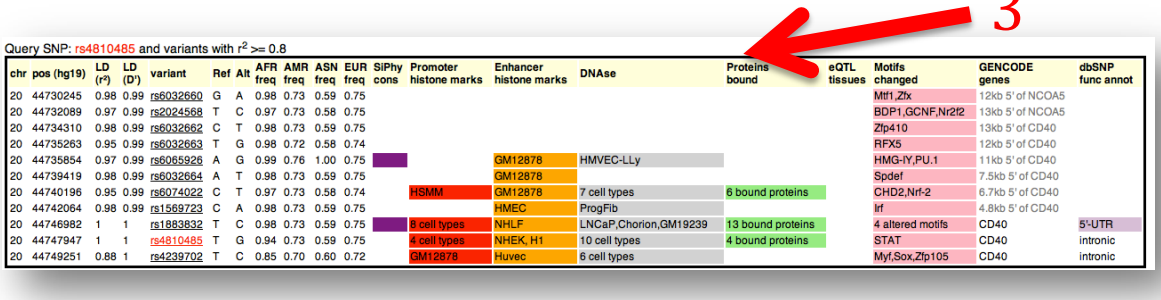
HaploReg and RegulomeDB are ENCODE-funded tools described in recent publications that retrieve ENCODE annotations at SNPs of interest, as well as annotations from work by other researchers and projects.

HaploReg v2:

Go to the HaploReg site, and enter the name of the SNP of interest (Arrow 1).
(Using the "Set Options" tab, the user can configure values such as the LD threshold and the population used from 1000 Genomes data used to calculate LD.)
Click on the submit button (Arrow 2).



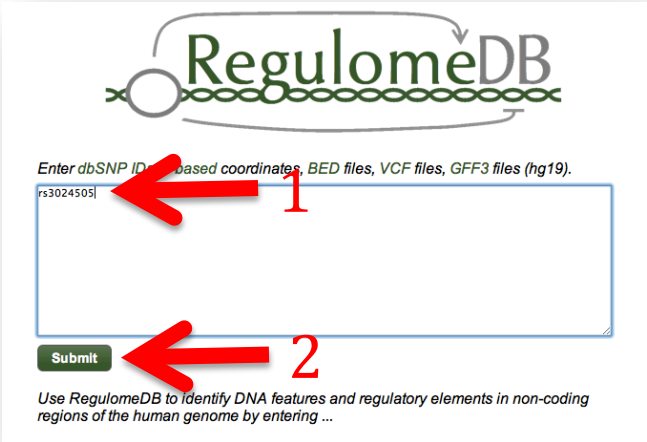
HaploReg retrieves the ENCODE and Roadmap Epigenomics annotations for the selected SNP, as well as other SNPs in LD (arrow 3).



chr	pos (hg19)	LD (r²)	LD (D')	variant	Ref	Alt	AFR freq	AMR freq	ASN freq	EUR freq	SiPhy cons	Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	eQTL tissues	Motifs changed	GENCODE genes	dbSNP func annot
20	44730245	0.98	0.99	rs6032660	G	A	0.98	0.73	0.59	0.75							Mtf1,Zfx	12kb 5' of NCOA5	
20	44732089	0.97	0.99	rs2024568	T	C	0.97	0.73	0.58	0.75							BDP1,GCNF,Nr2f2	13kb 5' of NCOA5	
20	44734310	0.98	0.99	rs6032662	C	T	0.98	0.73	0.59	0.75							Zfp410	13kb 5' of CD40	
20	44735263	0.95	0.99	rs6032663	T	G	0.98	0.72	0.58	0.74							RFX5	12kb 5' of CD40	
20	44735854	0.97	0.99	rs6065926	A	G	0.99	0.76	1.00	0.75			GM12878	HMVEC-LLy			HMG-IY,PU.1	11kb 5' of CD40	
20	44739419	0.98	0.99	rs6032664	A	T	0.98	0.73	0.59	0.75			GM12878				Spdef	7.5kb 5' of CD40	
20	44740196	0.95	0.99	rs6074022	C	T	0.97	0.73	0.58	0.74		HSMM	GM12878	7 cell types	6 bound proteins		CHD2,Nrf-2	6.7kb 5' of CD40	
20	44742064	0.98	0.99	rs1569723	C	A	0.98	0.73	0.59	0.75			HMEC	ProgFib			Irf	4.8kb 5' of CD40	
20	44746982	1	1	rs1883832	T	C	0.98	0.73	0.59	0.75		8 cell types	NHLF	LNCaP,Chorion,GM19239	13 bound proteins		4 altered motifs	CD40	5'-UTR
20	44747947	1	1	rs4810485	T	G	0.94	0.73	0.59	0.75		4 cell types	NHEK, H1	10 cell types	4 bound proteins		STAT	CD40	intronic
20	44749251	0.88	1	rs4239702	T	C	0.85	0.70	0.60	0.72		GM12878	Huvec	6 cell types			Myf,Sox,Zfp105	CD40	intronic

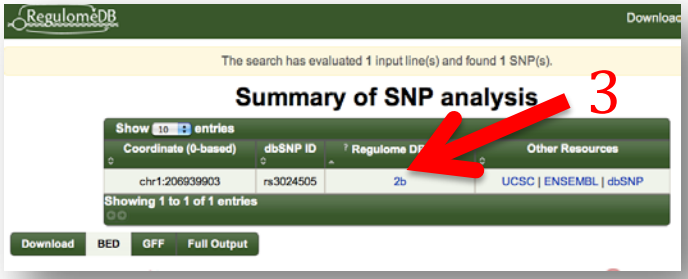
RegulomeDB:

Go to the RegulomeDB site and enter the name of the SNP of interest (Arrow 1).

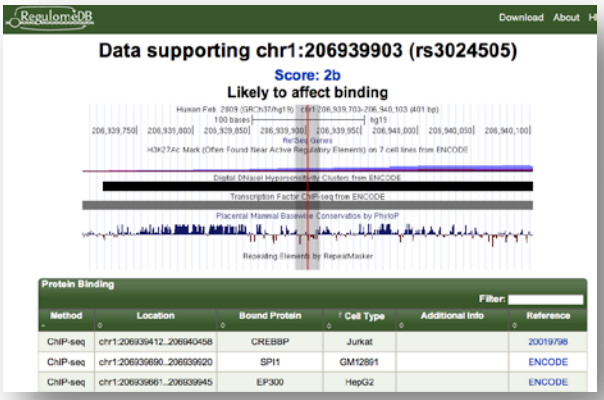


Click on the submit button (Arrow 2).

RegulomeDB calculates a score for the regulatory potential of this region.



Clicking on the score (arrow 3) retrieves the ENCODE (and other) annotation for the region, including transcription factor binding, chromatin structure (DNase, FAIRE, and histone modifications), transcription factor motifs and eQTL.



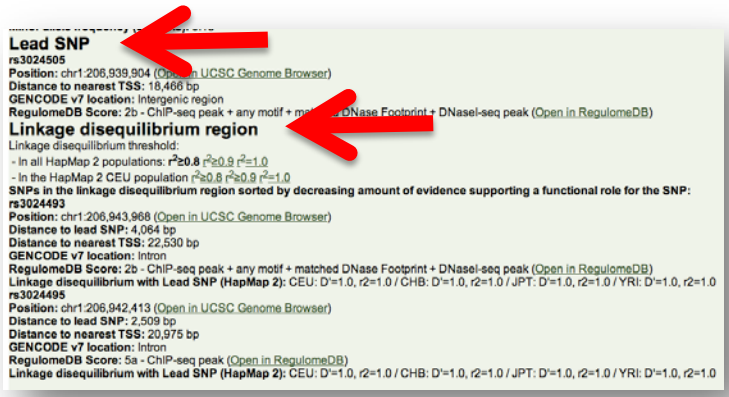
RegulomeDB Disease Association Database, a database of predicted functional SNPs, organized by disease/trait and by SNP, is available at: http://regulome.stanford.edu/GWAS

There is a list of over 4700 SNPs associated with human traits and disease (arrow 1), as well as a list of over 470 human traits and diseases (arrow 2).

Clicking on a trait/disease returns a list of SNPs that have been associated with that trait or disease:

Clicking on a SNP (red arrow) returns the evidence for the association:

As well as the annotation for the lead SNP, and other SNPs in LD that, based on functional annotation, are candidates for the functional variant:



One can follow the links to view the genomic annotation of these SNPs in the genome browser.