Variant Annotation Using HaploReg

Wouter Meuleman

MIT / Broad Institute

Altius Institute (starting July)

Motivation

- The majority of variants reported by GWAS are in noncoding regions of the genome

- Using data from ENCODE, we can annotate noncoding regions of the genome and predict the function of disease associated noncoding variants

- The variant reported by the GWAS (lead/tagged variant) may not be causal but is in high linkage disequilibrium with the causal variant

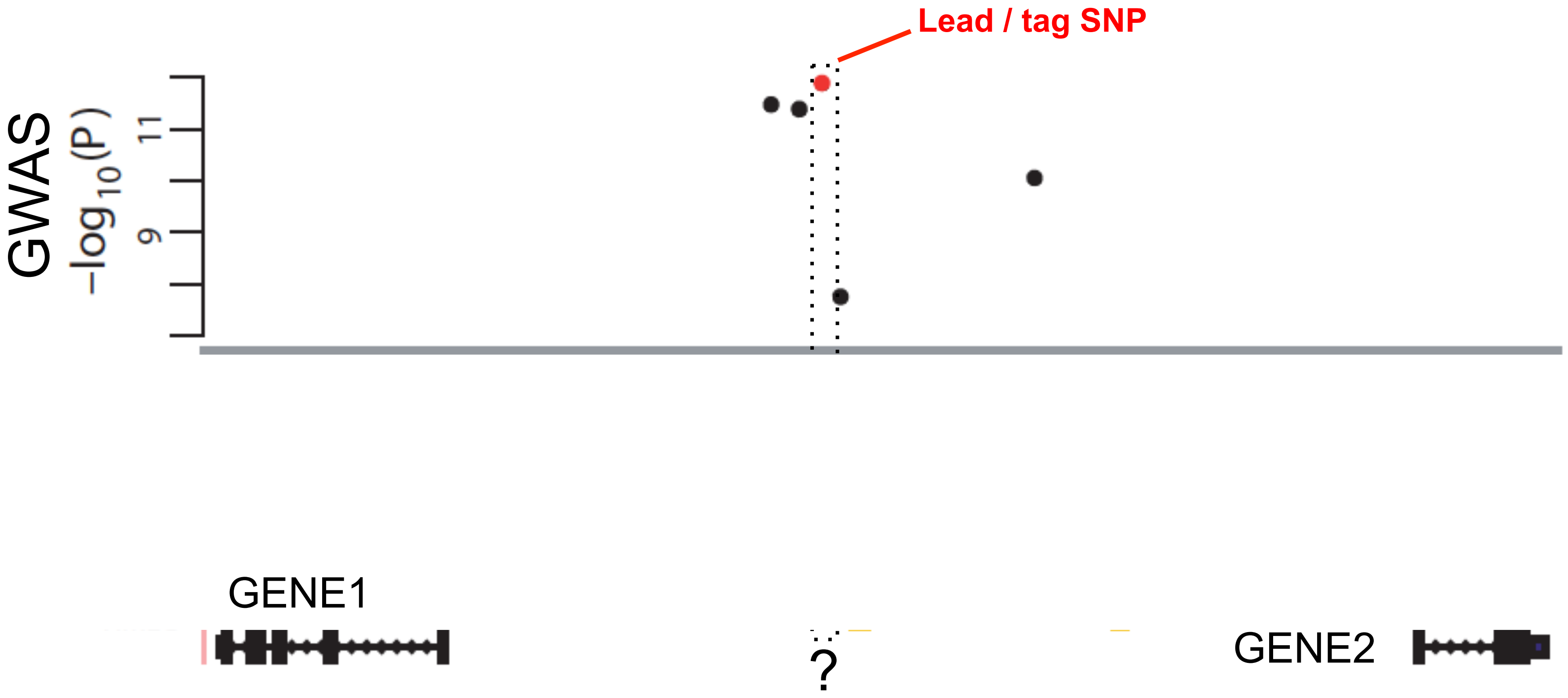
Motivation

- The majority of variants reported by GWAS are in noncoding regions of the genome

- Using data from ENCODE, we can annotate noncoding regions of the genome and predict the function of disease associated noncoding variants

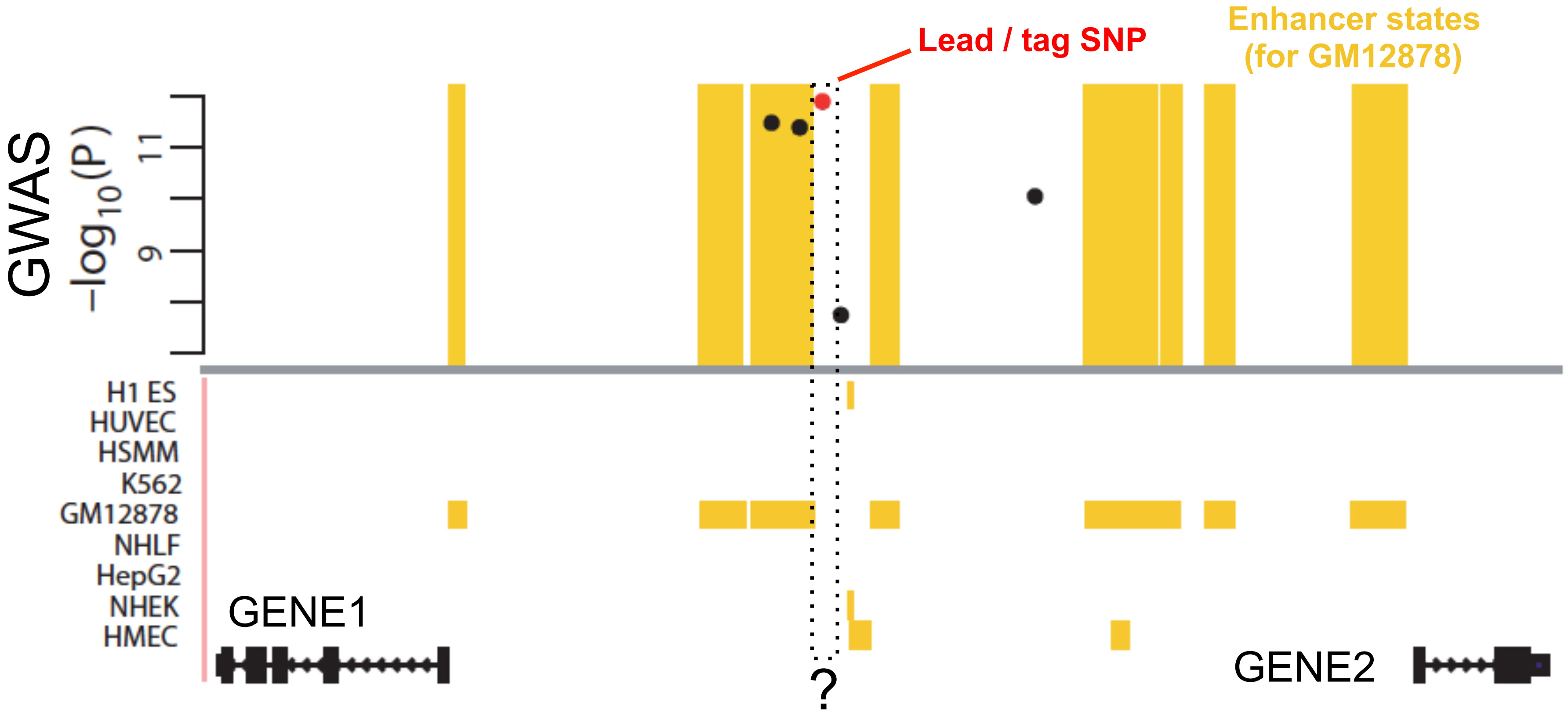
- ***The variant reported by the GWAS (lead/tagged variant) may not be causal but is in high linkage disequilibrium with the causal variant***

Conceptual example – using chromatin states only



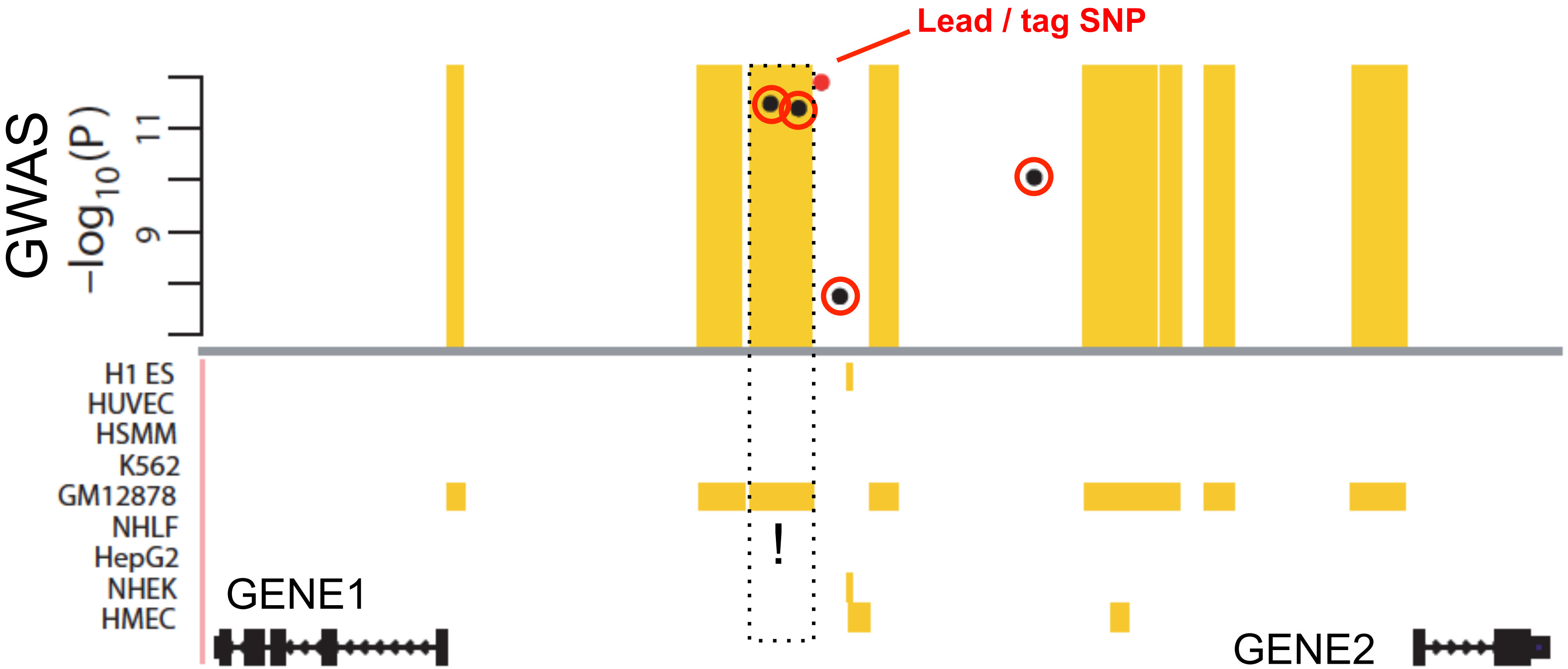
- Highly significant association of **SNP** with a particular trait
- Lack of **chromatin state annotation** hinders interpretation

Conceptual example – using chromatin states only



- Highly significant association of **SNP** with a particular trait
- Lack of **chromatin state annotation** hinders interpretation

Key insight Haploreg: exploit LD-structure



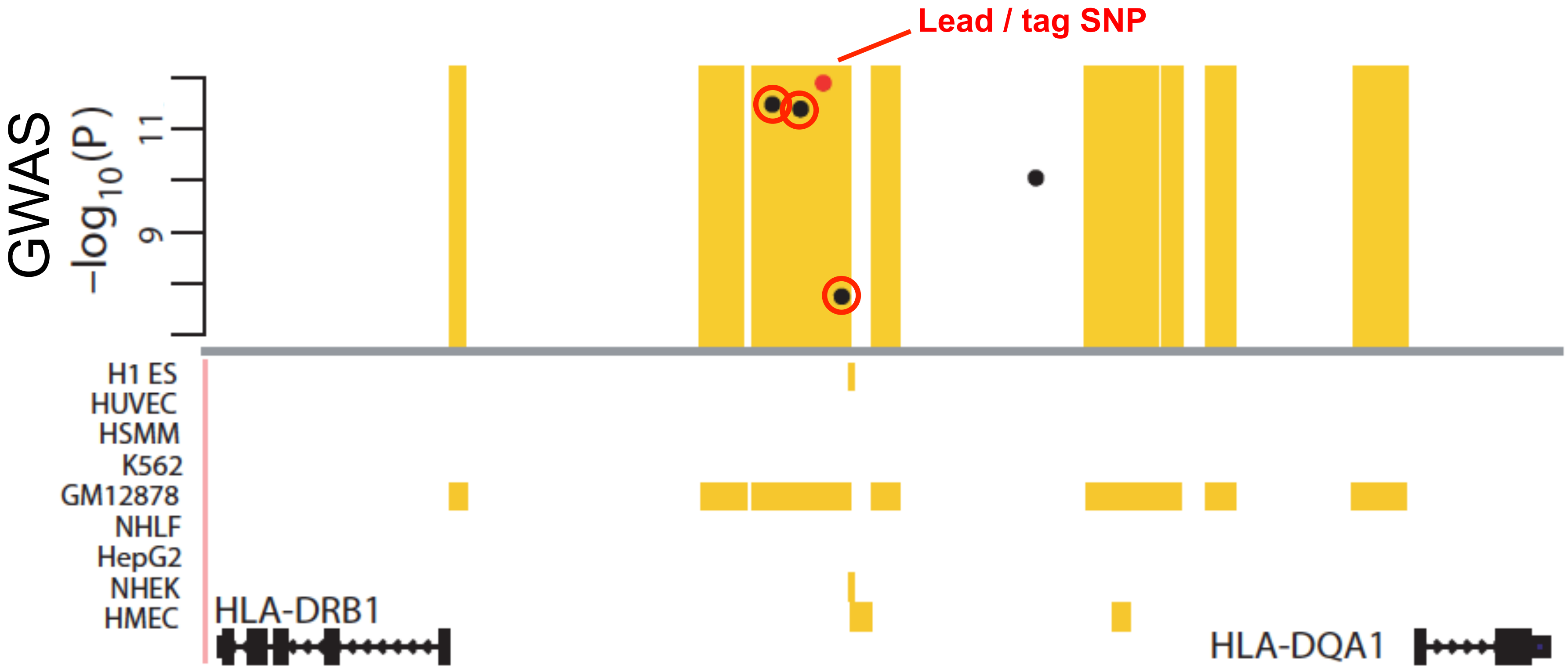
haplotype block

(i.e., SNPs are in strong LD / highly correlated)

● now include these SNPs as well -- guilt by correlation!

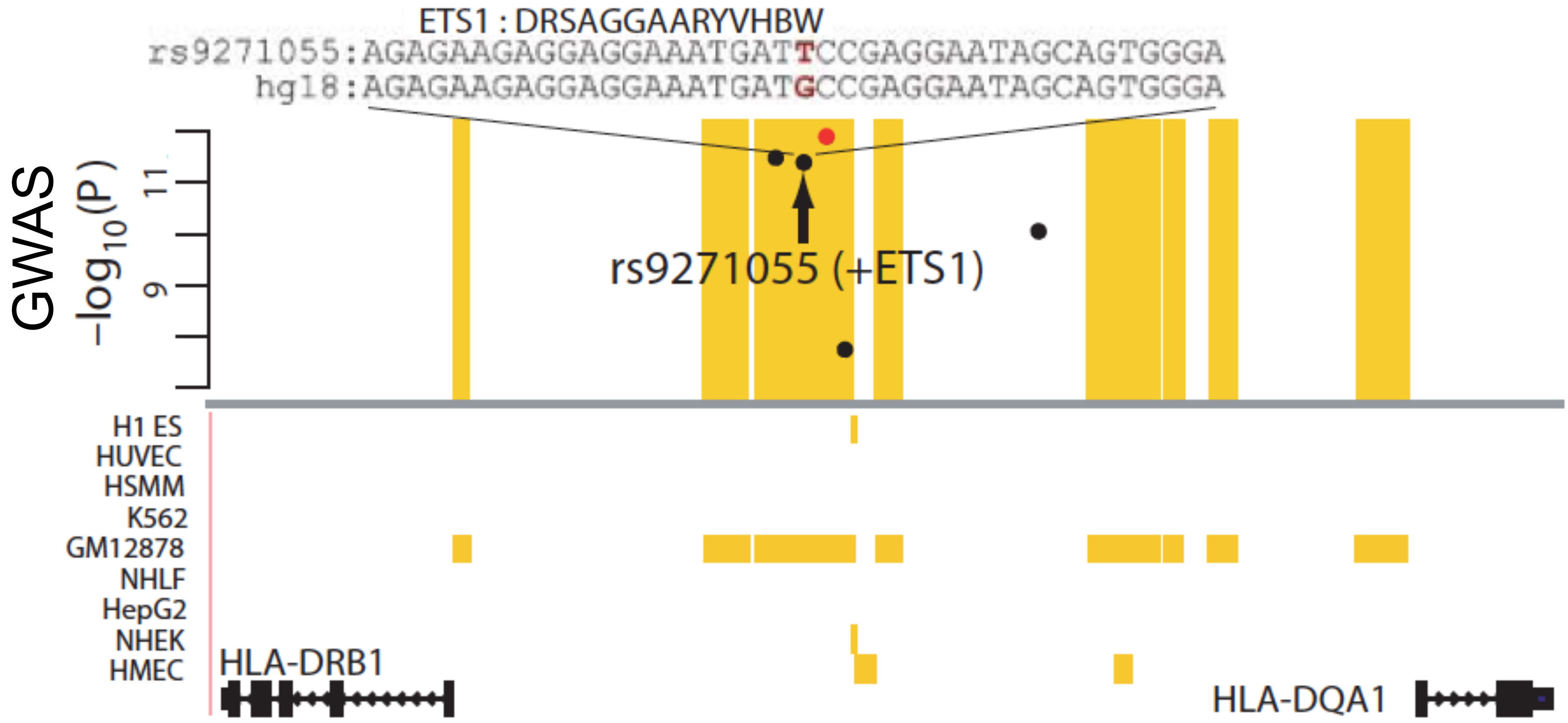
Real example – beyond chromatin states only

(locus associated with systemic lupus erythematosus)



- Here, SNP located in a GM12878-specific enhancer
- But, no further trace of mechanistic explanation
- Solution: also consider other (enhancer) SNPs in LD!

Bingo: LD-SNP strengthens an ETS1 motif



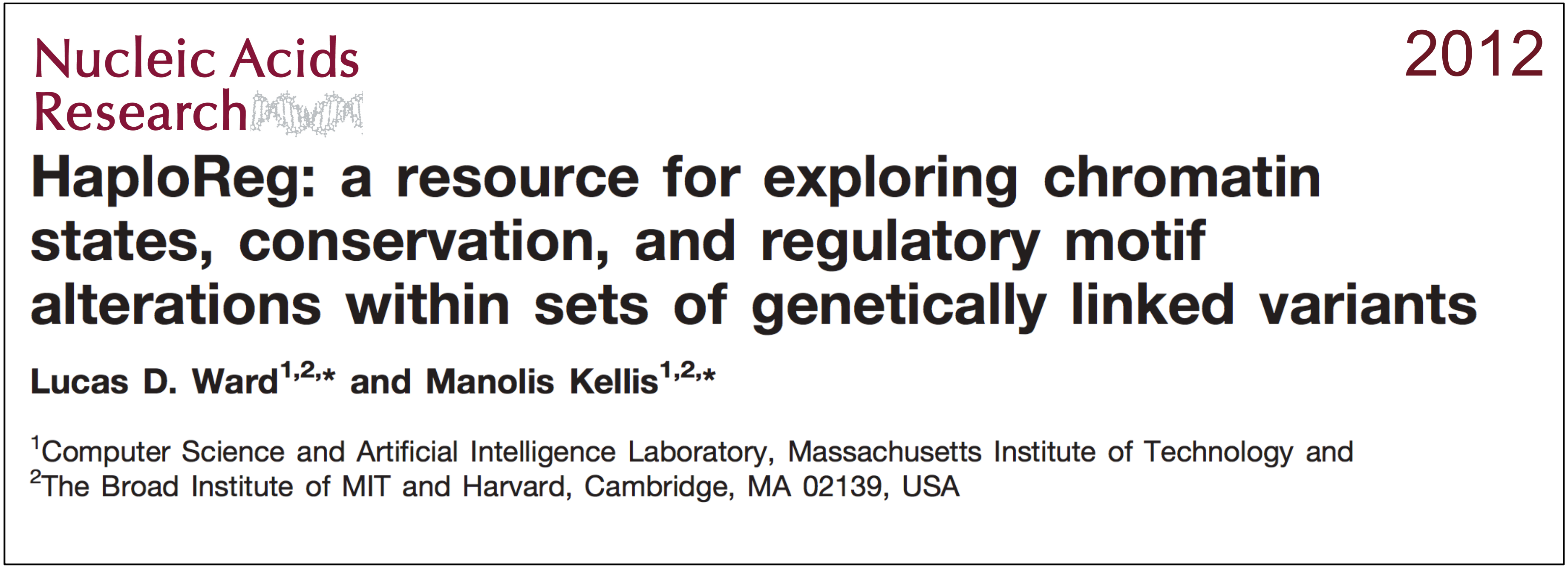
- ETS1 is predicted activator of lymphoblastoid enhancers
- Other lupus-associated variants affecting ETS1 locus

Motivation

- The majority of variants reported by GWAS are in noncoding regions of the genome

- Using data from ENCODE, we can annotate noncoding regions of the genome and predict the function of disease associated noncoding variants

- ***The variant reported by the GWAS (lead/tagged variant) may not be causal but is in high linkage disequilibrium with the causal variant***



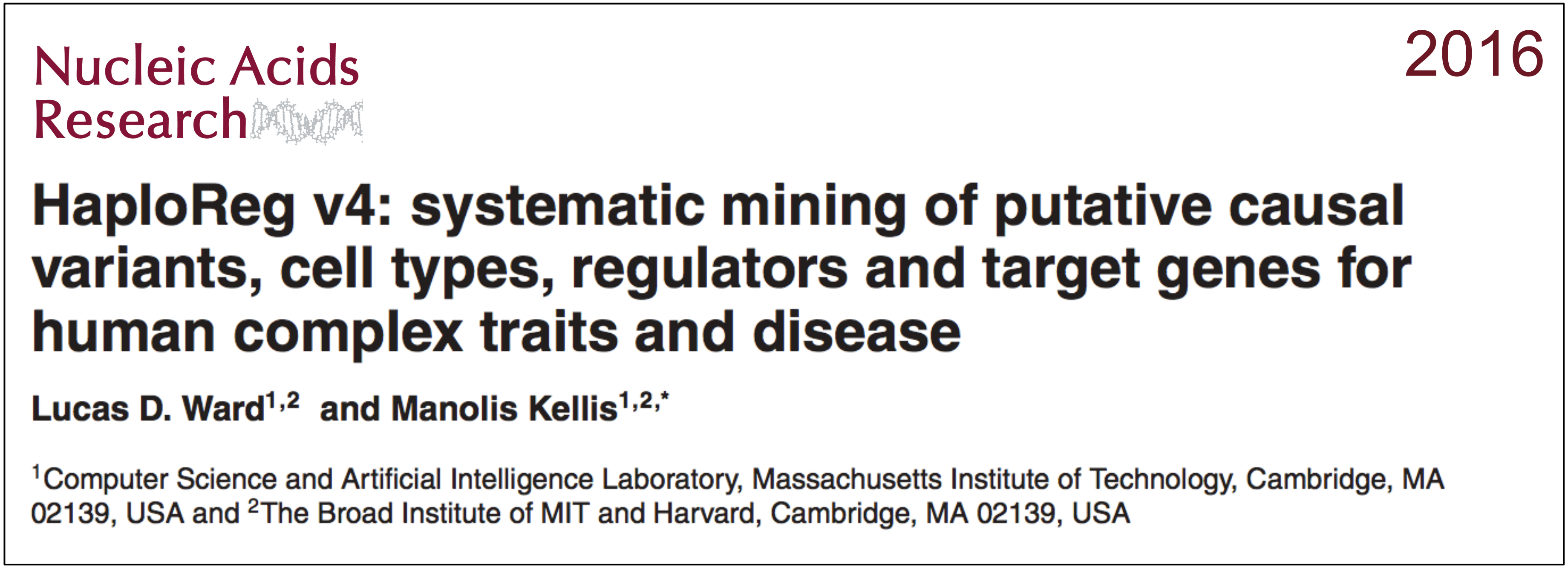
2012

Nucleic Acids Research

HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants

Lucas D. Ward[1,2,*] and Manolis Kellis[1,2,*]

[1]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology and [2]The Broad Institute of MIT and Harvard, Cambridge, MA 02139, USA



2016

Nucleic Acids Research

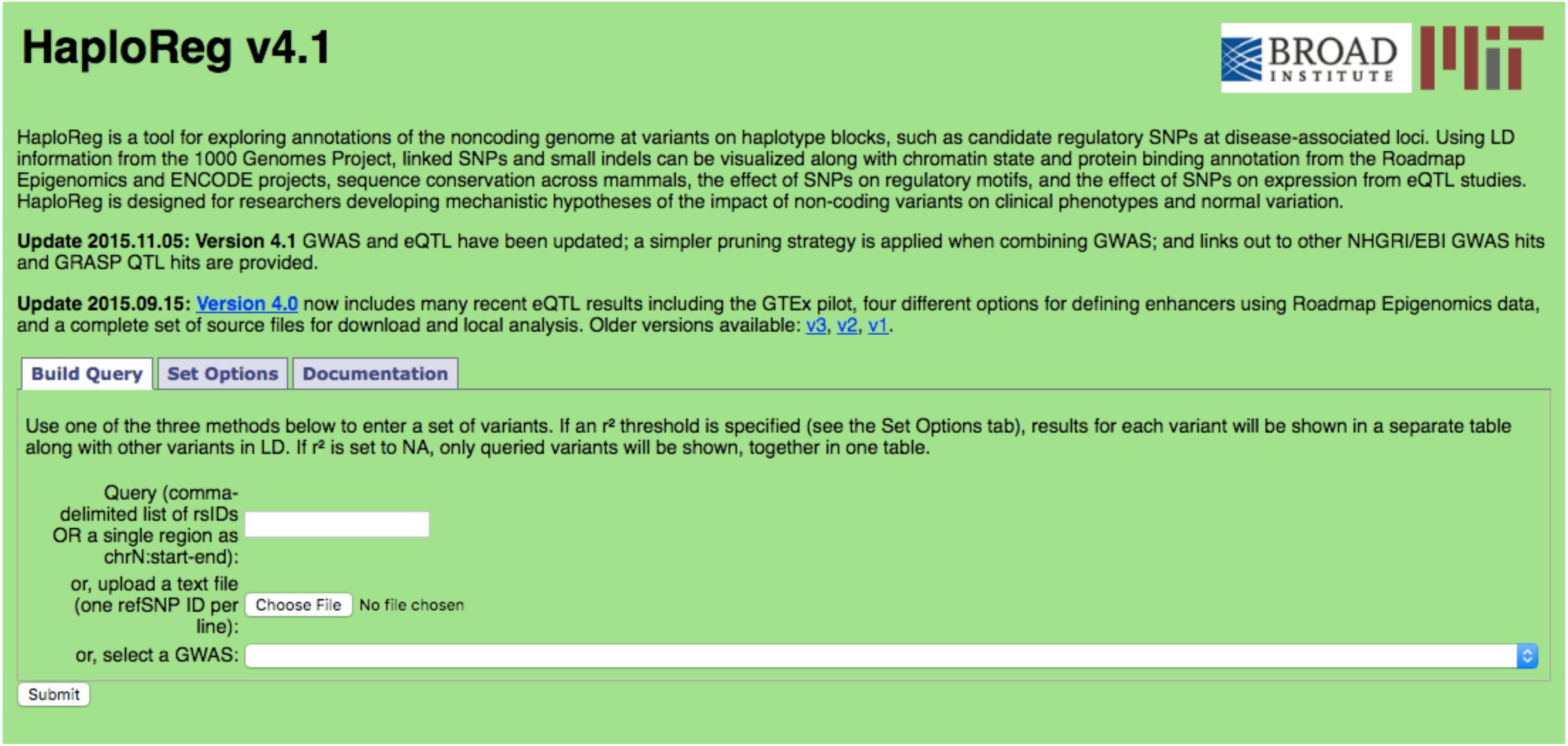
HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease

Lucas D. Ward[1,2] and Manolis Kellis[1,2,*]

[1]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA and [2]The Broad Institute of MIT and Harvard, Cambridge, MA 02139, USA

What’s in HaploReg v4.1 (Updated 5 Nov 2015):

- Roadmap epigenomes (HMM segmentation of histone modification ChIP on 127 tissues/lines; DNase peaks on 53 tissues/lines)
- Regulatory protein binding (ChIP-seq peaks) and regulatory motifs (PWM score change) from ENCODE
- Mammalian-conserved sequence elements (SiPhy and GERP elements – not scores)
- eQTL from GTEx (NIH RNA-seq project on multiple tissues from cadavers), GEUVADIS (EU RNA-seq project + WGS on 1000 genomes LCLs), and 11 other papers
- **See Ward and Kellis (NAR, 2016) for methods and tutorial**



HaploReg v4.1

HaploReg is a tool for exploring annotations of the noncoding genome at variants on haplotype blocks, such as candidate regulatory SNPs at disease-associated loci. Using LD information from the 1000 Genomes Project, linked SNPs and small indels can be visualized along with chromatin state and protein binding annotation from the Roadmap Epigenomics and ENCODE projects, sequence conservation across mammals, the effect of SNPs on regulatory motifs, and the effect of SNPs on expression from eQTL studies. HaploReg is designed for researchers developing mechanistic hypotheses of the impact of non-coding variants on clinical phenotypes and normal variation.

Update 2015.11.05: Version 4.1 GWAS and eQTL have been updated; a simpler pruning strategy is applied when combining GWAS; and links out to other NHGRI/EBI GWAS hits and GRASP QTL hits are provided.

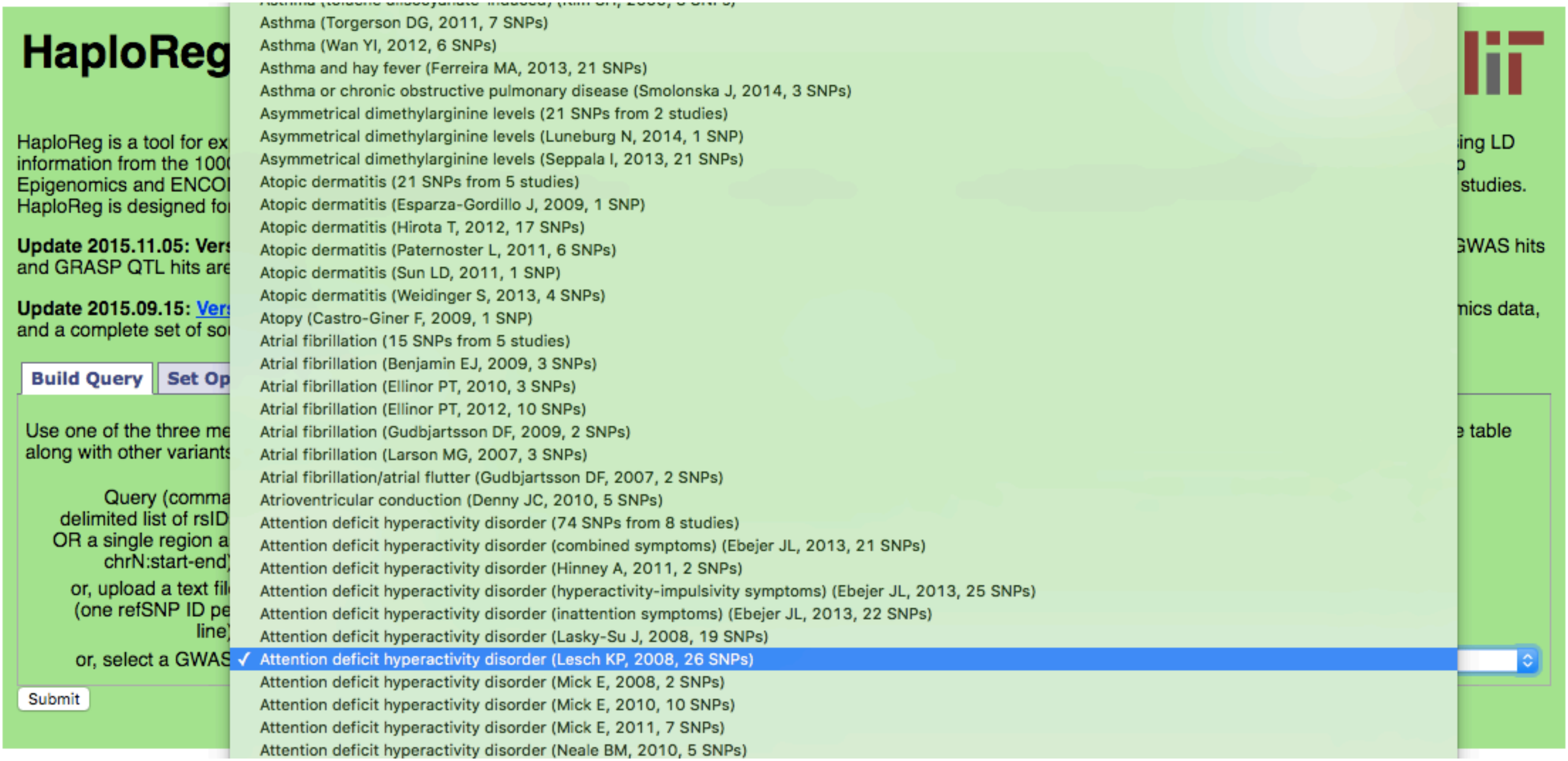
Update 2015.09.15: Version 4.0 now includes many recent eQTL results including the GTEx pilot, four different options for defining enhancers using Roadmap Epigenomics data, and a complete set of source files for download and local analysis. Older versions available: v3, v2, v1.

Use one of the three methods below to enter a set of variants. If an r^2 threshold is specified (see the Set Options tab), results for each variant will be shown in a separate table along with other variants in LD. If r^2 is set to NA, only queried variants will be shown, together in one table.

Query (comma-delimited list of rsIDs OR a single region as chrN:start-end):

or, upload a text file (one refSNP ID per line): Choose File No file chosen

or, select a GWAS:



HaploReg

HaploReg is a tool for ex … ing LD
information from the 100(… o
Epigenomics and ENCO| … studies.
HaploReg is designed fo|

Update 2015.11.05: Vers … GWAS hits
and GRASP QTL hits are

Update 2015.09.15: Vers … nics data,
and a complete set of so|

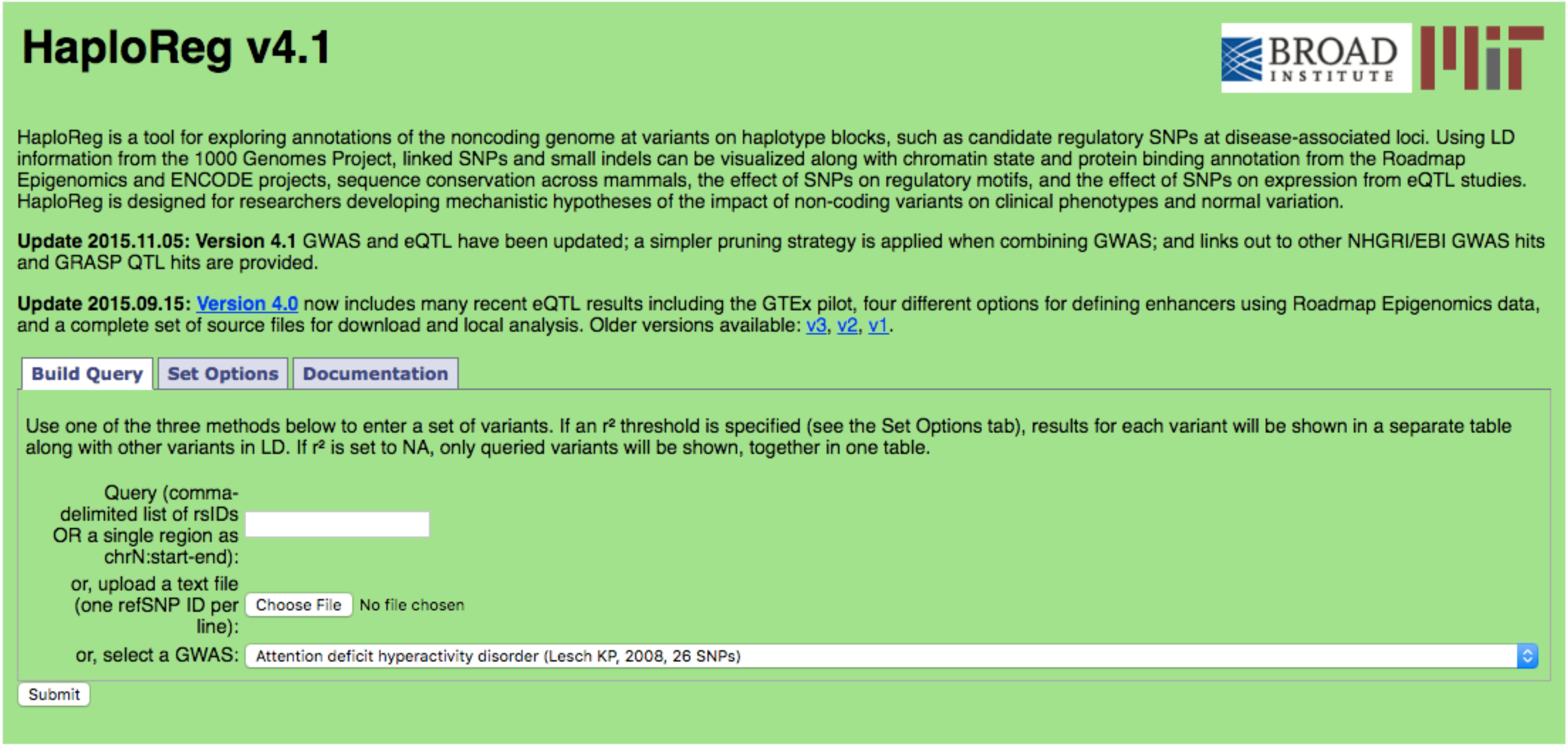
Use one of the three me … e table
along with other variants

Query (comma
delimited list of rsID
OR a single region a
chrN:start-end)

or, upload a text fil
(one refSNP ID pe
line)

or, select a GWAS

Asthma (Torgerson DG, 2011, 7 SNPs)
Asthma (Wan YI, 2012, 6 SNPs)
Asthma and hay fever (Ferreira MA, 2013, 21 SNPs)
Asthma or chronic obstructive pulmonary disease (Smolonska J, 2014, 3 SNPs)
Asymmetrical dimethylarginine levels (21 SNPs from 2 studies)
Asymmetrical dimethylarginine levels (Luneburg N, 2014, 1 SNP)
Asymmetrical dimethylarginine levels (Seppala I, 2013, 21 SNPs)
Atopic dermatitis (21 SNPs from 5 studies)
Atopic dermatitis (Esparza-Gordillo J, 2009, 1 SNP)
Atopic dermatitis (Hirota T, 2012, 17 SNPs)
Atopic dermatitis (Paternoster L, 2011, 6 SNPs)
Atopic dermatitis (Sun LD, 2011, 1 SNP)
Atopic dermatitis (Weidinger S, 2013, 4 SNPs)
Atopy (Castro-Giner F, 2009, 1 SNP)
Atrial fibrillation (15 SNPs from 5 studies)
Atrial fibrillation (Benjamin EJ, 2009, 3 SNPs)
Atrial fibrillation (Ellinor PT, 2010, 3 SNPs)
Atrial fibrillation (Ellinor PT, 2012, 10 SNPs)
Atrial fibrillation (Gudbjartsson DF, 2009, 2 SNPs)
Atrial fibrillation (Larson MG, 2007, 3 SNPs)
Atrial fibrillation/atrial flutter (Gudbjartsson DF, 2007, 2 SNPs)
Atrioventricular conduction (Denny JC, 2010, 5 SNPs)
Attention deficit hyperactivity disorder (74 SNPs from 8 studies)
Attention deficit hyperactivity disorder (combined symptoms) (Ebejer JL, 2013, 21 SNPs)
Attention deficit hyperactivity disorder (Hinney A, 2011, 2 SNPs)
Attention deficit hyperactivity disorder (hyperactivity-impulsivity symptoms) (Ebejer JL, 2013, 25 SNPs)
Attention deficit hyperactivity disorder (inattention symptoms) (Ebejer JL, 2013, 22 SNPs)
Attention deficit hyperactivity disorder (Lasky-Su J, 2008, 19 SNPs)
✓ Attention deficit hyperactivity disorder (Lesch KP, 2008, 26 SNPs)
Attention deficit hyperactivity disorder (Mick E, 2008, 2 SNPs)
Attention deficit hyperactivity disorder (Mick E, 2010, 10 SNPs)
Attention deficit hyperactivity disorder (Mick E, 2011, 7 SNPs)
Attention deficit hyperactivity disorder (Neale BM, 2010, 5 SNPs)



HaploReg v4.1

HaploReg is a tool for exploring annotations of the noncoding genome at variants on haplotype blocks, such as candidate regulatory SNPs at disease-associated loci. Using LD information from the 1000 Genomes Project, linked SNPs and small indels can be visualized along with chromatin state and protein binding annotation from the Roadmap Epigenomics and ENCODE projects, sequence conservation across mammals, the effect of SNPs on regulatory motifs, and the effect of SNPs on expression from eQTL studies. HaploReg is designed for researchers developing mechanistic hypotheses of the impact of non-coding variants on clinical phenotypes and normal variation.

Update 2015.11.05: Version 4.1 GWAS and eQTL have been updated; a simpler pruning strategy is applied when combining GWAS; and links out to other NHGRI/EBI GWAS hits and GRASP QTL hits are provided.

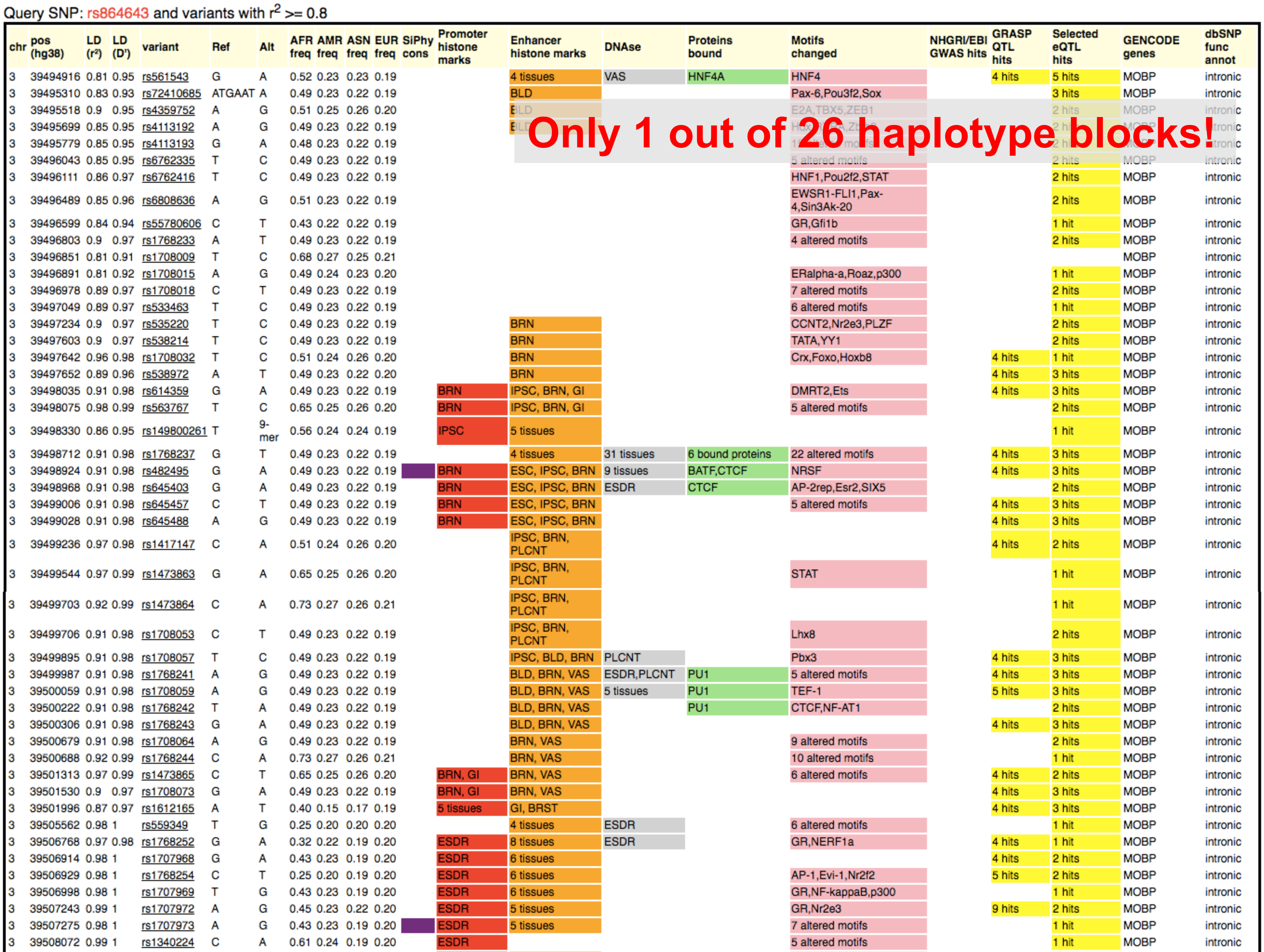
Update 2015.09.15: Version 4.0 now includes many recent eQTL results including the GTEx pilot, four different options for defining enhancers using Roadmap Epigenomics data, and a complete set of source files for download and local analysis. Older versions available: v3, v2, v1.

Use one of the three methods below to enter a set of variants. If an r^2 threshold is specified (see the Set Options tab), results for each variant will be shown in a separate table along with other variants in LD. If r^2 is set to NA, only queried variants will be shown, together in one table.

Query (comma-delimited list of rsIDs OR a single region as chrN:start-end):

or, upload a text file (one refSNP ID per line): No file chosen

or, select a GWAS: Attention deficit hyperactivity disorder (Lesch KP, 2008, 26 SNPs)

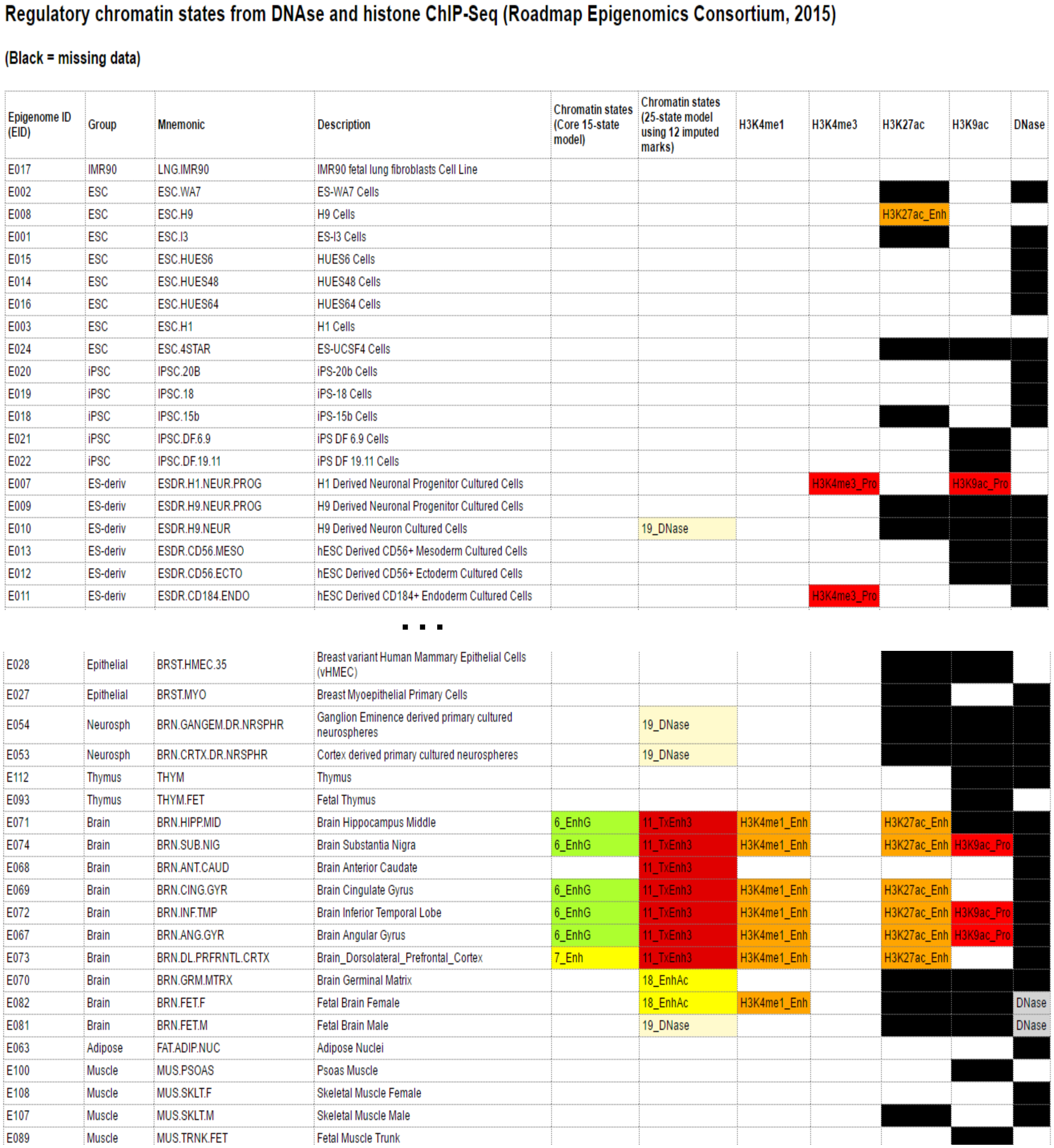


Query SNP: rs864643 and variants with r^2 >= 0.8

chr	pos (hg38)	LD (r^2)	LD (D')	variant	Ref	Alt	AFR freq	AMR freq	ASN freq	EUR freq	SiPhy cons	Promoter histone marks	Enhancer histone marks	DNAse	Proteins bound	Motifs changed	NHGRI/EBI GWAS hits	GRASP QTL hits	Selected eQTL hits	GENCODE genes	dbSNP func annot
3	39494916	0.81	0.95	rs561543	G	A	0.52	0.23	0.23	0.19			4 tissues	VAS	HNF4A	HNF4		4 hits	5 hits	MOBP	intronic
3	39495310	0.83	0.93	rs72410685	ATGAAT	A	0.49	0.23	0.22	0.19			BLD			Pax-6,Pou3f2,Sox			3 hits	MOBP	intronic
3	39495518	0.9	0.95	rs4359752	A	G	0.51	0.25	0.26	0.20			BLD			E2A,TBX5,ZEB1			2 hits	MOBP	intronic
3	39495699	0.85	0.95	rs4113192	A	G	0.49	0.23	0.22	0.19			BLD			Hdx,RXRA,Zbtb3			2 hits	MOBP	intronic
3	39495779	0.85	0.95	rs4113193	G	A	0.48	0.23	0.22	0.19						15 altered motifs			2 hits	MOBP	intronic
3	39496043	0.85	0.95	rs6762335	T	C	0.49	0.23	0.22	0.19						5 altered motifs			2 hits	MOBP	intronic
3	39496111	0.86	0.97	rs6762416	T	C	0.49	0.23	0.22	0.19						HNF1,Pou2f2,STAT			2 hits	MOBP	intronic
3	39496489	0.85	0.96	rs6808636	A	G	0.51	0.23	0.22	0.19						EWSR1-FLI1,Pax-4,Sin3Ak-20			2 hits	MOBP	intronic
3	39496599	0.84	0.94	rs55780606	C	T	0.43	0.22	0.22	0.19						GR,Gfi1b			1 hit	MOBP	intronic
3	39496803	0.9	0.97	rs1768233	A	T	0.49	0.23	0.22	0.19						4 altered motifs			2 hits	MOBP	intronic
3	39496851	0.81	0.91	rs1708009	T	C	0.68	0.27	0.25	0.21										MOBP	intronic
3	39496891	0.81	0.92	rs1708015	A	G	0.49	0.24	0.23	0.20						ERalpha-a,Roaz,p300			1 hit	MOBP	intronic
3	39496978	0.89	0.97	rs1708018	C	T	0.49	0.23	0.22	0.19						7 altered motifs			2 hits	MOBP	intronic
3	39497049	0.89	0.97	rs533463	T	C	0.49	0.23	0.22	0.19						6 altered motifs			1 hit	MOBP	intronic
3	39497234	0.9	0.97	rs535220	T	C	0.49	0.23	0.22	0.19			BRN			CCNT2,Nr2e3,PLZF			2 hits	MOBP	intronic
3	39497603	0.9	0.97	rs538214	T	C	0.49	0.23	0.22	0.19			BRN			TATA,YY1			2 hits	MOBP	intronic
3	39497642	0.96	0.98	rs1708032	T	C	0.51	0.24	0.26	0.20			BRN			Crx,Foxo,Hoxb8		4 hits	1 hit	MOBP	intronic
3	39497652	0.89	0.96	rs538972	A	T	0.49	0.23	0.22	0.20			BRN					4 hits	3 hits	MOBP	intronic
3	39498035	0.91	0.98	rs614359	G	A	0.49	0.23	0.22	0.19		BRN	IPSC, BRN, GI			DMRT2,Ets		4 hits	3 hits	MOBP	intronic
3	39498075	0.98	0.99	rs563767	T	C	0.65	0.25	0.26	0.20		BRN	IPSC, BRN, GI			5 altered motifs			2 hits	MOBP	intronic
3	39498330	0.86	0.95	rs149800261	T	9-mer	0.56	0.24	0.24	0.19		IPSC	5 tissues						1 hit	MOBP	intronic
3	39498712	0.91	0.98	rs1768237	G	T	0.49	0.23	0.22	0.19			4 tissues	31 tissues	6 bound proteins	22 altered motifs		4 hits	3 hits	MOBP	intronic
3	39498924	0.91	0.98	rs482495	G	A	0.49	0.23	0.22	0.19	🟪	BRN	ESC, IPSC, BRN	9 tissues	BATF,CTCF	NRSF		4 hits	3 hits	MOBP	intronic
3	39498968	0.91	0.98	rs645403	G	A	0.49	0.23	0.22	0.19		BRN	ESC, IPSC, BRN	ESDR	CTCF	AP-2rep,Esr2,SIX5			2 hits	MOBP	intronic
3	39499006	0.91	0.98	rs645457	C	T	0.49	0.23	0.22	0.19		BRN	ESC, IPSC, BRN			5 altered motifs		4 hits	3 hits	MOBP	intronic
3	39499028	0.91	0.98	rs645488	A	G	0.49	0.23	0.22	0.19		BRN	ESC, IPSC, BRN					4 hits	3 hits	MOBP	intronic
3	39499236	0.97	0.98	rs1417147	C	A	0.51	0.24	0.26	0.20			IPSC, BRN, PLCNT					4 hits	2 hits	MOBP	intronic
3	39499544	0.97	0.99	rs1473863	G	A	0.65	0.25	0.26	0.20			IPSC, BRN, PLCNT			STAT			1 hit	MOBP	intronic
3	39499703	0.92	0.99	rs1473864	C	A	0.73	0.27	0.26	0.21			IPSC, BRN, PLCNT						1 hit	MOBP	intronic
3	39499706	0.91	0.98	rs1708053	C	T	0.49	0.23	0.22	0.19			IPSC, BRN, PLCNT			Lhx8			2 hits	MOBP	intronic
3	39499895	0.91	0.98	rs1708057	T	C	0.49	0.23	0.22	0.19			IPSC, BLD, BRN	PLCNT		Pbx3		4 hits	3 hits	MOBP	intronic
3	39499987	0.91	0.98	rs1768241	A	G	0.49	0.23	0.22	0.19			BLD, BRN, VAS	ESDR,PLCNT	PU1	5 altered motifs		4 hits	3 hits	MOBP	intronic
3	39500059	0.91	0.98	rs1708059	A	G	0.49	0.23	0.22	0.19			BLD, BRN, VAS	5 tissues	PU1	TEF-1		5 hits	3 hits	MOBP	intronic
3	39500222	0.91	0.98	rs1768242	T	A	0.49	0.23	0.22	0.19			BLD, BRN, VAS		PU1	CTCF,NF-AT1			2 hits	MOBP	intronic
3	39500306	0.91	0.98	rs1768243	G	A	0.49	0.23	0.22	0.19			BLD, BRN, VAS					4 hits	3 hits	MOBP	intronic
3	39500679	0.91	0.98	rs1708064	A	G	0.49	0.23	0.22	0.19			BRN, VAS			9 altered motifs			2 hits	MOBP	intronic
3	39500688	0.92	0.99	rs1768244	C	A	0.73	0.27	0.26	0.21			BRN, VAS			10 altered motifs			1 hit	MOBP	intronic
3	39501313	0.97	0.99	rs1473865	C	T	0.65	0.25	0.26	0.20		BRN, GI	BRN, VAS			6 altered motifs		4 hits	2 hits	MOBP	intronic
3	39501530	0.9	0.97	rs1708073	G	A	0.49	0.23	0.22	0.19		BRN, GI	BRN, VAS					4 hits	3 hits	MOBP	intronic
3	39501996	0.87	0.97	rs1612165	A	T	0.40	0.15	0.17	0.19		5 tissues	GI, BRST					4 hits	3 hits	MOBP	intronic
3	39505562	0.98	1	rs559349	T	G	0.25	0.20	0.20	0.20			4 tissues	ESDR		6 altered motifs			1 hit	MOBP	intronic
3	39506768	0.97	0.98	rs1768252	G	A	0.32	0.22	0.19	0.20		ESDR	8 tissues	ESDR		GR,NERF1a		4 hits	1 hit	MOBP	intronic
3	39506914	0.98	1	rs1707968	G	A	0.43	0.23	0.19	0.20		ESDR	6 tissues					4 hits	2 hits	MOBP	intronic
3	39506929	0.98	1	rs1768254	C	T	0.25	0.20	0.19	0.20		ESDR	6 tissues			AP-1,Evi-1,Nr2f2		5 hits	2 hits	MOBP	intronic
3	39506998	0.98	1	rs1707969	T	G	0.43	0.23	0.19	0.20		ESDR	6 tissues			GR,NF-kappaB,p300			1 hit	MOBP	intronic
3	39507243	0.99	1	rs1707972	A	G	0.45	0.23	0.22	0.20		ESDR	5 tissues			GR,Nr2e3		9 hits	2 hits	MOBP	intronic
3	39507275	0.98	1	rs1707973	A	G	0.43	0.23	0.19	0.20	🟪	ESDR	5 tissues			7 altered motifs			1 hit	MOBP	intronic
3	39508072	0.99	1	rs1340224	C	A	0.61	0.24	0.19	0.20		ESDR				5 altered motifs			1 hit	MOBP	intronic

Only 1 out of 26 haplotype blocks!

Result page for the strongest catalog SNP, rs864643



Regulatory chromatin states from DNAse and histone ChIP-Seq (Roadmap Epigenomics Consortium, 2015)

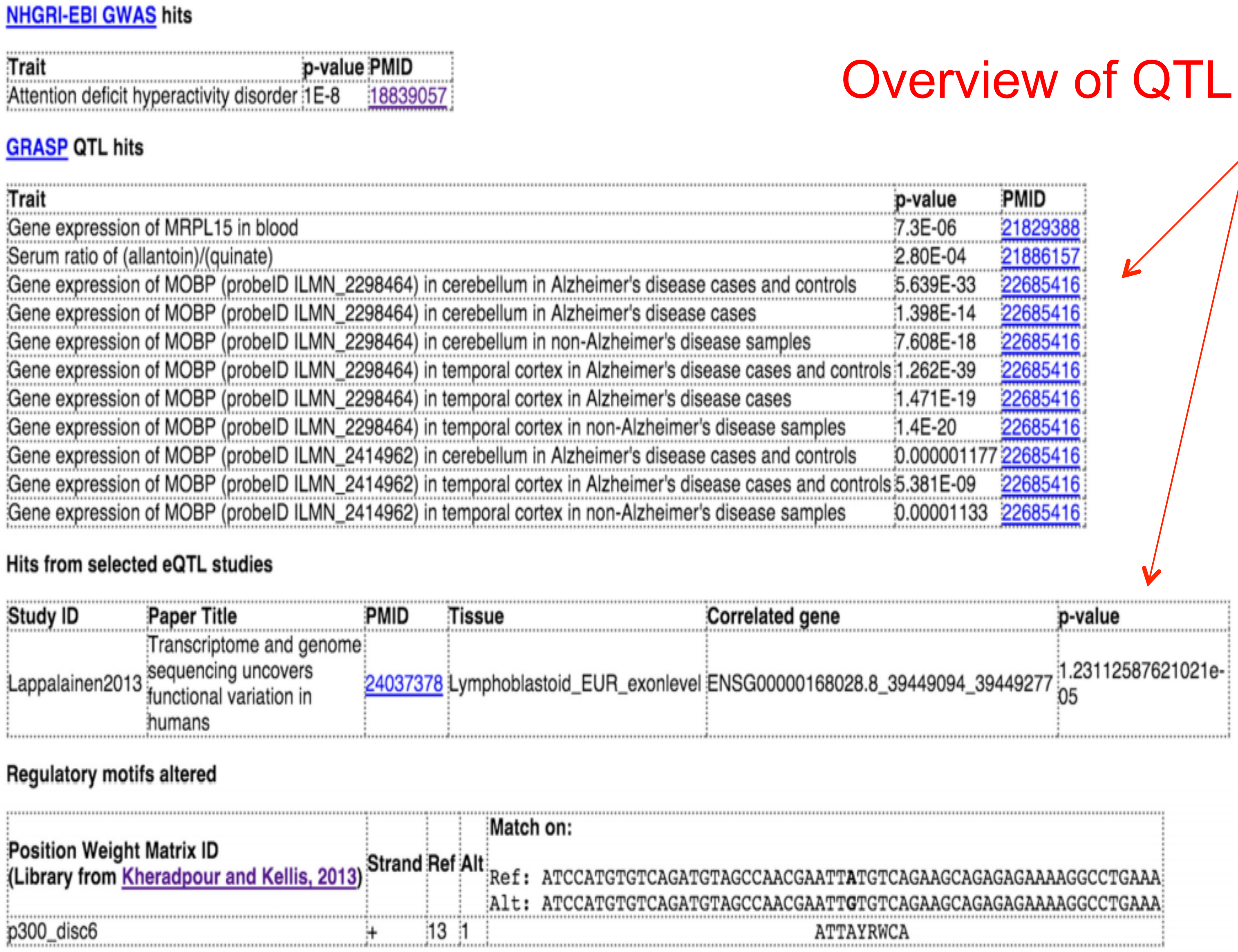
(Black = missing data)

Epigenome ID (EID)	Group	Mnemonic	Description	Chromatin states (Core 15-state model)	Chromatin states (25-state model using 12 imputed marks)	H3K4me1	H3K4me3	H3K27ac	H3K9ac	DNase
E017	IMR90	LNG.IMR90	IMR90 fetal lung fibroblasts Cell Line							
E002	ESC	ESC.WA7	ES-WA7 Cells					■		■
E008	ESC	ESC.H9	H9 Cells					H3K27ac_Enh		
E001	ESC	ESC.I3	ES-I3 Cells					■		■
E015	ESC	ESC.HUES6	HUES6 Cells							■
E014	ESC	ESC.HUES48	HUES48 Cells							■
E016	ESC	ESC.HUES64	HUES64 Cells							■
E003	ESC	ESC.H1	H1 Cells							
E024	ESC	ESC.4STAR	ES-UCSF4 Cells					■	■	■
E020	iPSC	IPSC.20B	iPS-20b Cells							■
E019	iPSC	IPSC.18	iPS-18 Cells							■
E018	iPSC	IPSC.15b	iPS-15b Cells					■		■
E021	iPSC	IPSC.DF.6.9	iPS DF 6.9 Cells						■	
E022	iPSC	IPSC.DF.19.11	iPS DF 19.11 Cells						■	
E007	ES-deriv	ESDR.H1.NEUR.PROG	H1 Derived Neuronal Progenitor Cultured Cells				H3K4me3_Pro		H3K9ac_Pro	
E009	ES-deriv	ESDR.H9.NEUR.PROG	H9 Derived Neuronal Progenitor Cultured Cells					■	■	■
E010	ES-deriv	ESDR.H9.NEUR	H9 Derived Neuron Cultured Cells		19_DNase			■	■	■
E013	ES-deriv	ESDR.CD56.MESO	hESC Derived CD56+ Mesoderm Cultured Cells						■	■
E012	ES-deriv	ESDR.CD56.ECTO	hESC Derived CD56+ Ectoderm Cultured Cells						■	■
E011	ES-deriv	ESDR.CD184.ENDO	hESC Derived CD184+ Endoderm Cultured Cells				H3K4me3_Pro			■

. . .

E028	Epithelial	BRST.HMEC.35	Breast variant Human Mammary Epithelial Cells (vHMEC)					■	■	
E027	Epithelial	BRST.MYO	Breast Myoepithelial Primary Cells					■		■
E054	Neurosph	BRN.GANGEM.DR.NRSPHR	Ganglion Eminence derived primary cultured neurospheres		19_DNase			■	■	■
E053	Neurosph	BRN.CRTX.DR.NRSPHR	Cortex derived primary cultured neurospheres		19_DNase			■	■	■
E112	Thymus	THYM	Thymus						■	■
E093	Thymus	THYM.FET	Fetal Thymus						■	
E071	Brain	BRN.HIPP.MID	Brain Hippocampus Middle	6_EnhG	11_TxEnh3	H3K4me1_Enh		H3K27ac_Enh	■	■
E074	Brain	BRN.SUB.NIG	Brain Substantia Nigra	6_EnhG	11_TxEnh3	H3K4me1_Enh		H3K27ac_Enh	H3K9ac_Pro	■
E068	Brain	BRN.ANT.CAUD	Brain Anterior Caudate		11_TxEnh3					■
E069	Brain	BRN.CING.GYR	Brain Cingulate Gyrus	6_EnhG	11_TxEnh3	H3K4me1_Enh		H3K27ac_Enh		■
E072	Brain	BRN.INF.TMP	Brain Inferior Temporal Lobe	6_EnhG	11_TxEnh3	H3K4me1_Enh		H3K27ac_Enh	H3K9ac_Pro	■
E067	Brain	BRN.ANG.GYR	Brain Angular Gyrus	6_EnhG	11_TxEnh3	H3K4me1_Enh		H3K27ac_Enh	H3K9ac_Pro	■
E073	Brain	BRN.DL.PRFRNTL.CRTX	Brain_Dorsolateral_Prefrontal_Cortex	7_Enh	11_TxEnh3	H3K4me1_Enh		H3K27ac_Enh		■
E070	Brain	BRN.GRM.MTRX	Brain Germinal Matrix		18_EnhAc			■	■	■
E082	Brain	BRN.FET.F	Fetal Brain Female		18_EnhAc	H3K4me1_Enh		■	■	DNase
E081	Brain	BRN.FET.M	Fetal Brain Male		19_DNase			■	■	DNase
E063	Adipose	FAT.ADIP.NUC	Adipose Nuclei							■
E100	Muscle	MUS.PSOAS	Psoas Muscle						■	
E108	Muscle	MUS.SKLT.F	Skeletal Muscle Female							■
E107	Muscle	MUS.SKLT.M	Skeletal Muscle Male					■		■
E089	Muscle	MUS.TRNK.FET	Fetal Muscle Trunk						■	

Brain-specific enhancer



NHGRI-EBI GWAS hits

Trait	p-value	PMID
Attention deficit hyperactivity disorder	1E-8	18839057

GRASP QTL hits

Trait	p-value	PMID
Gene expression of MRPL15 in blood	7.3E-06	21829388
Serum ratio of (allantoin)/(quinate)	2.80E-04	21886157
Gene expression of MOBP (probeID ILMN_2298464) in cerebellum in Alzheimer's disease cases and controls	5.639E-33	22685416
Gene expression of MOBP (probeID ILMN_2298464) in cerebellum in Alzheimer's disease cases	1.398E-14	22685416
Gene expression of MOBP (probeID ILMN_2298464) in cerebellum in non-Alzheimer's disease samples	7.608E-18	22685416
Gene expression of MOBP (probeID ILMN_2298464) in temporal cortex in Alzheimer's disease cases and controls	1.262E-39	22685416
Gene expression of MOBP (probeID ILMN_2298464) in temporal cortex in Alzheimer's disease cases	1.471E-19	22685416
Gene expression of MOBP (probeID ILMN_2298464) in temporal cortex in non-Alzheimer's disease samples	1.4E-20	22685416
Gene expression of MOBP (probeID ILMN_2414962) in cerebellum in Alzheimer's disease cases and controls	0.000001177	22685416
Gene expression of MOBP (probeID ILMN_2414962) in temporal cortex in Alzheimer's disease cases and controls	5.381E-09	22685416
Gene expression of MOBP (probeID ILMN_2414962) in temporal cortex in non-Alzheimer's disease samples	0.00001133	22685416

Hits from selected eQTL studies

Study ID	Paper Title	PMID	Tissue	Correlated gene	p-value
Lappalainen2013	Transcriptome and genome sequencing uncovers functional variation in humans	24037378	Lymphoblastoid_EUR_exonlevel	ENSG00000168028.8_39449094_39449277	1.23112587621021e-05

Regulatory motifs altered

Position Weight Matrix ID (Library from Kheradpour and Kellis, 2013)	Strand	Ref	Alt	Match on: Ref: ATCCATGTGTCAGATGTAGCCAACGAATT**A**TGTCAGAAGCAGAGAGAAAAGGCCTGAAA Alt: ATCCATGTGTCAGATGTAGCCAACGAATT**G**TGTCAGAAGCAGAGAGAAAAGGCCTGAAA
p300_disc6	+	13	1	ATTAYRWCA

Overview of QTL study hits

Dramatically altered p300 binding

Nucleic Acids Research

2016

HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease

Lucas D. Ward[1,2] and Manolis Kellis[1,2,*]

[1] Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA and [2] The Broad Institute of MIT and Harvard, Cambridge, MA 02139, USA

http://compbio.mit.edu/haploreg

Many thanks to Jill Moore and Luke Ward for help with slides!