

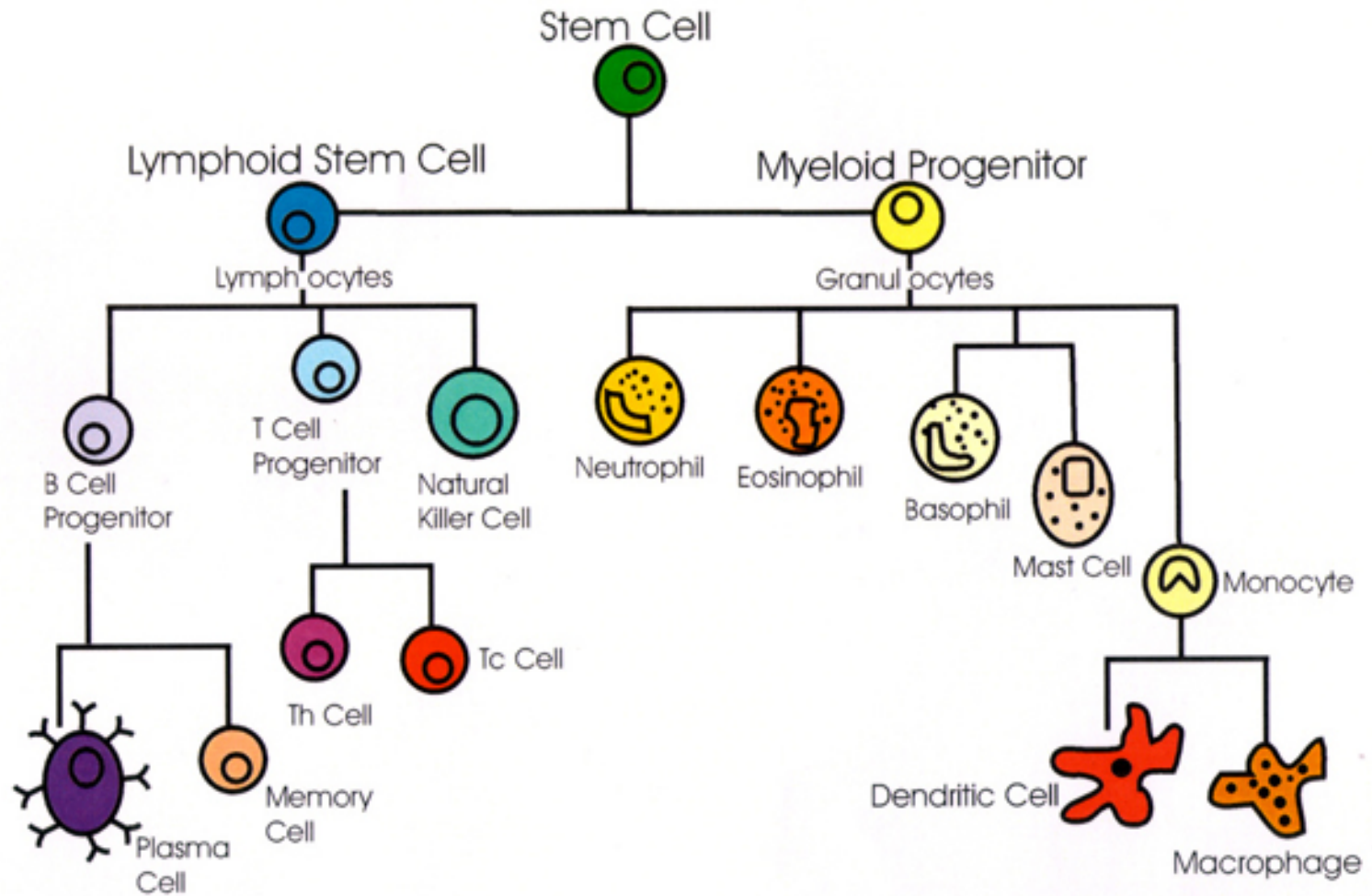
Identifying Dysregulated Genes in Autoimmune Disease

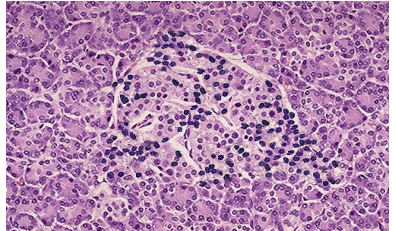
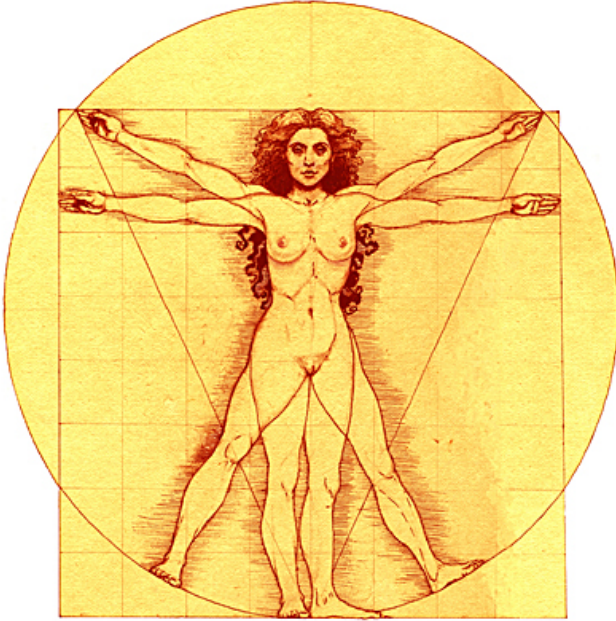
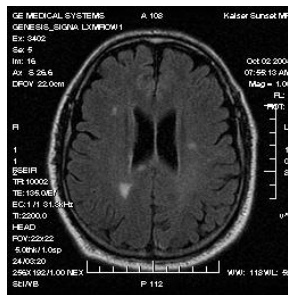
Chris Cotsapas PhD
Yale Neurology/Genetics
Broad Institute
cotsapas@broadinstitute.org

Causal
Identifying ~~Dysregulated~~ Genes
in Autoimmune Disease

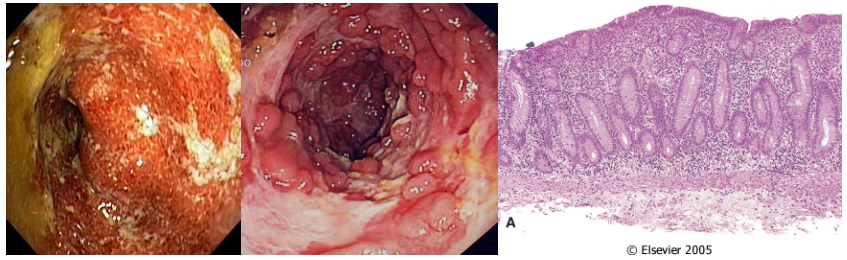
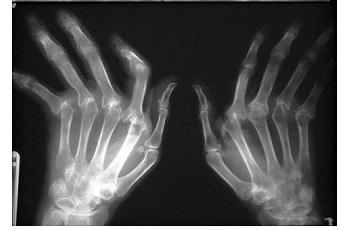
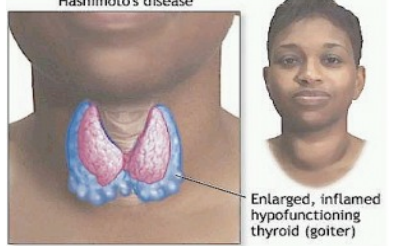
Chris Cotsapas PhD
Yale Neurology/Genetics
Broad Institute
cotsapas@broadinstitute.org

Cells of the Immune System



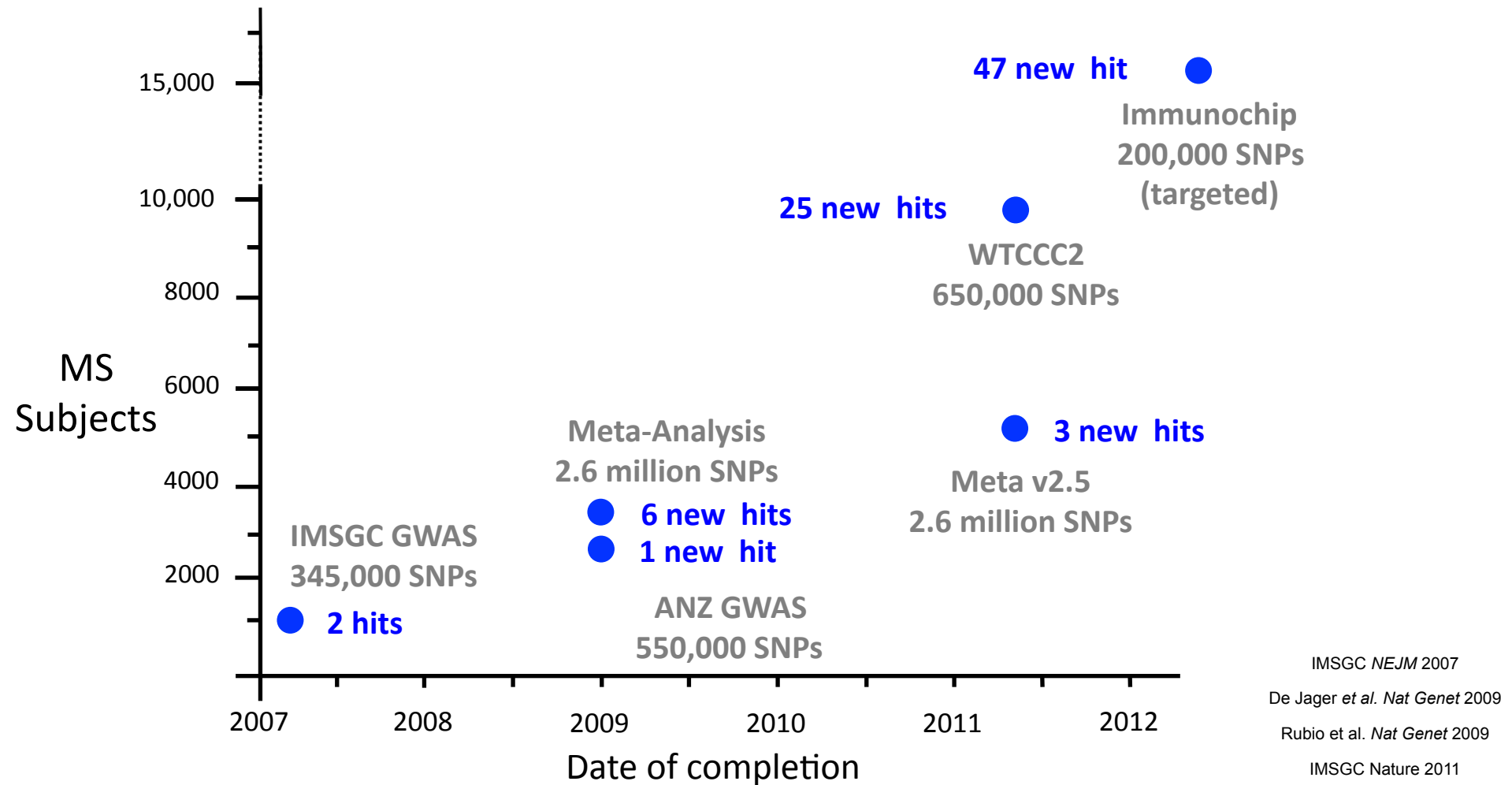


Hashimoto's disease



© Elsevier 2005

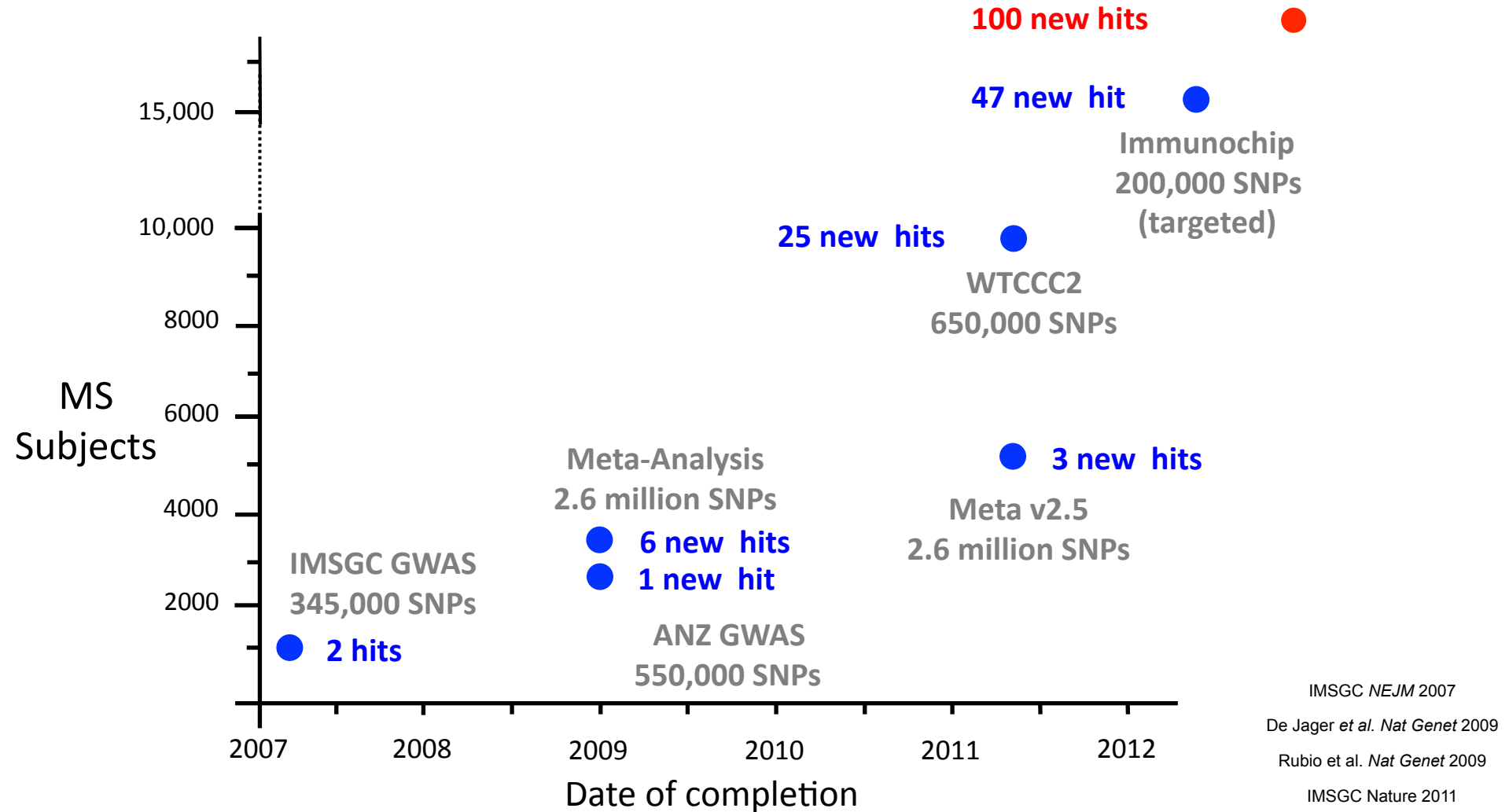
Multiple sclerosis GWAS



... and that's not all!

Meta-Analysis v3.0
16K MS cases / 26K controls

Replication
18K MS cases / 18K controls



GWAS signals are enriched in regulatory DNA

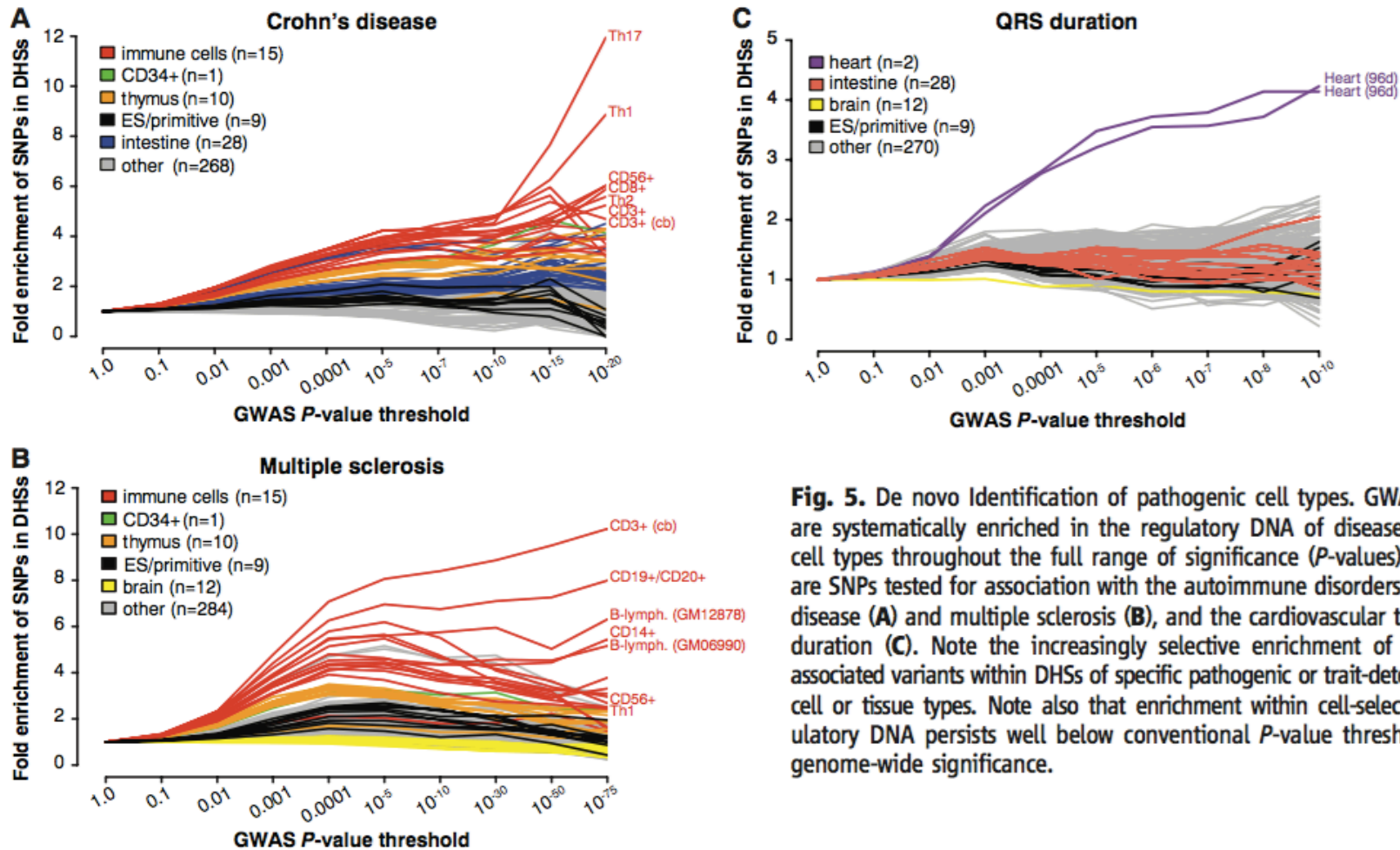
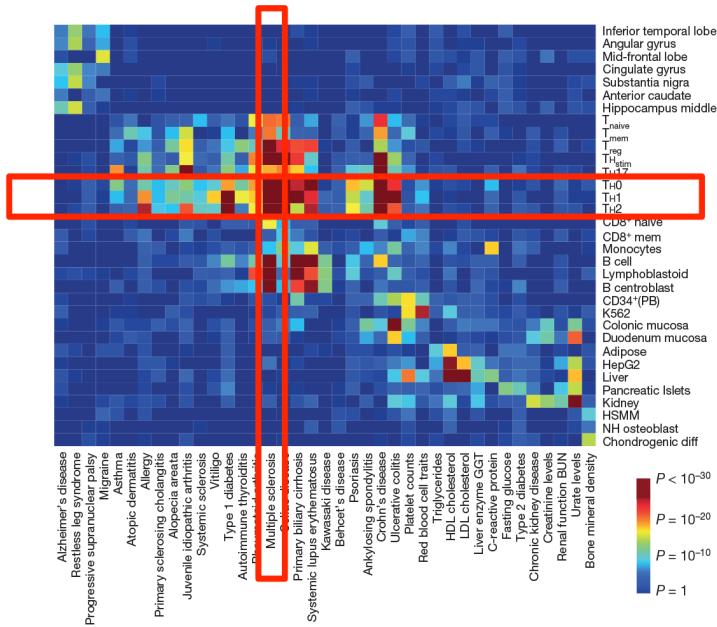


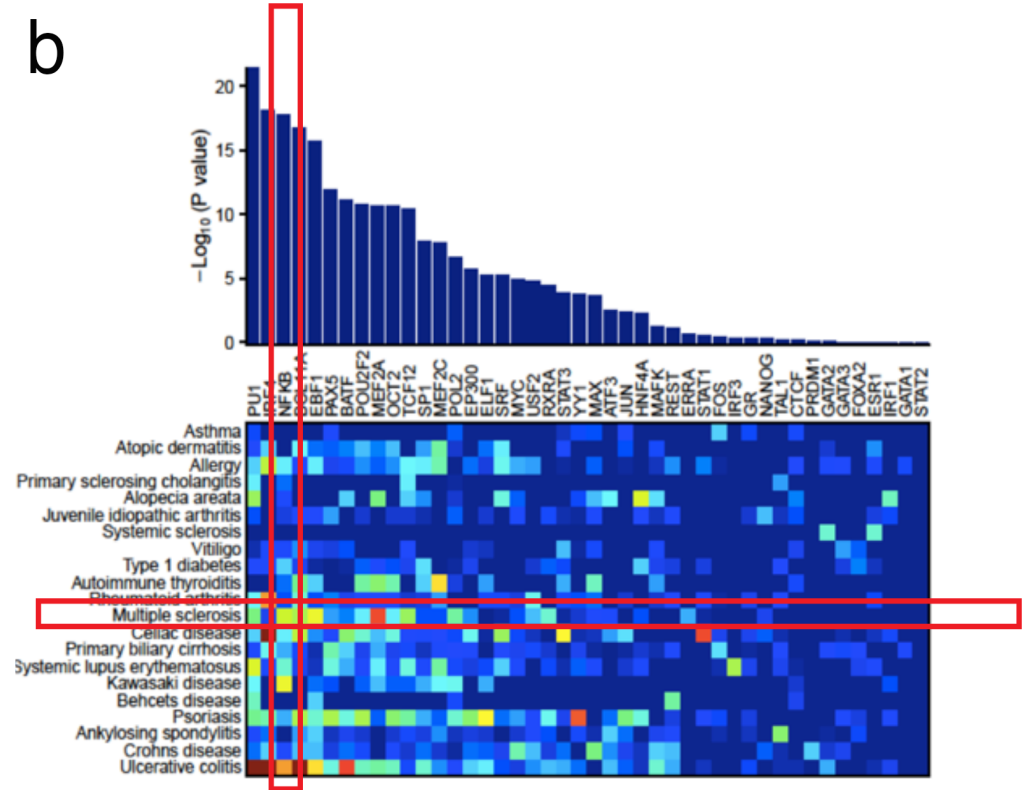
Fig. 5. De novo Identification of pathogenic cell types. GWAS SNPs are systematically enriched in the regulatory DNA of disease-specific cell types throughout the full range of significance (P -values). Shown are SNPs tested for association with the autoimmune disorders Crohn's disease (A) and multiple sclerosis (B), and the cardiovascular trait QRS duration (C). Note the increasingly selective enrichment of disease-associated variants within DHSs of specific pathogenic or trait-determining cell or tissue types. Note also that enrichment within cell-selective regulatory DNA persists well below conventional P -value thresholds for genome-wide significance.

MS GWAS hits enriched in transcription factor binding sites

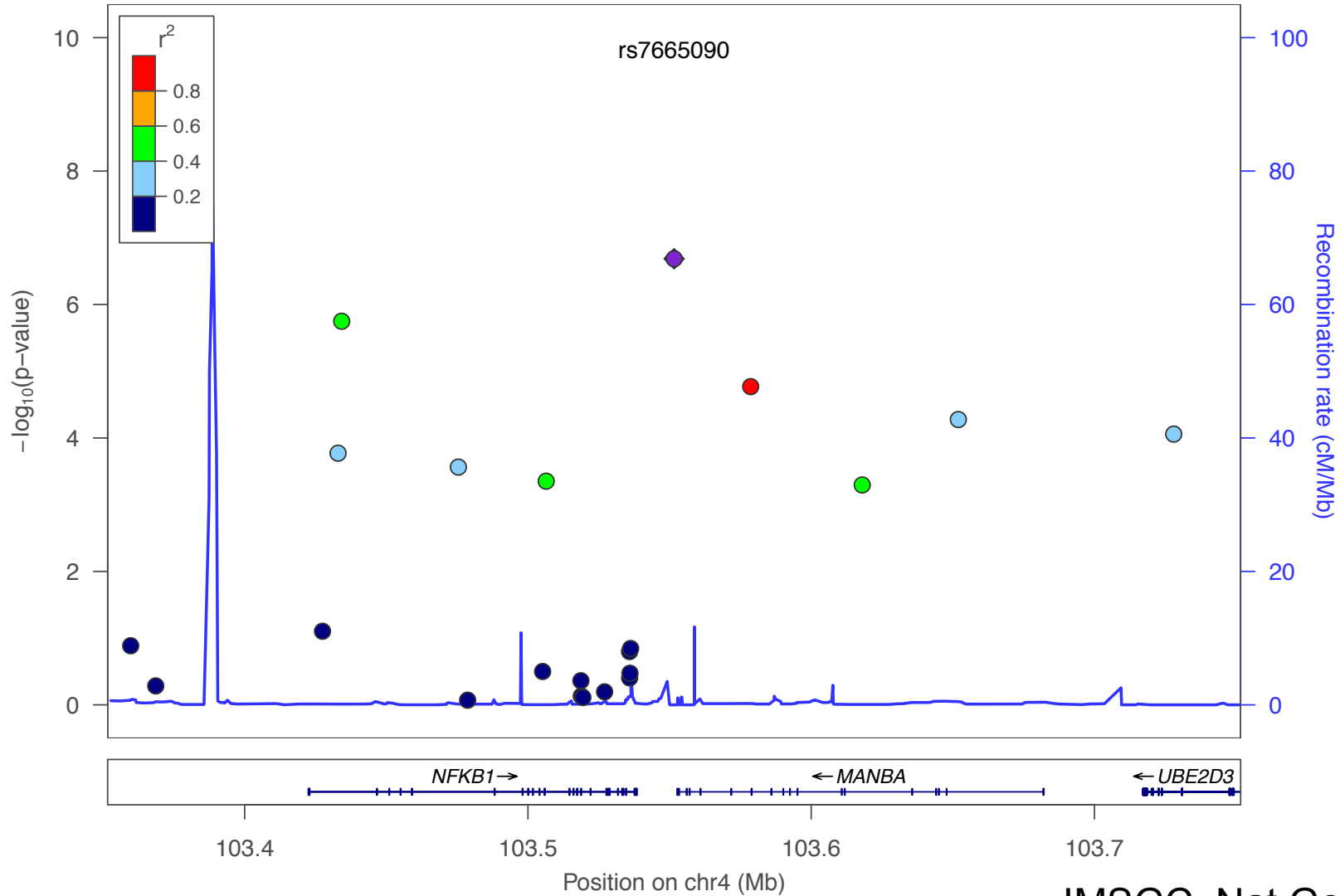
a



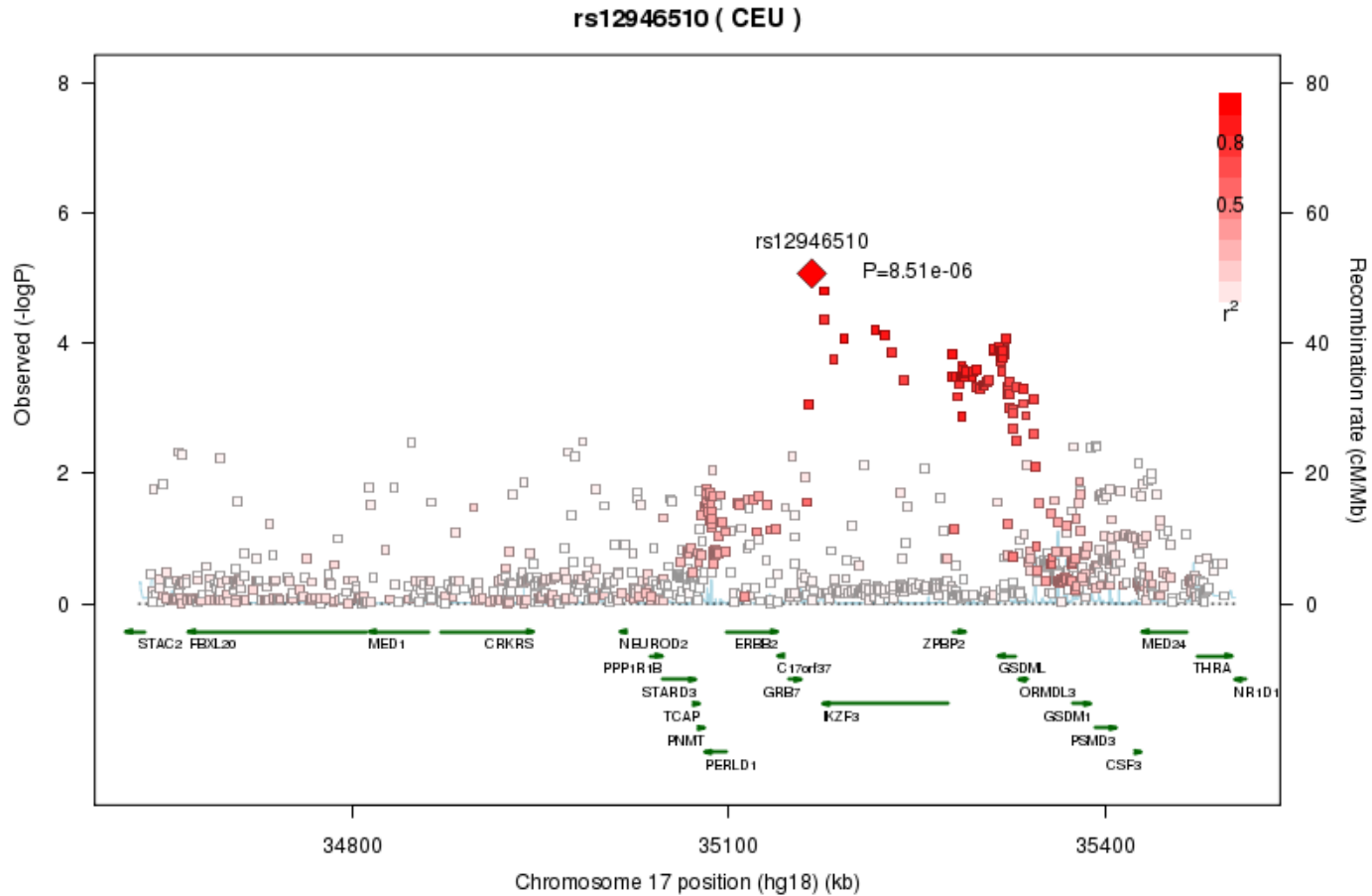
b



NFKB1 locus in MS GWAS



IKZF3/ORMDL3 locus in MS GWAS



Approach

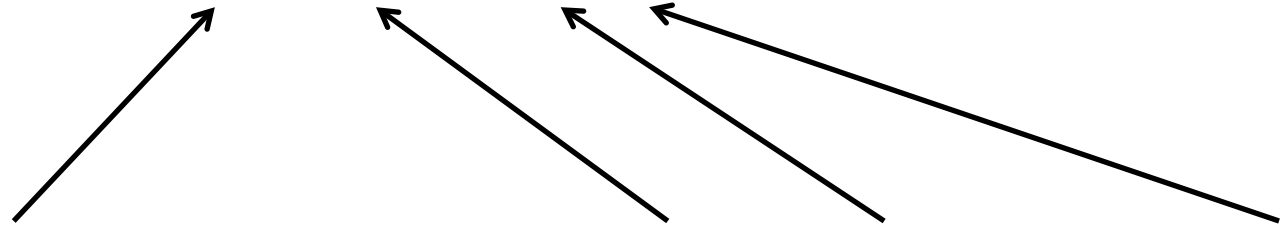
Total gene posterior

$$GP = \sum(CP_{DHS, gene} \times PPA_{DHS})$$

Gene



Gene-DHS
correlation
posterior
CP x PPA



DHS



Regulatory
posterior
 $RP = \sum(PPA_{DHS})$

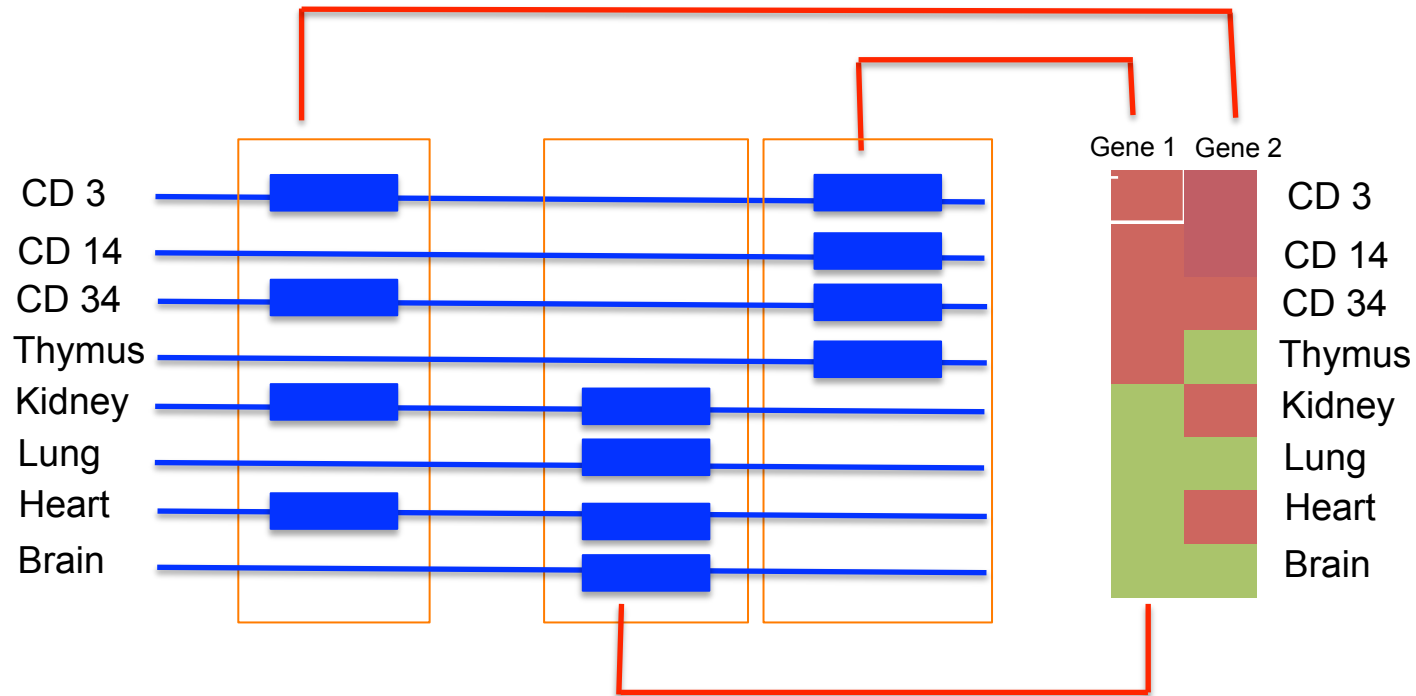
Posterior
probability
of association
PPA



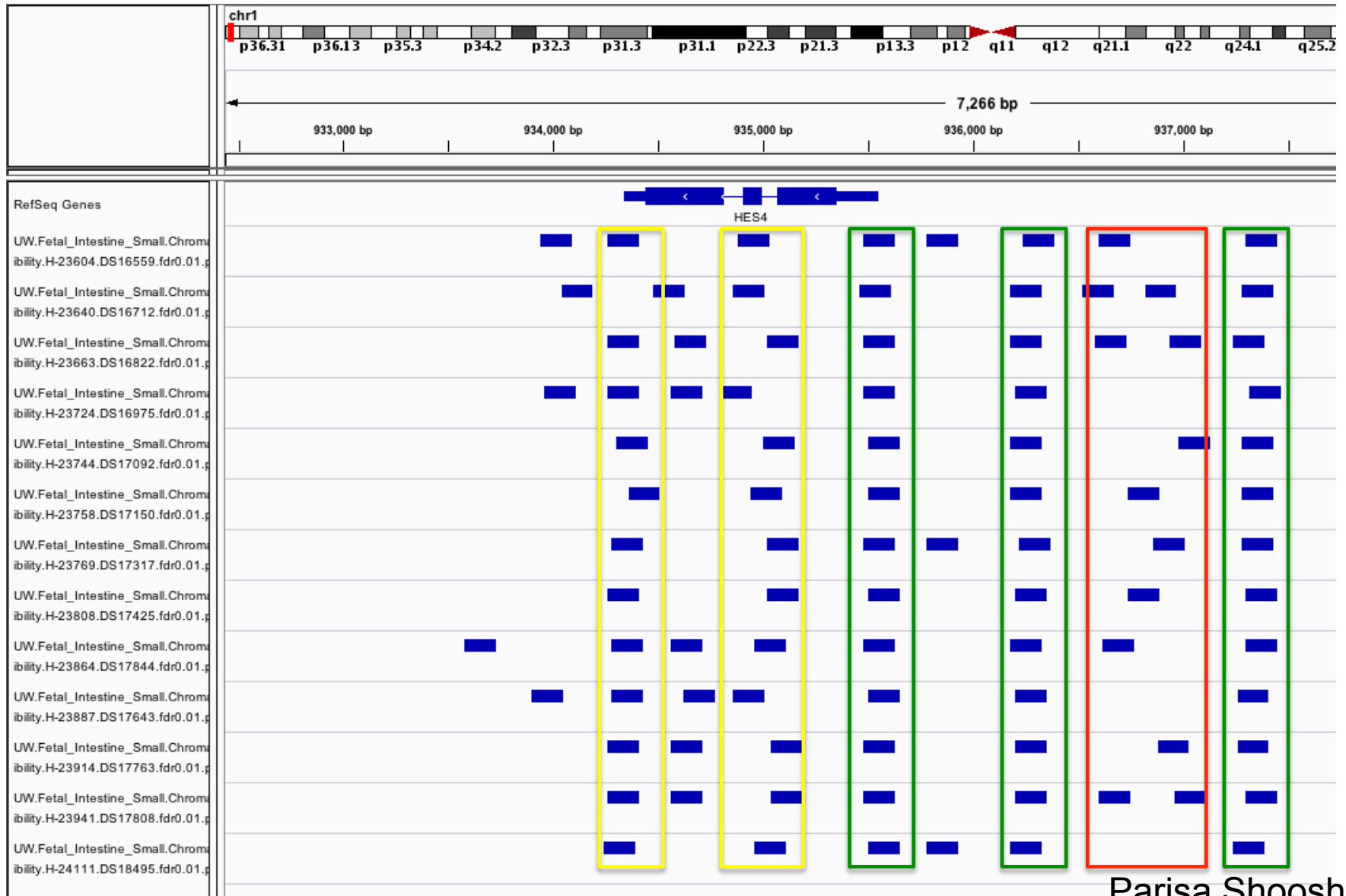
SNP



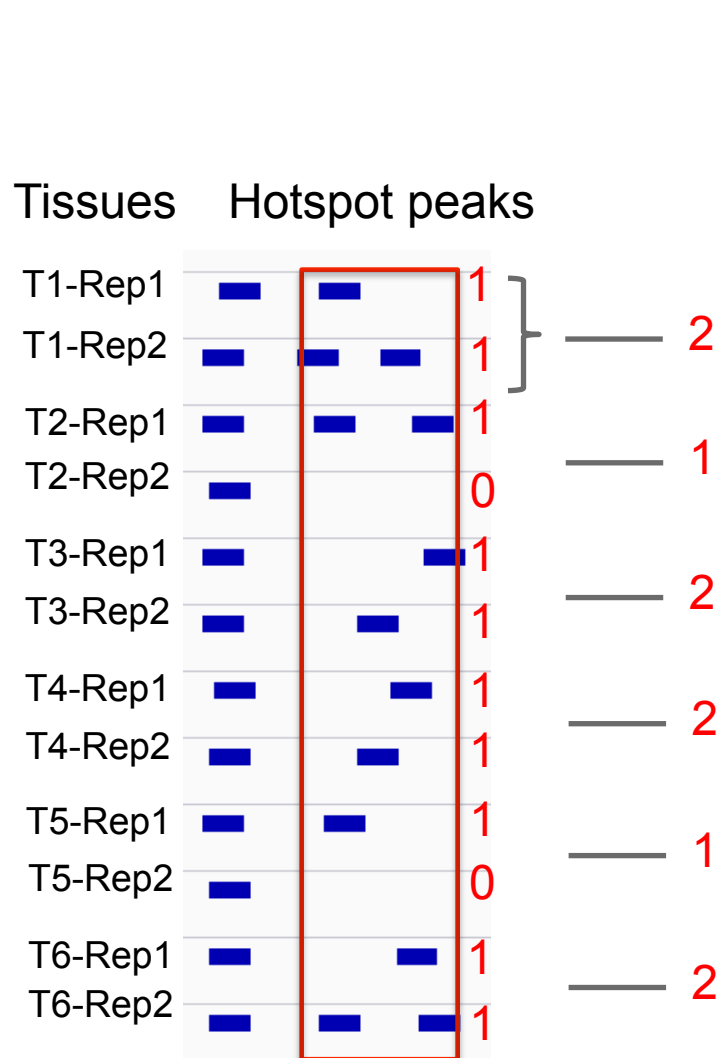
Problem 1: DHS-gene correlations



Aligning DHSs Over Samples



Identify detectable DHS clusters



Scenario 1		Scenario 2	
C1	1	C1	2
C2	1	C2	2
C3	1	C3	1
C4	0	C4	0
C5	0	C5	0
C6	1	C6	2

Unreliable Cluster

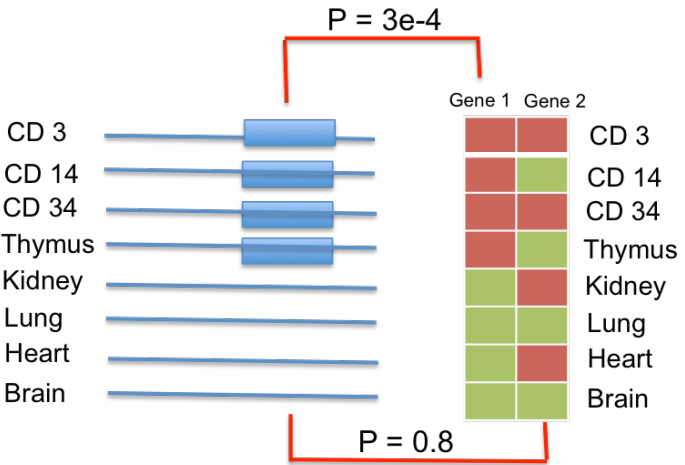
Reliable Cluster

Align over 57 tissue replicates from REP

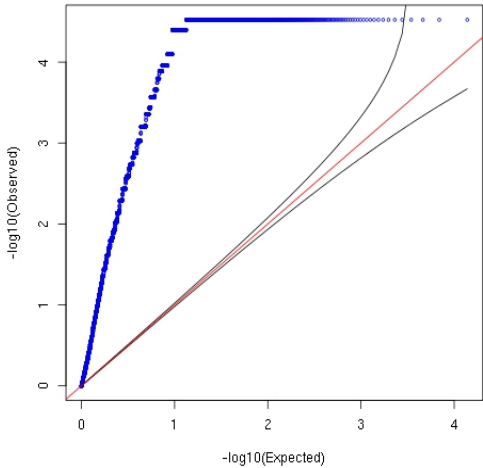
1,079,138/1,994,675 (54.1%) clusters pass
Cover 8% of genome (cf. 14% of all DHS)

NB singletons, low power

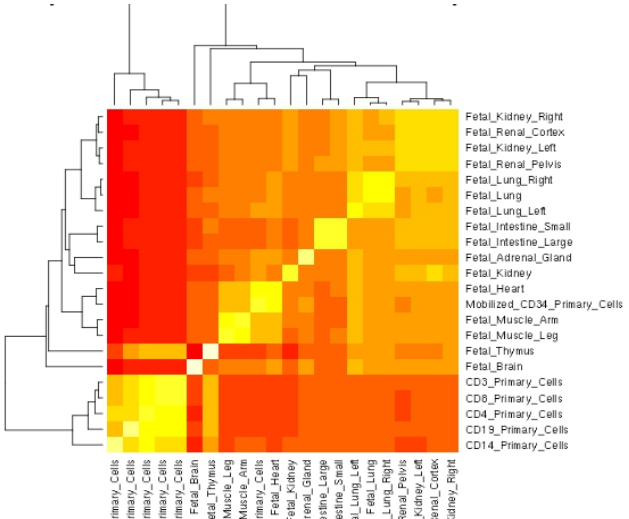
Challenge 2: Gene expression correlation



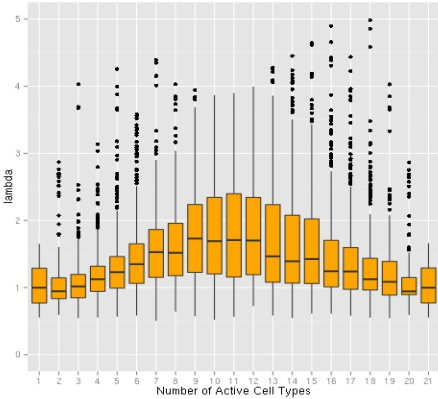
QQ plot for P Value of Correlation Between One DHS and 14000 Genes



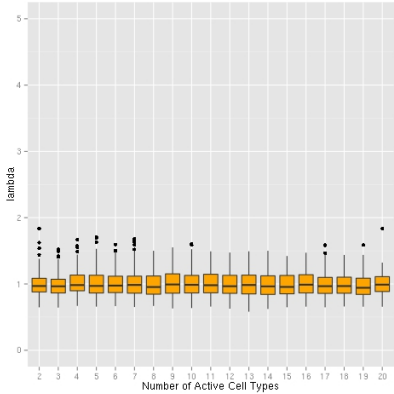
Correlation Structure of the Gene Expression Data



Before Correction



After



Approach

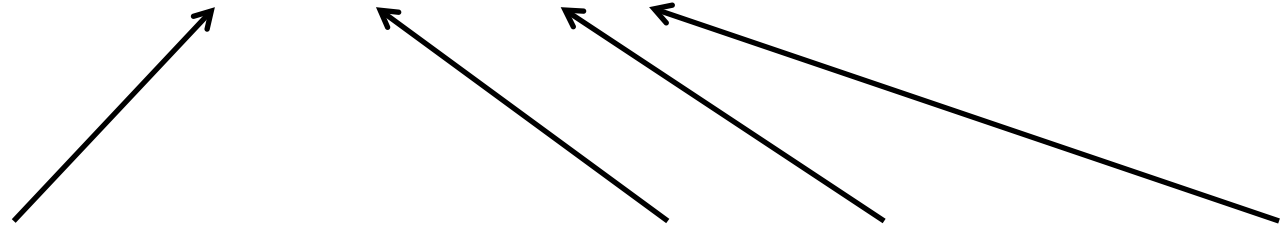
Total gene posterior

$$GP = \sum(CP_{DHS, gene} \times PPA_{DHS})$$

Gene



Gene-DHS
correlation
posterior
CP x PPA

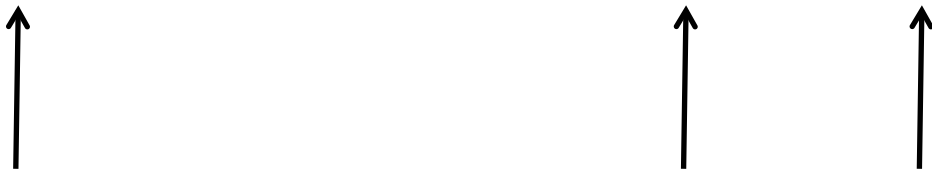


DHS



Regulatory
posterior
 $RP = \sum(PPA_{DHS})$

Posterior
probability
of association
PPA



SNP





Application to MS GWAS

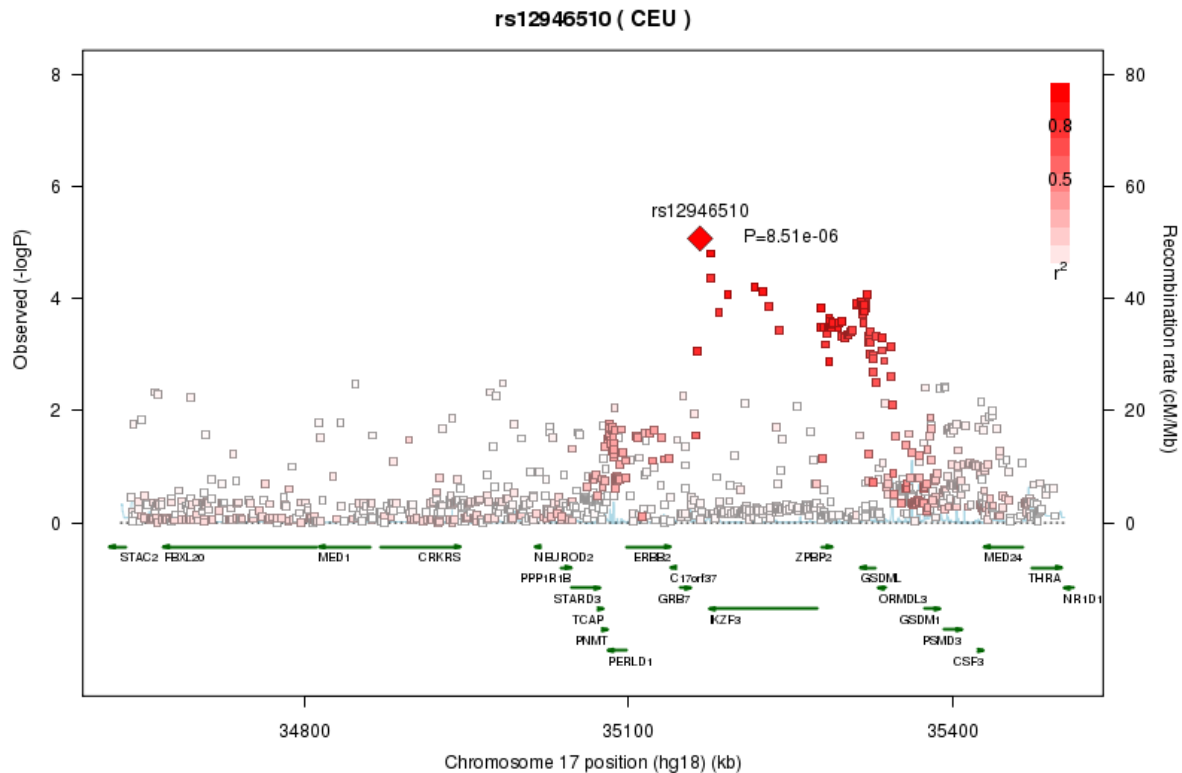
	Gene	GP
Chr 6	MDN1	0.555
90.5-91.5Mb	BACH2	0.162
RP = 0.945	GABRR2	0.106
	RRAGD	0.065
	GJA10	0.029
	MAP3K7	0.028



IKZF3/ORMDL3 locus

Gene	GP
------	----

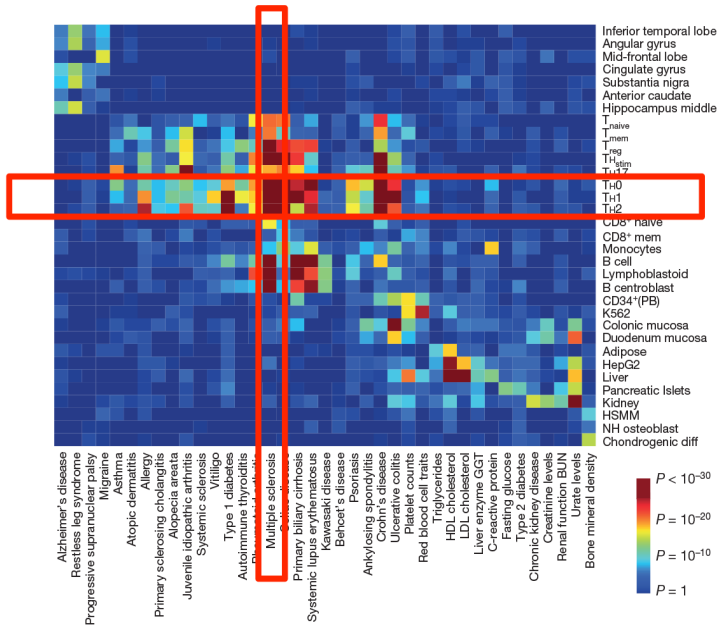
Chr 17	ORMDL3	0.029
34.5-35.5Mb	PIP4K2B	0.022
RP = 0.295	IGFBP4	0.018
	IKZF3	0.015
	GSDMB	0.014
	SMARCE1	0.013
	CCR7	0.013
	TNS4	0.01
	ZPBP2	0.009
	MED1	0.009
	MED24	0.009
	KRT24	0.009
	PNMT	0.008
	CDK12	0.007
	RPL23	0.007
	PSMD3	0.007
	PLXDC1	0.006
	TOP2A	0.006
	RARA	0.006



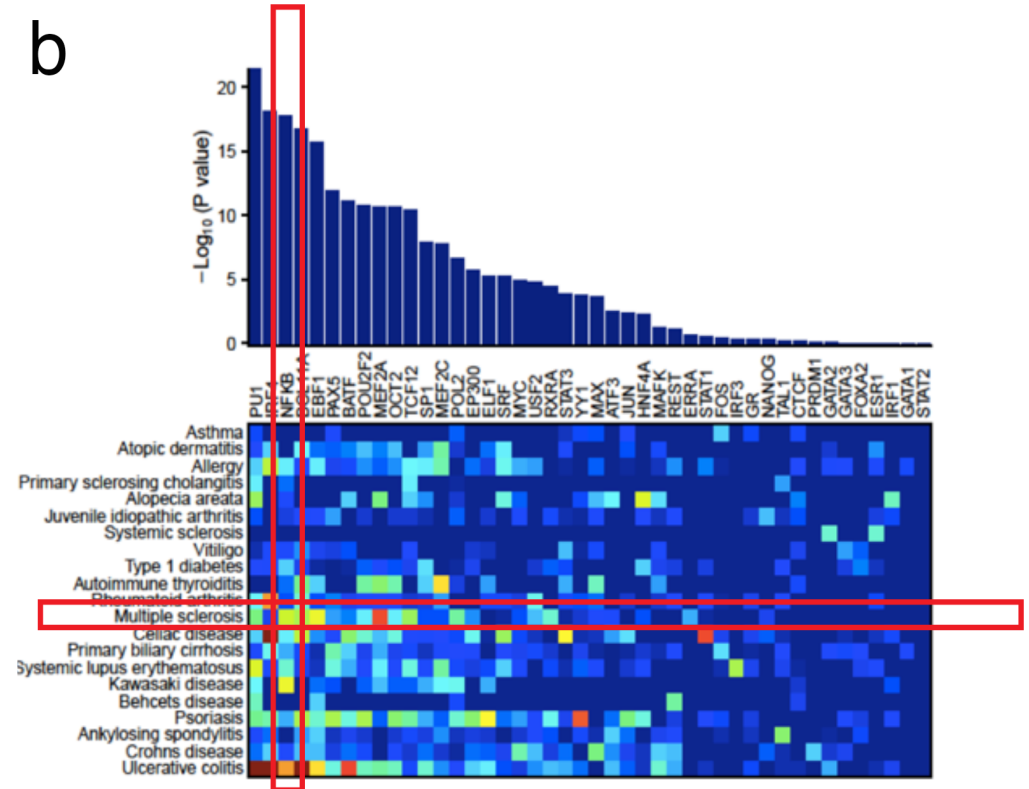


MS GWAS hits enriched in transcription factor binding sites

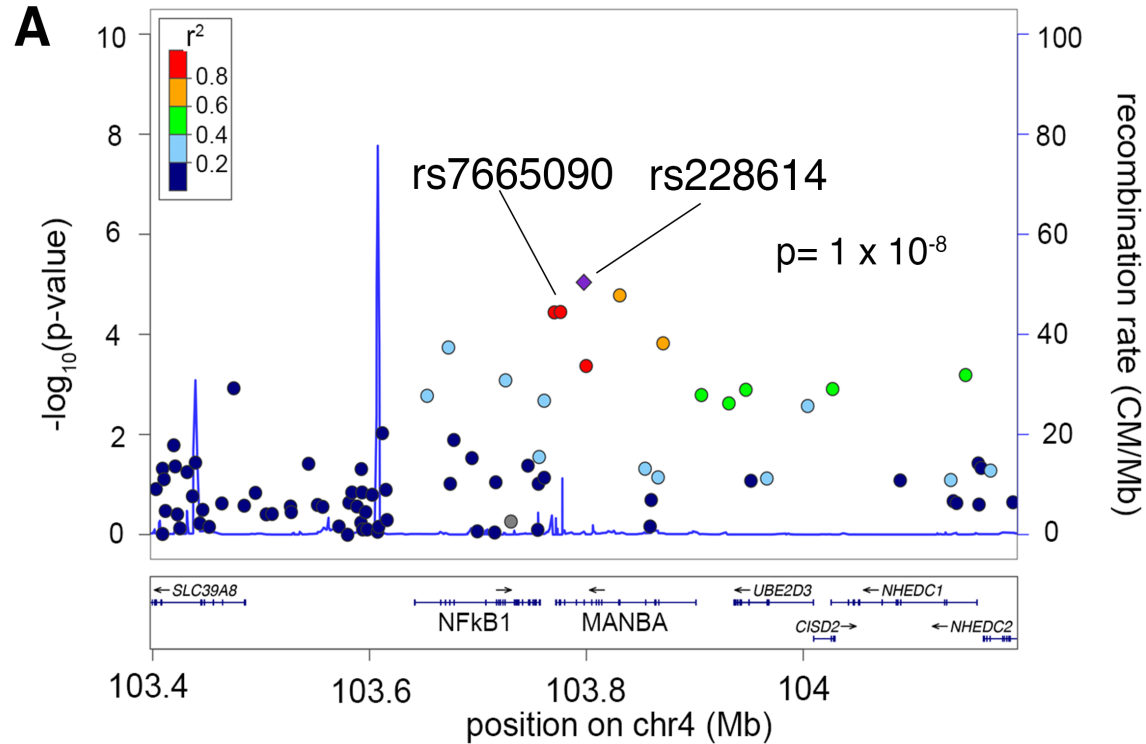
a



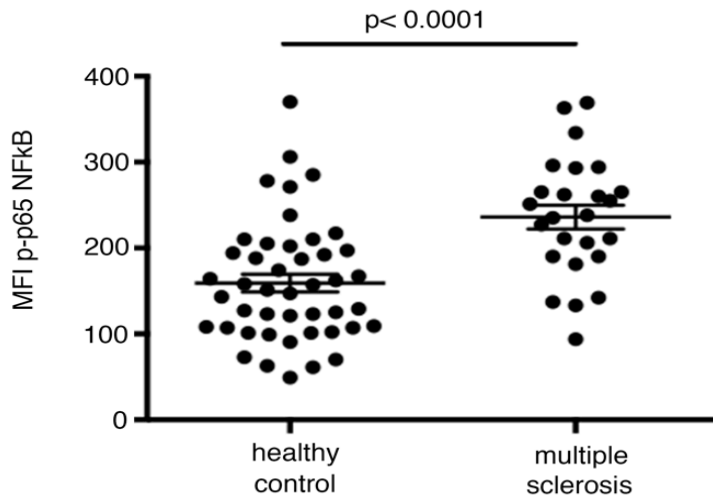
b



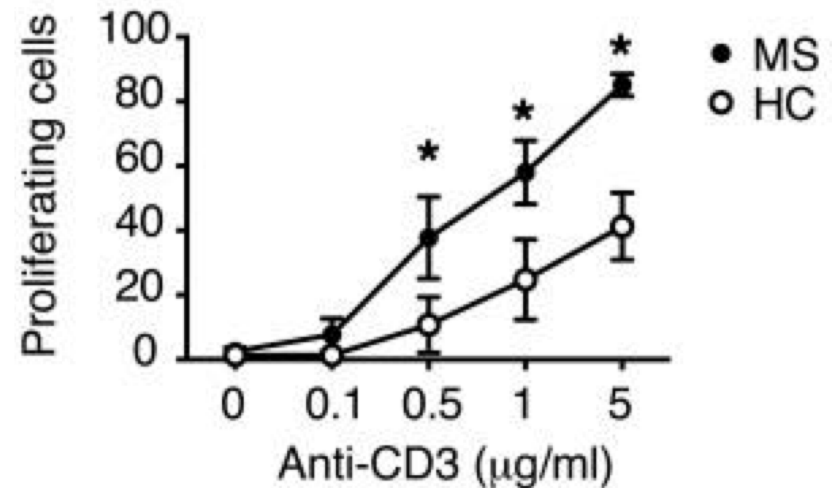
MS GWAS risk effect: NFKB1 locus



MS patients show altered NFκB signaling in CD4⁺ T cells



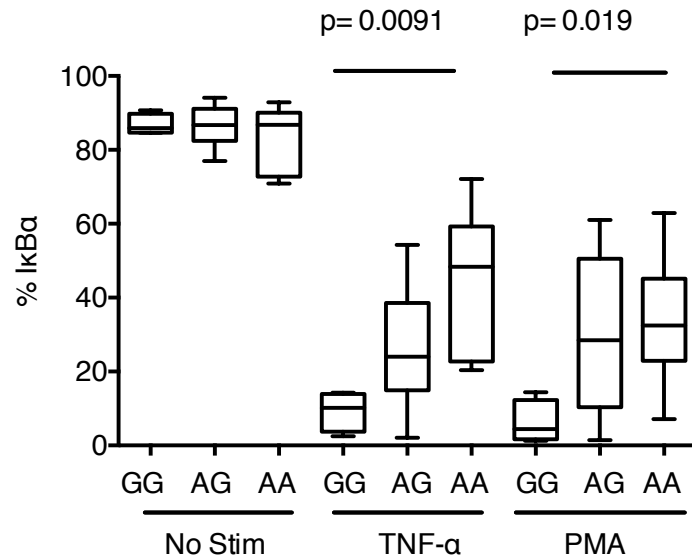
***ex vivo* CD4⁺ T cells show higher p-p65 (Housley *et al*, STM 2015)**



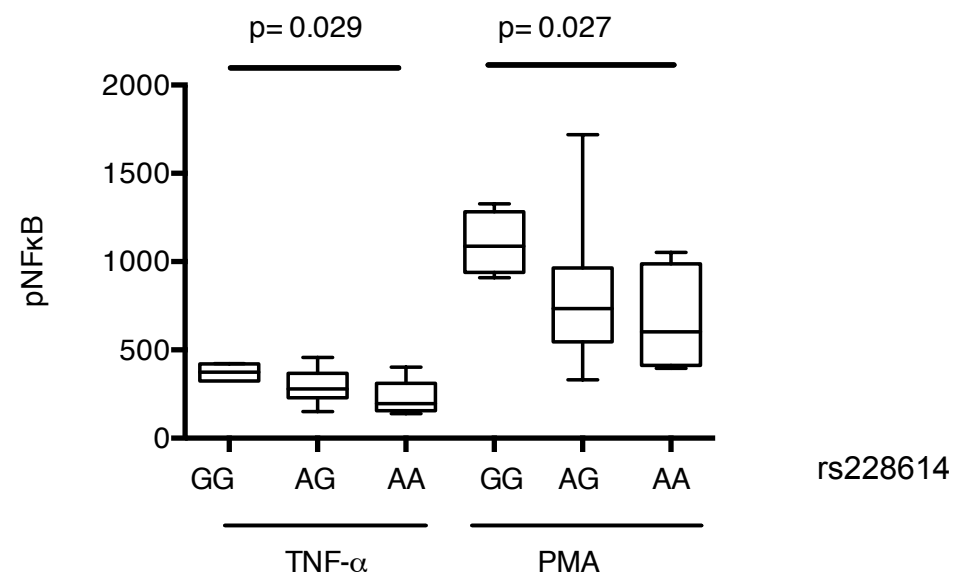
CD4⁺ T cells from MS patients proliferate more rapidly after stimulus (Kofler *et al* JCI 2014)

MS risk effect near NFKB1 alters signaling in CD4+ cells

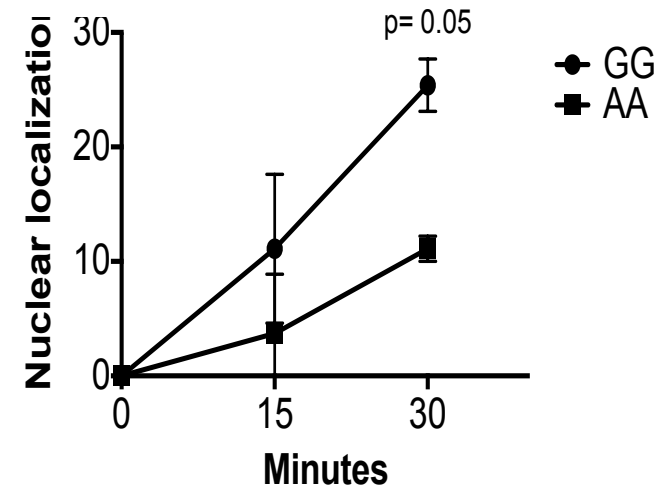
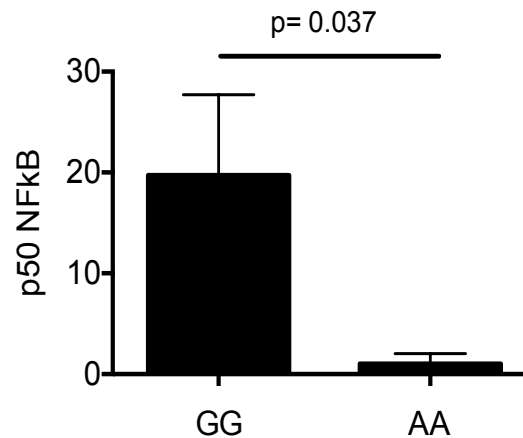
a



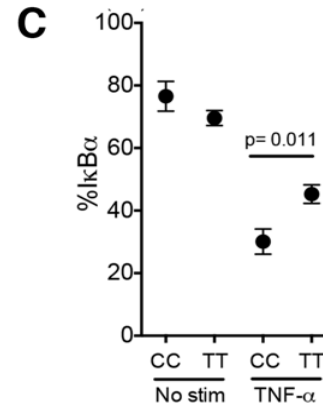
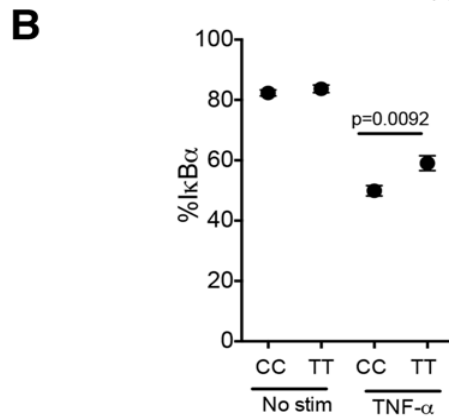
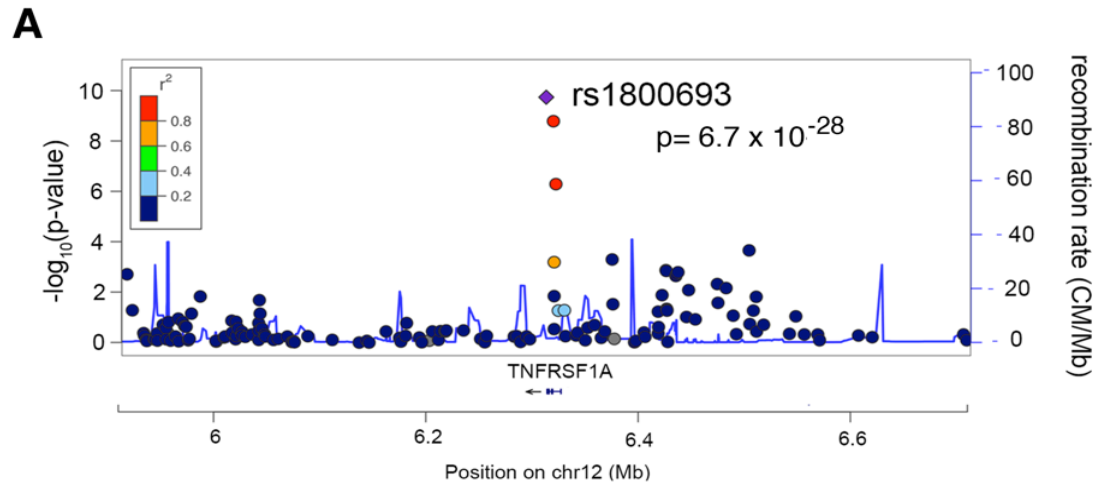
b



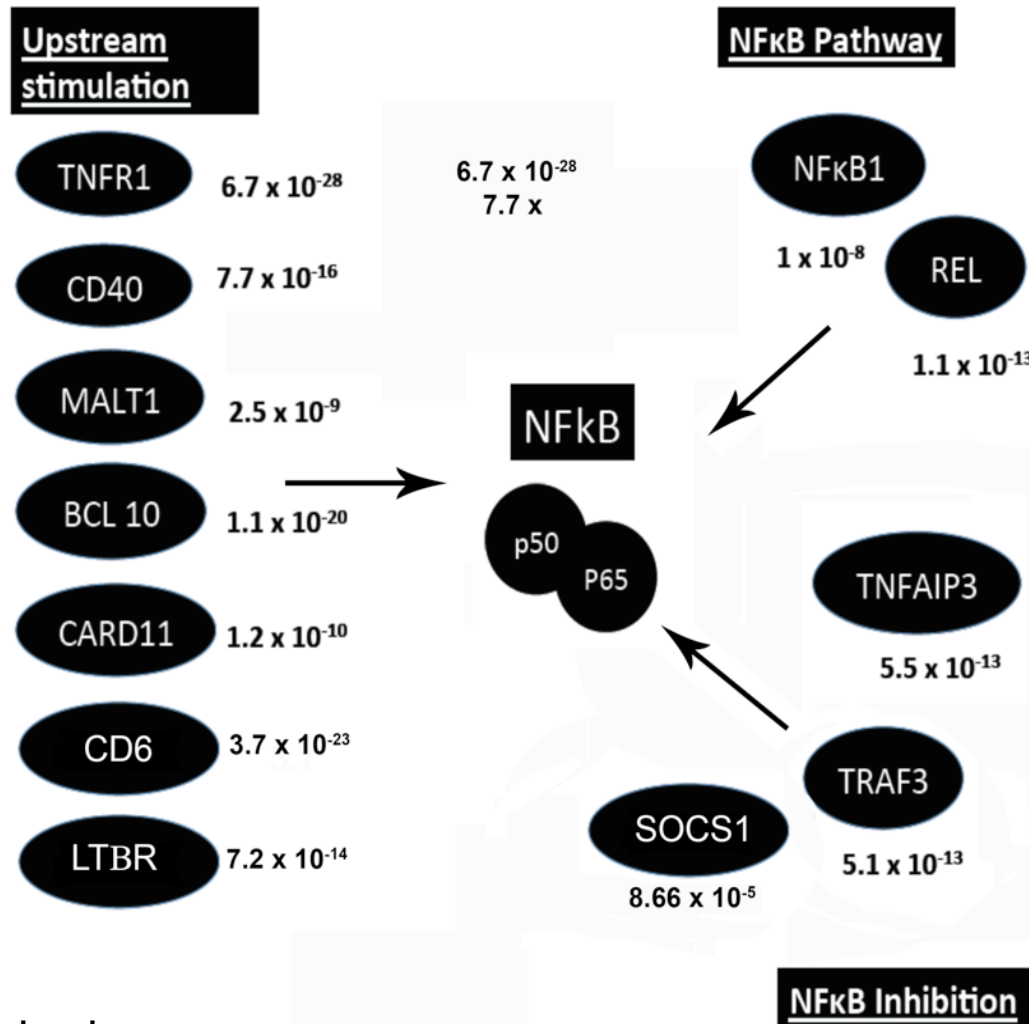
rs228614



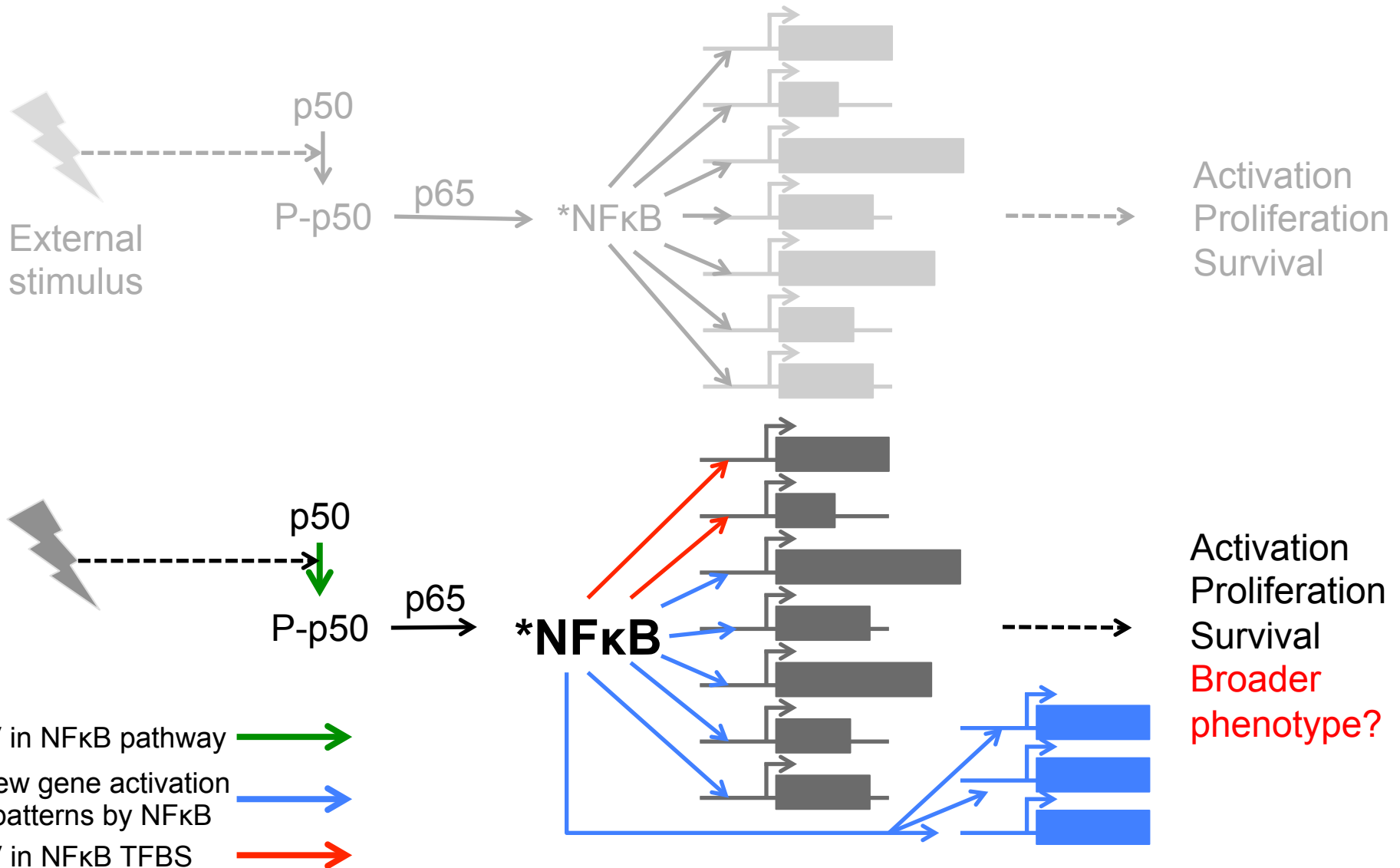
MS variant in TNFRSF1A alters TNF α -dependent NF κ B signaling



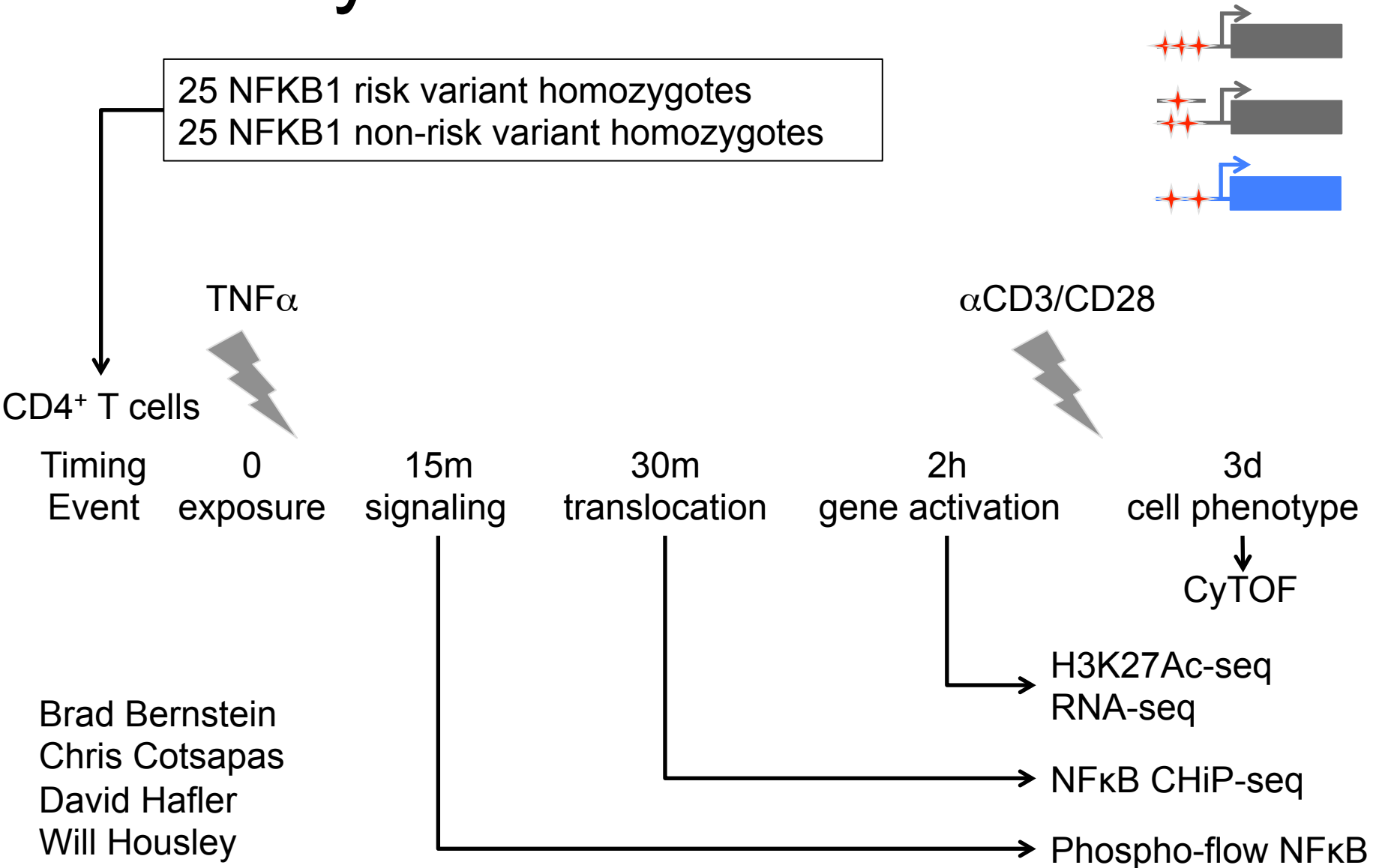
GWAS loci harbor many NFκB genes



Model: NFκB signaling variation



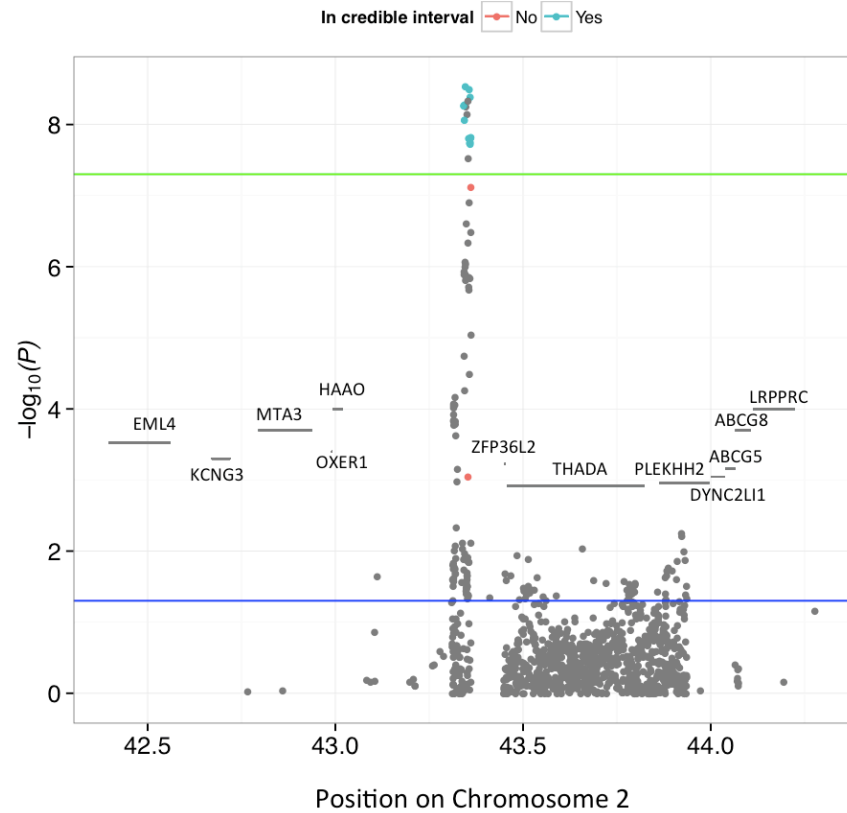
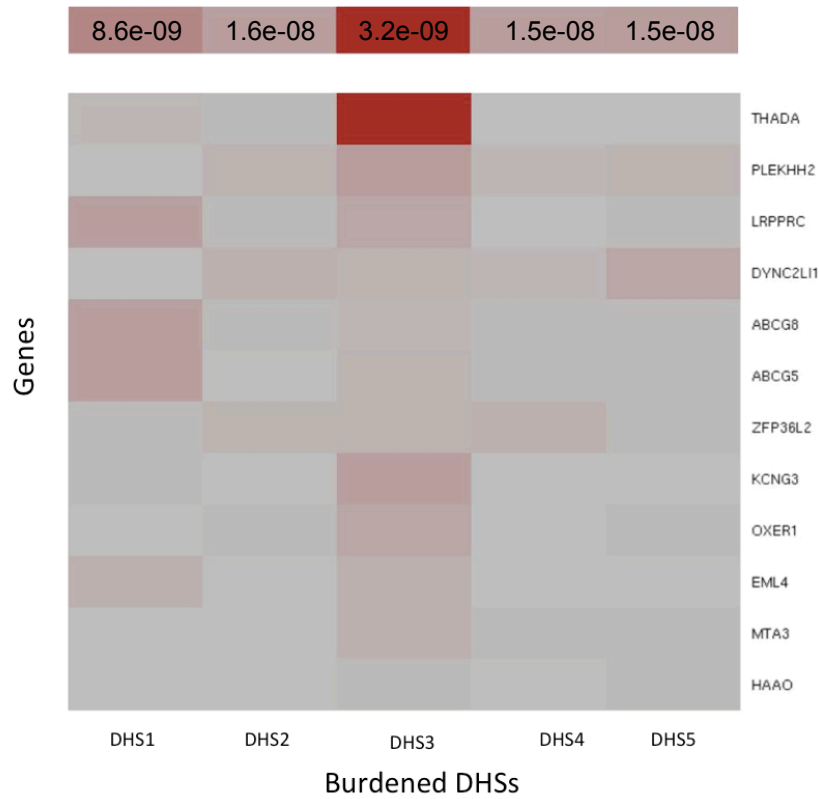
Systematic dissection I



Acknowledgements

- IMSGC
 - David Hafler
 - Phil De Jager
 - Steve Hauser
 - Adrian Ivinson
 - Nikos Patsopoulos
 - Many, many others
- Yale labs
 - **Parisa Shooshtari**
 - Mitja Mitrovic
 - Alex Casparino
 - **Will Housley**
- Partners
 - David Hafler
 - Phil De Jager
 - Brad Bernstein
 - John Stamatoyannopoulos





Color Key



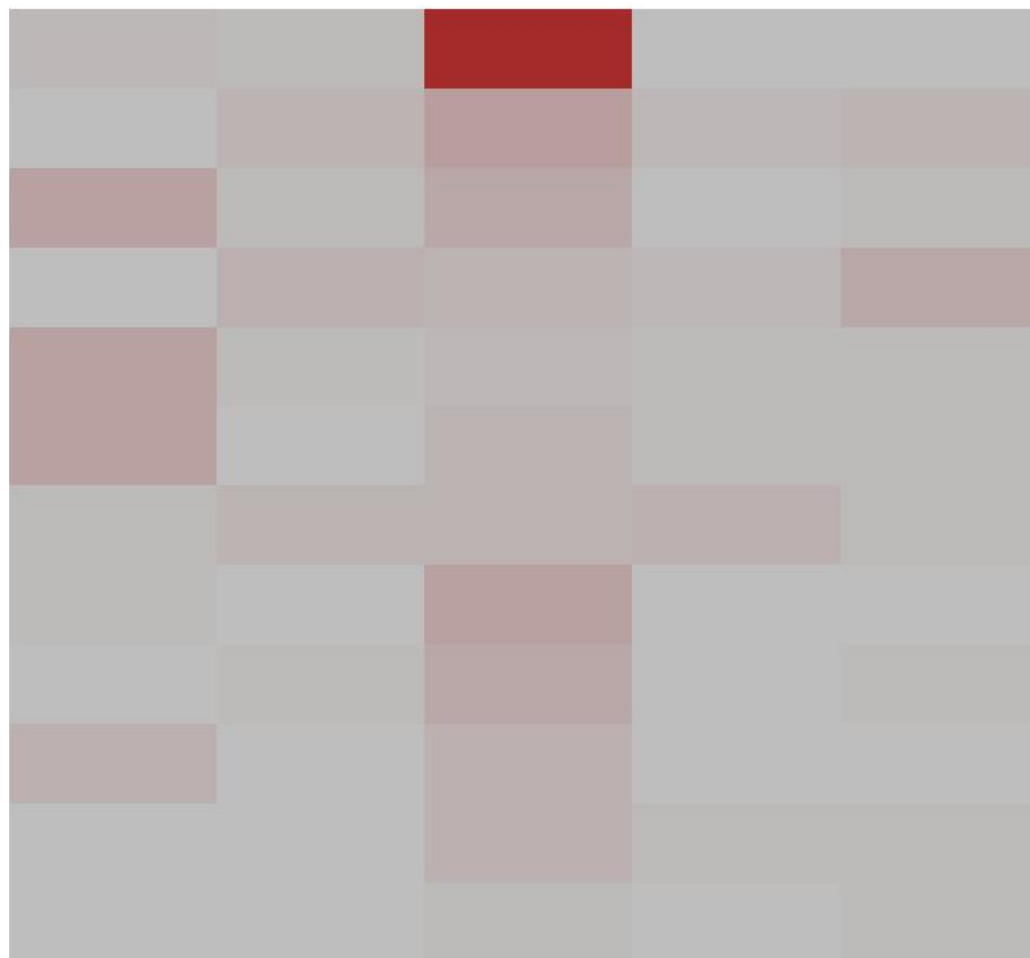
0.05 0.15

Color Key



0.02 0.06

Genes



DHS1

DHS2

DHS3

DHS4

DHS5

Burdened DHSs

THADA

PLEKHH2

LRPPRC

DYNC2LI1

ABCG8

ABCG5

ZFP36L2

KCNG3

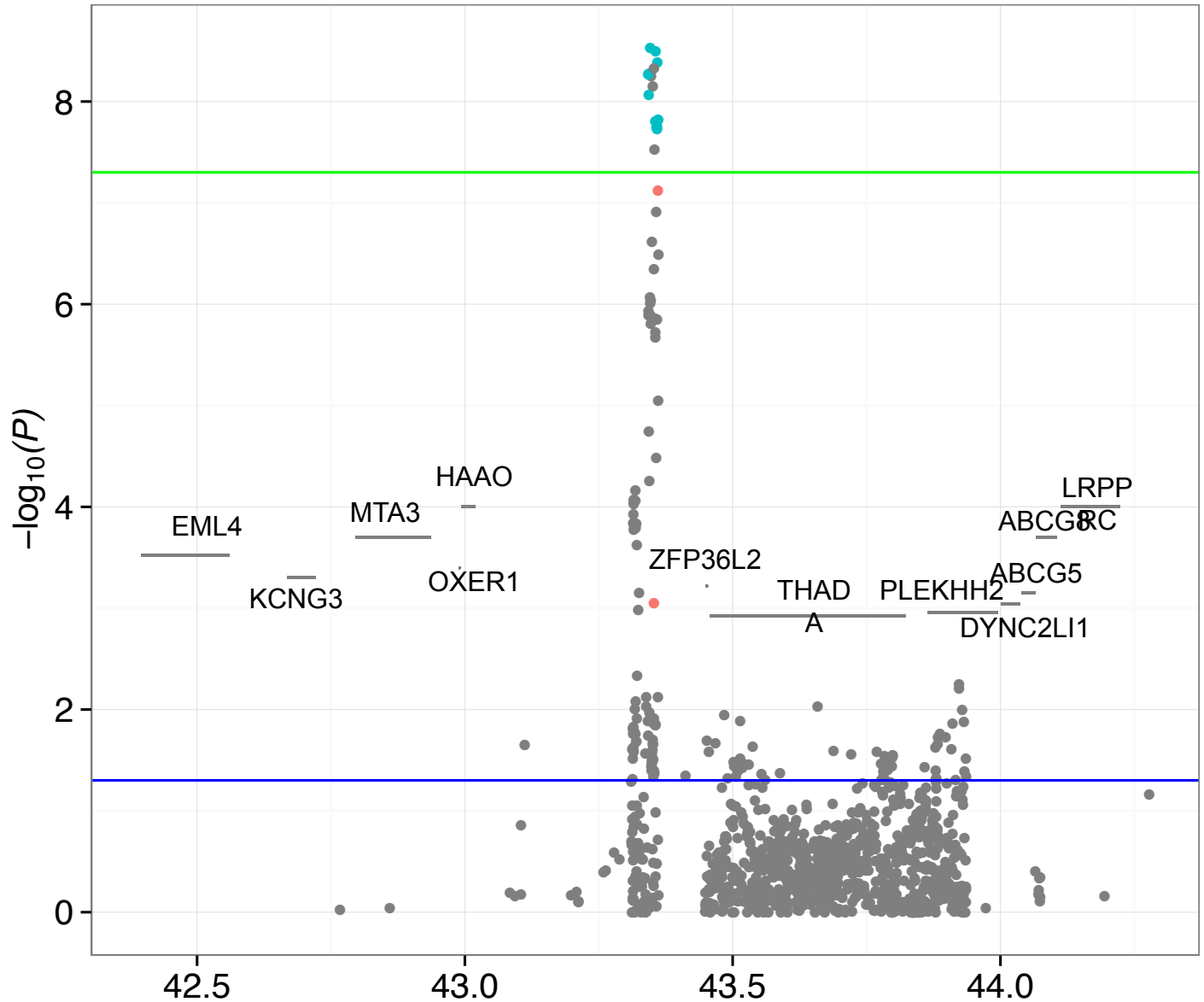
OXER1

EML4

MTA3

HAAO

In credible interval —●— No —●— Yes



Position on Chromosome 2