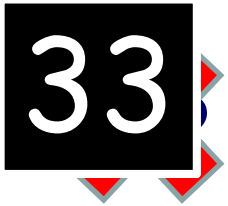


new Crohn's disease susceptibility genes and loci identified by the International IBD Genetics Consortium

Miles Parkes on behalf of IIBDGC

www.ibdgenetics.org

Acknowledgements

Andre Franke
A Hillary Steinhart
Albert Cohen
Amir Levine
Andre Van Gossum
Anne Griffiths
Anne Phillips
Carl Anderson
Carsten Bunning
Cathryn Edwards
Cecile Libioulle
Charlie Lees
Christopher Mathew
Cisca Wijmenga
Craig Mowat
Daan Hommes
Debby Laukens
Deborah D Proctor
Denis Franchimont
Derek P Jewell
Dermot McGovern
Edouard Louis
Frank Seibold
G Radford Smith
Grant Montgomery
Hakon Hakonarson
Ian Lawrance
Jack Satsangi
James Lee
Jean-Fred Colombel
Jean-Pierre Hugot
Jeff Barrett
Jeremy Sanderson
Jerry Rotter
John Mansfield
John Rioux
Jonas Halfvarson
Judy Cho
Julián Panés
Jürgen Glas
Kai Wang
Kent Taylor
Leif Torkvist
Lisa Simms
Marc Lemann
Mario Cottone
Mark Daly
Mark Silverberg
Marla Dubinsky
Martine De vos
Mauro D'Amato
Michel Georges
Miquel Regueiro
Miquel Sans
Murray Barclay
Paul Rutgeerts
Phil Schumm
Pieter Stokkers
Rebecca Roberts
Renata D'Inca
Richard Duerr
Richard Gearry
Rinse Weersma
Robert Baldassano
Salvator Cucchiara
Severine Vermeire
Stefan Schreiber
Stephan Brand
Stephan Targan
Steve Guthery
Steven R Brant
Subra Kugathasan
Tariq Ahmad
Ted Denson
Theodore Bayless
Thomas Balschun
Tim Florin
Vito Annese
William Newman
Yashoda Sharma

Funding Support from NIH, the Wellcome Trust, Giuliani SpA and Abbott Laboratories

Crohn's disease: Pathogenesis

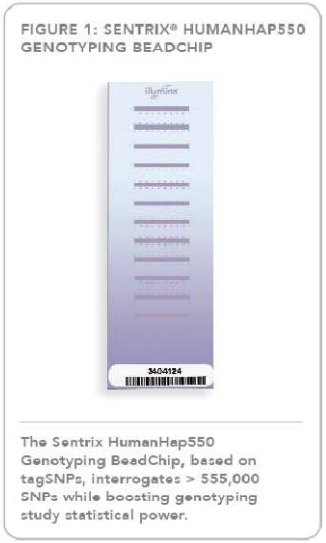
CD results from a dysregulated immune response to gut bacteria in genetically susceptible individuals.

Characterising the susceptibility genes should

1. Identify primary pathogenic pathways
2. Identify new targets for drug therapy
3. Help understand the environmental drivers

Genome Wide Association Scanning

- Genotype ≥0.5 million SNPs in thousands of individuals



- hypothesis-free, unbiased survey of genome for susceptibility loci

GWAS studies in Crohn's disease

Vol 447|7 June 2007|doi:10.1038/nature05911

nature

ARTICLES

Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls

The Wellcome Trust Case Control Consortium*

There is increasing evidence that genome-wide association (GWA) studies represent a powerful approach to the identification of genes involved in common human diseases. We describe a joint GWA study (using the Affymetrix GeneChip 500K Mapping Array Set) undertaken in the British population, which has examined ~2,000 individuals for each of 7 major diseases and a shared set of ~3,000 controls. Case-control comparisons identified 24 independent association signals at $P < 5 \times 10^{-7}$: 1 in bipolar disorder, 1 in coronary artery disease, 9 in Crohn's disease, 3 in rheumatoid arthritis, 7 in type 1

GWAS studies in Crohn's disease

- Identified importance of autophagy and Th17 pathways etc.

1st International Crohn's disease GWAS Meta-analysis: 32 confirmed loci

Barrett et al. *Nature Genetics* 2008

ARTICLES

nature genetics

Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease

Jeffrey C Barrett[*1], Sarah Hansoul[2], Dan L Nicolae[3], Judy H Cho[4], Richard H Duerr[5,6], John D Rioux[7,8], Steven R Brant[9,10], Mark S Silverberg[11], Kent D Taylor[12], M Michael Barmada[6], Alain Bitton[13], Themistocles Dassopoulos[9], Lisa Wu Datta[9], Todd Green[8], Anne M Griffiths[14], Emily O Kistner[15], Michael T Murtha[4], Miguel D Regueiro[5], Jerome I Rotter[12], L Philip Schumm[15], A Hillary Steinhart[11], Stephan R Targan[12], Ramnik J Xavier[16], the NIDDK IBD Genetics Consortium[33], Cécile Libioulle[2], Cynthia Sandor[2], Mark Lathrop[17], Jacques Belaiche[18], Olivier Dewit[19], Ivo Gut[17], Simon Heath[17], Debby Laukens[20], Myriam Mni[2], Paul Rutgeerts[21], André Van Gossum[22], Diana Zelenika[17], Denis Franchimont[22], Jean-Pierre Hugot[23], Martine de Vos[20], Severine Vermeire[21], Edouard Louis[18], the Belgian-French IBD Consortium[33], the Wellcome Trust Case Control Consortium[33,34], Lon R Cardon[1], Carl A Anderson[1], Hazel Drummond[24], Elaine Nimmo[24], Tariq Ahmad[25], Natalie J Prescott[26], Clive M Onnie[26], Sheila A Fisher[26], Jonathan Marchini[27], Jilur Ghori[28], Suzannah Bumpstead[28], Rhian Gwilliam[28], Mark Tremelling[29], Panos Deloukas[28], John Mansfield[30], Derek Jewell[31], Jack Satsangi[24],

But....

- Only 20% of genetic variance explained...

Aim

- Identify additional Crohn's disease susceptibility genes and loci which surpass a stringent genome-wide significance threshold of $P<5x10^{-8}$

Methods - overview

- meta-analysis of all available Crohn's disease GWAS'd case-controls ('discovery panel')
- Follow-up new signals in independent replication panel

NB: expanded ++ consortium has much greater power to detect 'typical' loci of modest effect

Subjects: GWAS discovery panel for meta-analysis

Index GWAS	**Crohn's disease cases**	**Healthy controls**	
Adolescent	1689	6197	Nat Genet 2009
German	479	1145	PLoS Genet 2009
USA (Cedars Sinai)	925	2882	
Belgium	537	913	PLoS Genet 2007
USA (NIDDK)	956	982	Nature Genet 2007
UK (WTCCC)	1747	2937	Nature 2007
TOTAL	**6,333**	**15,056**	

Methods: meta-analysis

- Used GWAS data +'Beagle' to impute genotypes for 953,000 HapMap3 SNPs / sample
- Standard 1 df allele-based tests of association – summarized as Z scores & p values for each SNP
- Each new locus meeting $P<1x10^{-5}$
  => 'focal' SNP +/- proxy genotyped in large replication panel

Subjects: replication panel

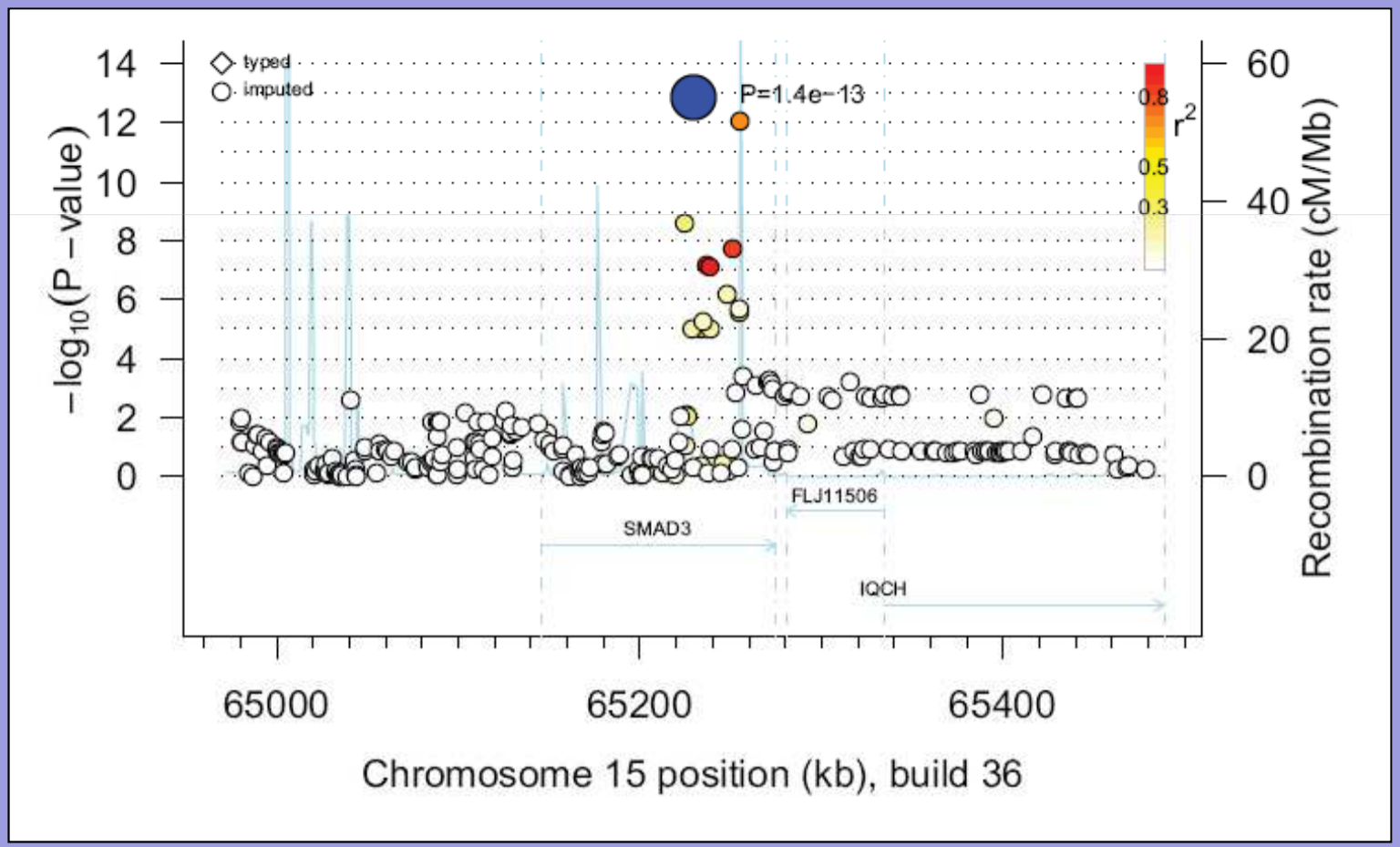
REPLICATION	cases	controls	platform	
Australia	1357	1923	Sequenom	
Belgium	1282	1682	SNPlex/Taqman	
France	414 trios		SNPlex/Taqman	
Germany	3808	2747	SNPlex/Taqman	
Israel	444	376	SNPlex/Taqman	
Italy	921	899	SNPlex/Taqman	
Netherlands	1101	269	SNPlex/Taqman	
New Zealand	514	457	Sequenom	
Spain	325	987	SNPlex/Taqman	
Sweden	724	992	SNPlex/Taqman	
UK	3243	2431	Sequenom	
USA (cedars)	1172	501	Sequenom	
USA (NIDDK)	803	762	Illumina	
TOTAL	**14,934**	**13,647**		

Results

- 52 new loci met $P<1x10^{-5}$ in discovery panel => focal SNP genotyped in replication panel
- SNPs from 33 distinct new loci showed evidence of association at $P<5x10^{-8}$ in combined 'discovery + replication' panel (with $P<0.05$ in replication)

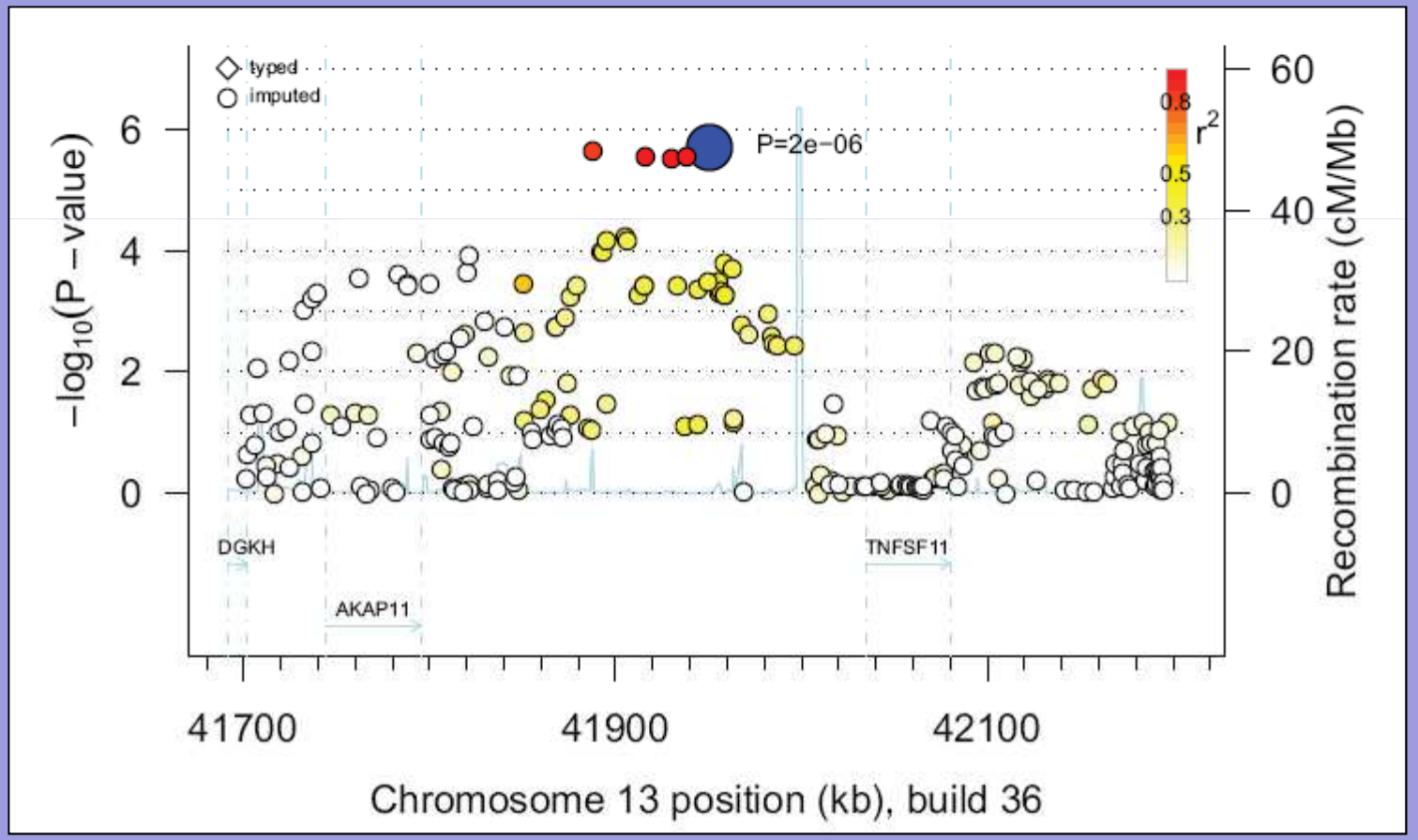
Representative association plot

‘one gene’ locus



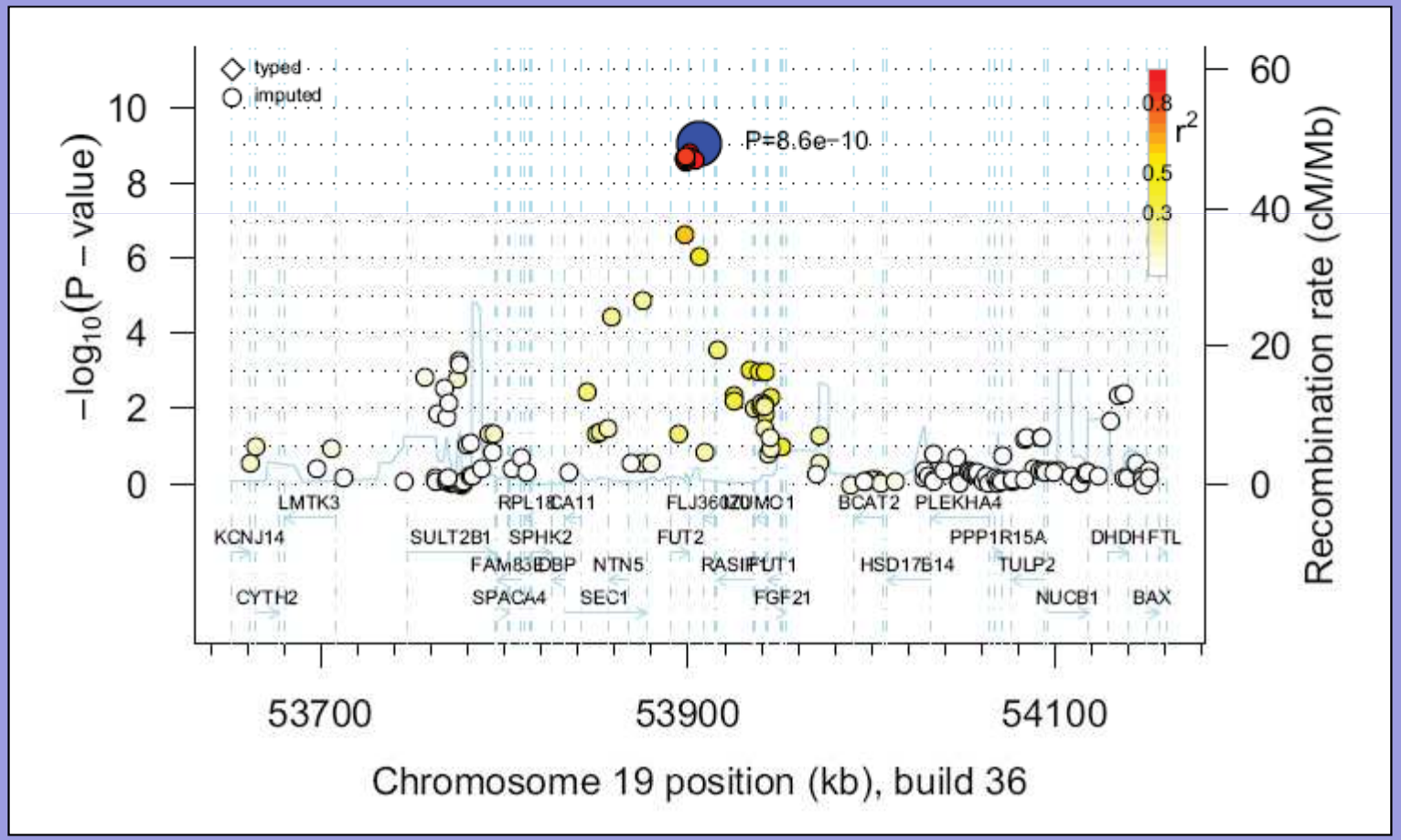
Representative association plot

‘Gene desert’



Representative association plot

‘Multi-gene’ locus



Loci with a single gene

SNP	CHR	meta-GWAS P	Rep P	Combined P	OR	Gene
rs17293632	15	1.41E-13	2.00E-08	**2.70E-19**	1.12	SMAD3
rs212388	6	1.41E-07	2.40E-05	**2.30E-11**	1.1	TAGAP
rs13428812	2	1.41E-08	0.00059	**8.50E-10**	1.06	DNMT3A
rs11167764	5	1.1E-09	0.0042	**2.00E-09**	1.06	NDFIP1
rs12722489	10	8.51E-06	5.20E-05	**2.90E-09**	1.11	IL2RA
rs1847472	6	3.63E-06	1.40E-04	**5.10E-09**	1.07	BACH2
rs1998598	1	4.9E-09	0.016	**8.70 E-09**	1.04	DENND1B

Loci with a single gene

SMAD3

- Transcriptional regulator downstream of TGF-β
- key role in TGF-β -mediated induction of Foxp3^{+} Tregs
- SMAD3 deficiency => incr Th17 differentiation
- Reduced SMAD3 phosphorylation seen in IBD
  - Tone et al. *Nat Immunol* 2008
  - Lu et al. *J Immunol* 2010
  - Martinez et al. *J Biol Chem*. 2009

Loci with a single gene

TAGAP

- T-cell Activation GTPase-Activating Protein
- Regulator of T cell activation / co-regulated with IL2
- Also associated with
  - celiac disease
  - rheumatoid arthritis
  - type 1 diabetes

Loci with a single gene

DNMT3A

- encodes DNA methyltransferase 3a - one of three key methyltransferases
- epigenetic regulator of gene transcription by methylating cytosine in CpG islands.
- many roles incl dynamic regulation of innate + adaptive immunity

  - Gamper et al *J Immunol*. 2009
  - El Gazzar et al *J Biol Chem.* 2009

Loci with a single gene

IL2RA

- encodes part of IL2 receptor
- regulates response to autoantigens by modulating Foxp3^{+} Tregs
- associated SNP correlates with IL2RA (CD25) expression on CD4+ naïve + memory T cells
- also associated with T1D + MS
  - Dendrouet al *Nature Genetics* 2009
  - Burchill et al *Immunol Lett*. 2007

Loci with a single gene

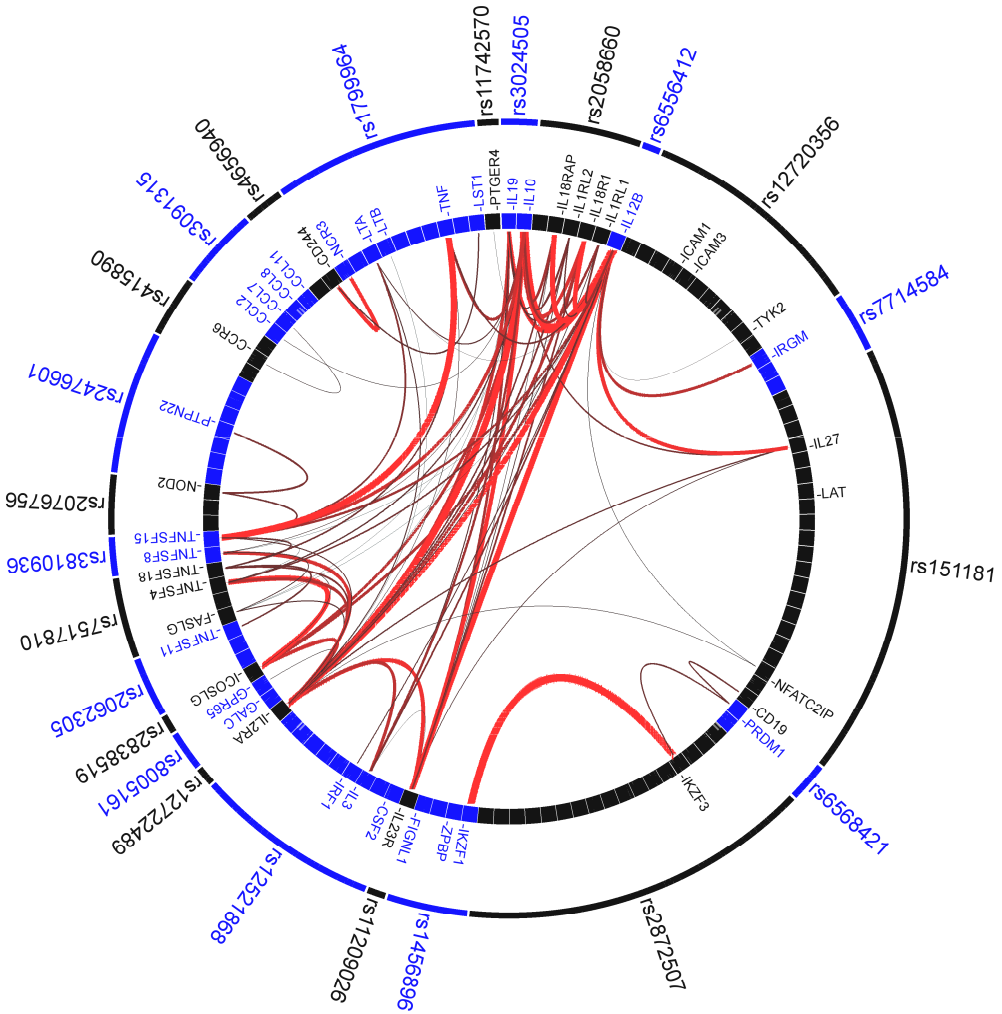
DENND1B

- expressed in dendritic and effector memory T cells
- represses TNFR1 signaling => modulates Th1-Th2 cytokines
- Recent assoc with asthma
  - Sleiman et al *NEJM* 2010

Methods to identify key gene within multi-gene loci

- eQTL analysis
  - to correlate focal SNP variation with gene expression
- Coding SNP (cSNP) analysis
  - using 1000G and Hapmap3 to identify genes containing cSNP's in LD with focal SNP
- GRAIL (pathway) analysis
  - Gene Relationships Among Implicated Loci
  - identify functional connectivity between genes in CD-associated loci

GRAIL connectivity for 69 CD loci



‘One from Many’ Analysis

SNP	CHR	meta-GWAS P	Rep P	Combined P	OR	Gene Highlighted
rs281379	19	8.60E-10	5.20E-05	**7.40E-12**	1.07	FUT2
rs2549794	5	4.47E-11	0.002	**1.10E-10**	1.05	ERAP2
rs12720356	19	9.20E-10	1.90E-05	**1.40E-12**	1.12	TYK2
rs3024505	1	8.32E-09	1.50E-07	**1.60E-14**	1.12	IL10

One from Many

FUT2 (cSNP)

- encodes alpha-(1,2)fucosyltransferase
- regulates expression of Lewis ABO(H) blood group antigens on epithelial cell surface
- 20% population *FUT2* non-secretor genotype => marked resistance to norovirus
- also partial resistance to H pylori and HIV infection
  - Thorven et al. J Virol. 2005

One from Many

ERAP2 (strong eQTL)

- endoplasmic reticulum aminopeptidase 2
- regulated by NFκB
- trims peptides for presentation on MHC class I => critically affects antigen presentation to T cells
- NB: ERAP1 associated with ank spond
  - Saveanu et al *Nature Immunology* 2005

One from Many

TYK2 (cSNP + GRAIL)

- encodes tyrosine kinase 2 (JAK family)
- transduces γ-IFN, IL12 and IL23 signaling

  => regulates Th1 / Th17 development
- also role in TLR-mediated responses in dendritic cells, incl IL12 + IL23 production

  – Ghoreschi et al Immunol Rev. 2009

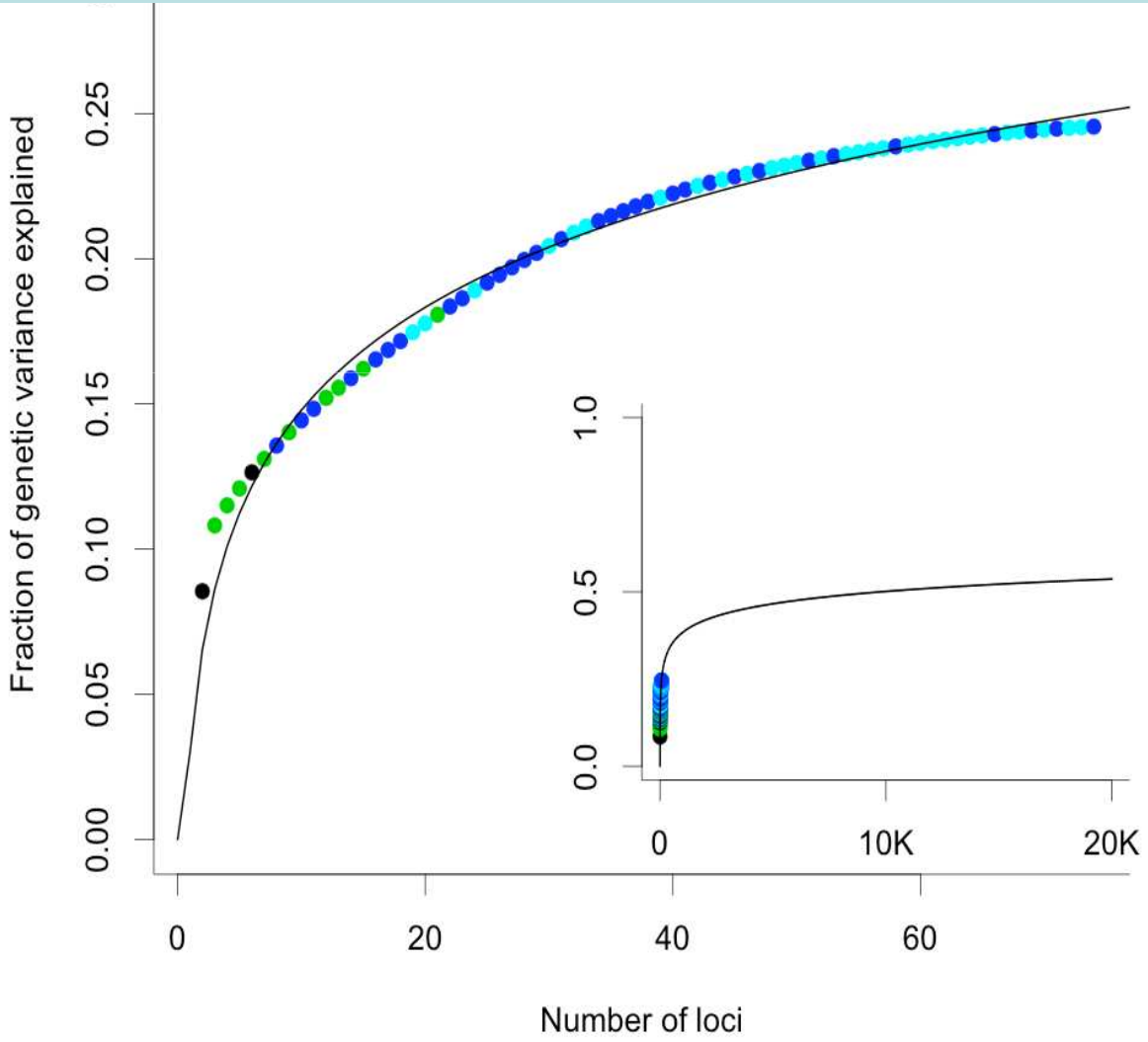
Gene desert

TNFSF11 = RANKL (eQTL)

- receptor activator of NFκB
- encodes a member of TNF family
- stimulates dendritic cells => proliferation of naive + systemic Treg populations
  - Loser et al. *Nature Med.* 2006
- also regulates osteoclast activity / bone loss
- previous studies => increased plasma levels in CD
  - Moschen et al. Gut 2005

What proportion of genetic variance now explained?

~25%



Conclusions

- 33 new Crohn's disease susceptibility loci identified
  - Total tally confirmed CD loci now = 69
- Specific innate and adaptive immune pathways highlighted
- Also role of
  - epigenetic regulation (DNMT3a)
  - pathogen resistance (FUT2)
- Fine mapping in progress to identify causal variants