

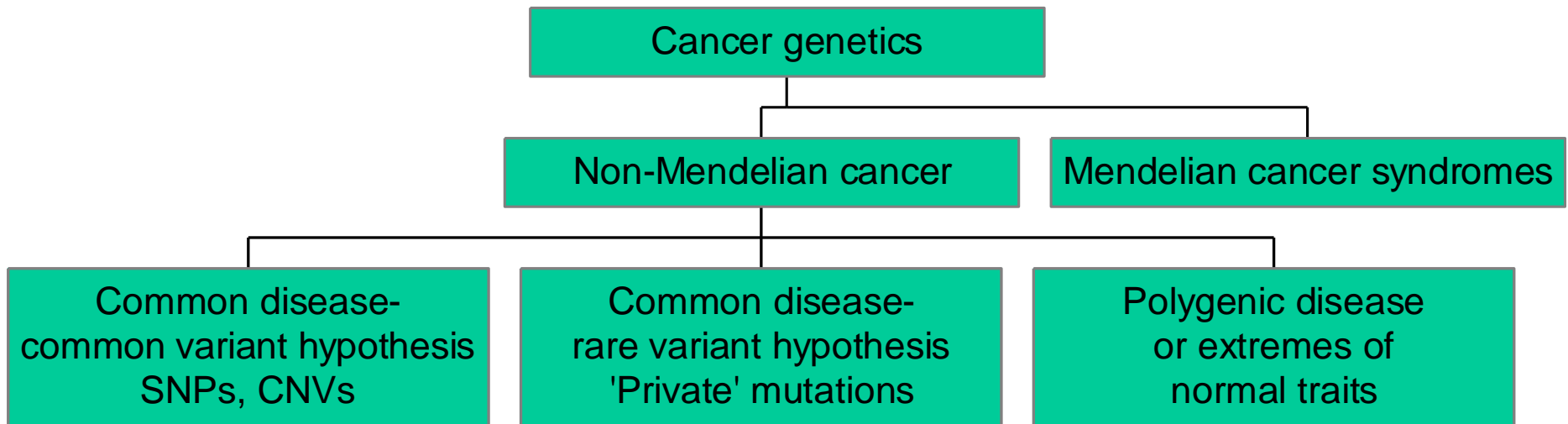
Genome Wide Association Studies and Colorectal Cancer Risk

Ian Tomlinson

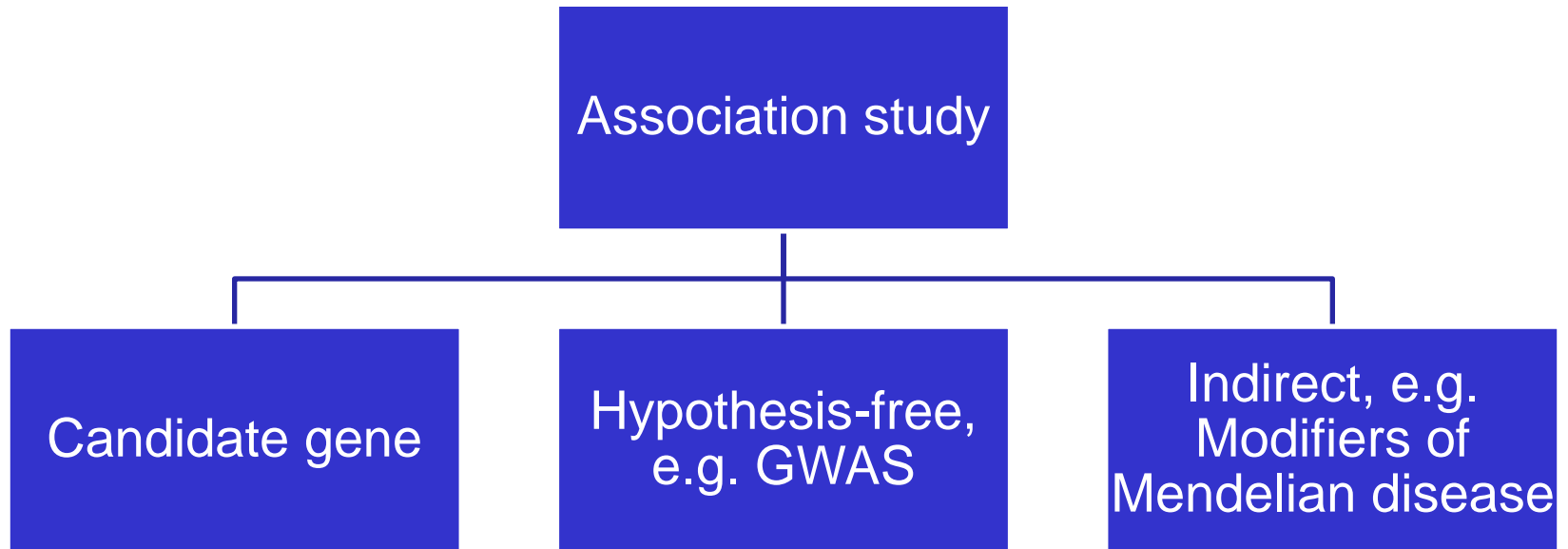
Wellcome Trust Centre for Human Genetics,
University of Oxford



Cancer as a genetic disease



Strategies to identify common cancer predisposition alleles



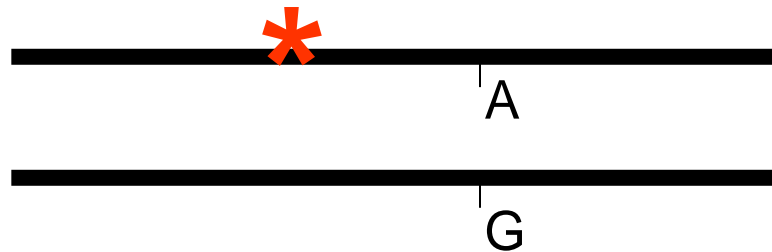
Meta-analysis of candidate CRC genes

- Xenobiotic metabolism
- Iron absorption/metabolism
- Oncogenes, e.g. Ha-ras VNTR rare alleles
- Tumour suppressor genes, e.g. p53
- Folate pathway, e.g. MTHFR C677T
- Signalling pathways, e.g. Wnt, TGF-beta
- Genome stability, e.g. BLM*Ash
- Predisposition genes for other cancers, e.g. CHEK2

Almost ALL negative in large sample sets

The HapMap project

- Catalogue of variation at single nucleotide polymorphisms (SNPs) genome-wide in different populations
- Provides important basic information
- Useful for disease gene identification via linkage disequilibrium mapping
- 'Tag' SNPs can cover whole genome



550K
SNPs

CORGI
1K + 1K

Edinburgh
1K + 1K

VICTOR/QUASAR2
1958 birth cohort
1.2K + 2.5K

50K
SNPs

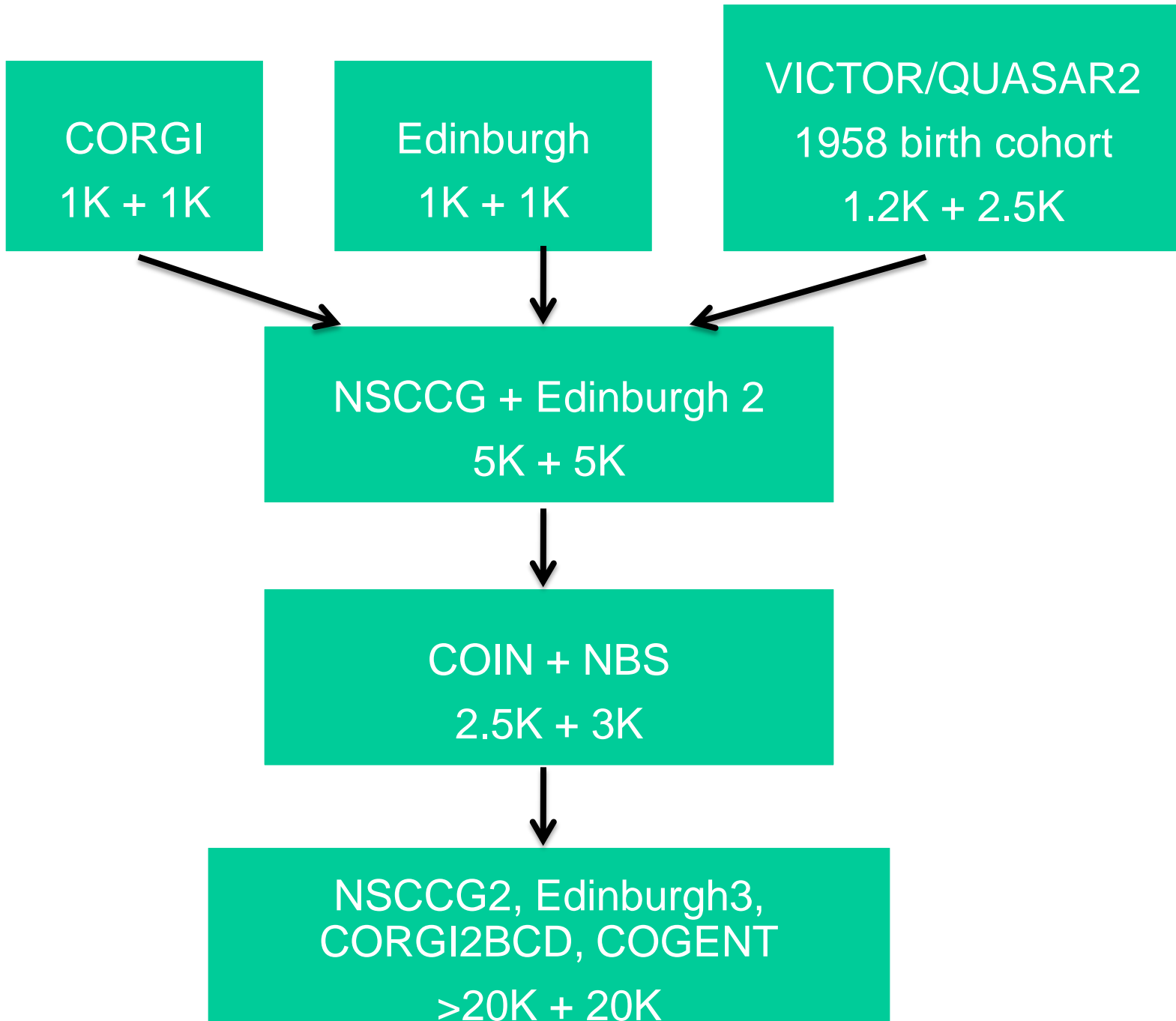
NSCCG + Edinburgh 2
5K + 5K

200
SNPs

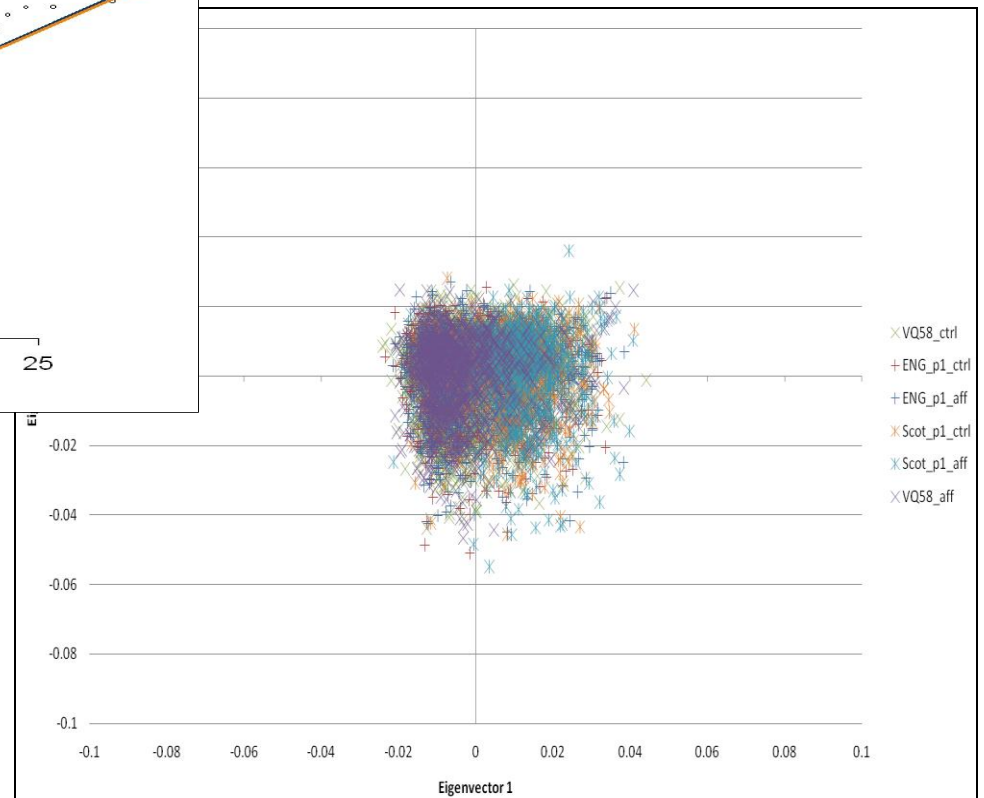
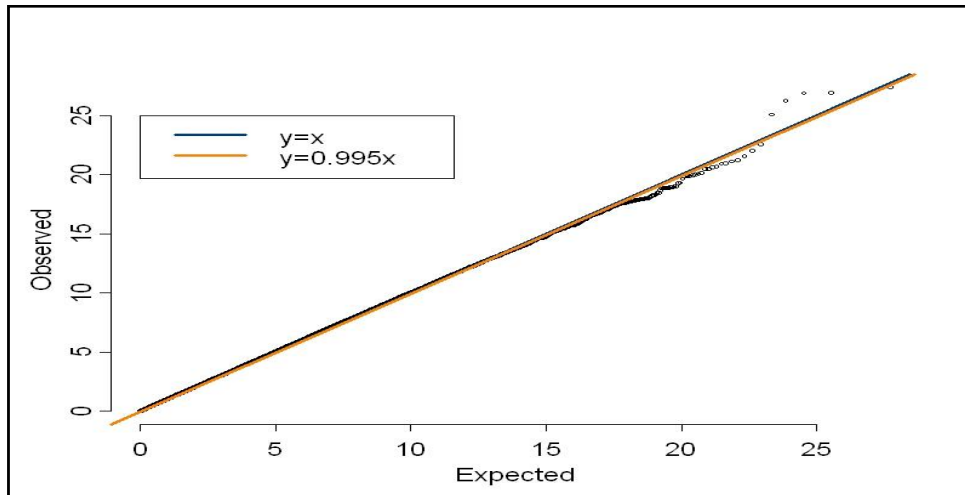
COIN + NBS
2.5K + 3K

~30
SNPs

NSCCG2, Edinburgh3,
CORGI2BCD, COGENT
>20K + 20K

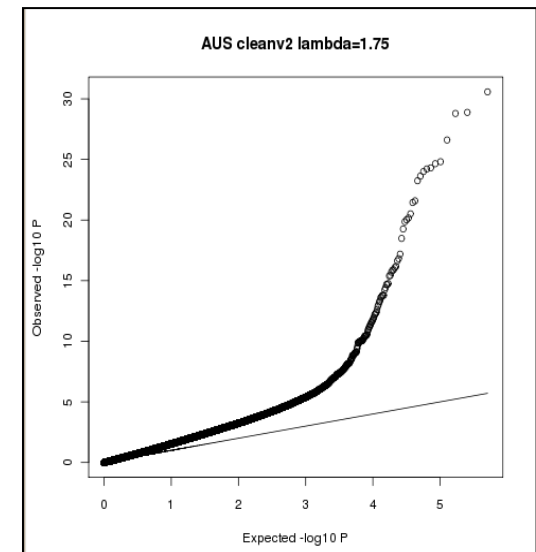
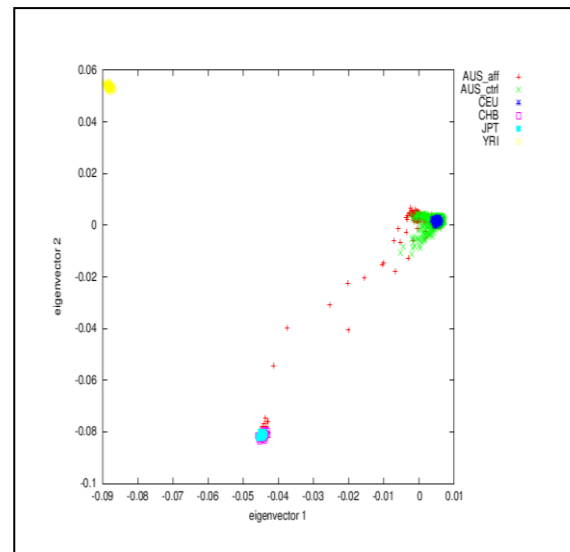
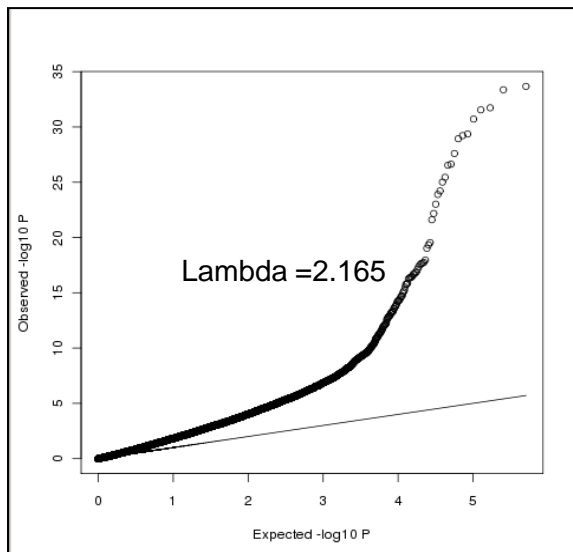


PCA reveals NW-SE cline in UK but cases and controls cluster together



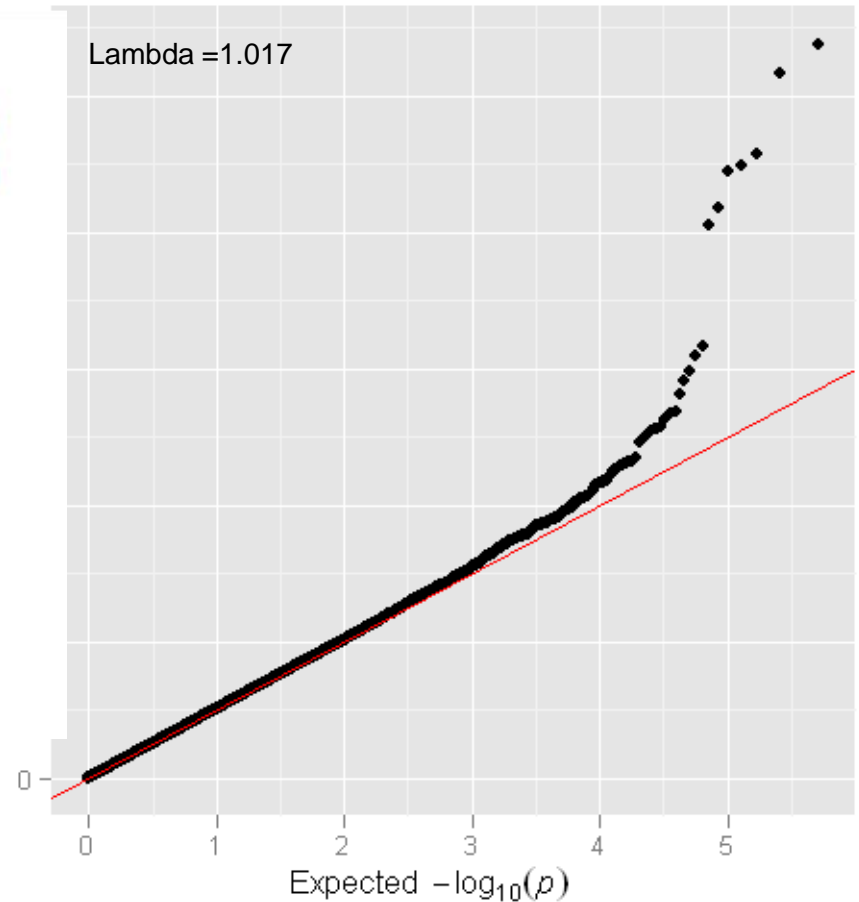
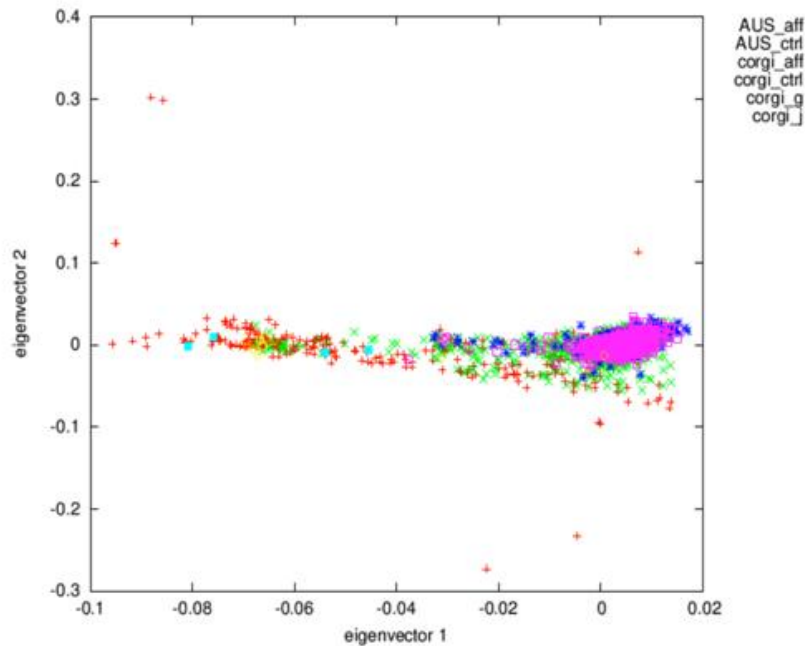
Other sample sets may be more stratified

- Genome-Wide association in Australian cohort reveals severe population stratification
- Need to preserve ethnic homogeneity or matching of cases and controls



Population Stratification II

- Further analysis of the principal components reveal the effect of the diversity of the European population



14 colorectal cancer SNPs

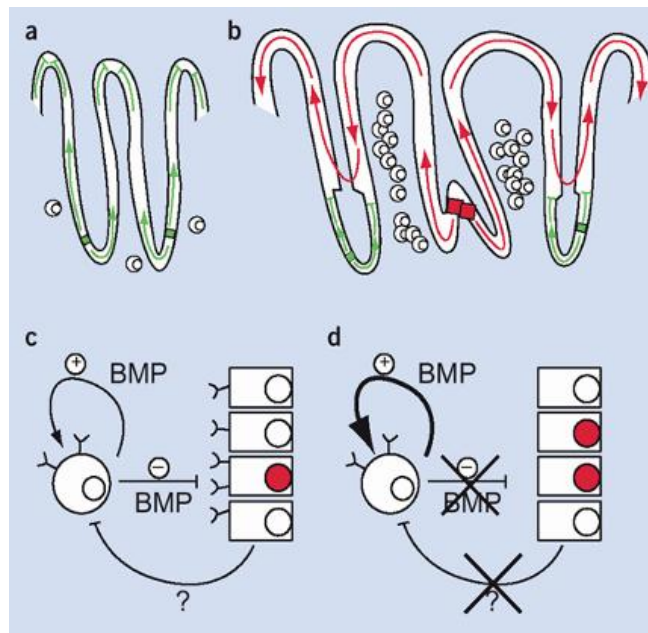
Chr	Nearby genes	OR per allele	Notes
8q24	MYC	1.26	May be functional
18q21.1	SMAD7	1.21	BMP/TGFB pathway
15q14	GREM1	1.20	BMP pathway (rs4779584)
8q23	EIF3H3	1.16	
10p14	None	0.84	Minor allele protective
11q23	POU2AF1	1.17	Rectal cancer assocn
19q13.1	RHPHN1	0.84	In intron
16q22.1	CDH1, CDH3	0.89	Good candidate genes
14q22.2	BMP4	1.12	BMP pathway
20p12.3	BMP2	1.13	BMP pathway
1q41	None	1.09	
3q26.2	MYNN	0.93	
12q13.13	DIP2B	1.06, 0.93	
20q.13.33	LAMA5	0.92	Includes BMP pathway

GWAS suggests most important influence on colorectal cancer risk is the bone morphogenetic protein pathway

Although the functional variants are not known, the probability of 5/14 CRC SNPs being near genes involved in this pathway is $\sim 8 \times 10^{10}$

BMP pathway in the normal colorectum

- BMPs are secreted proteins with vital roles for development and homoeostasis in multiple tissues
- Role in colorectum poorly understood
- May be involved in stem cell niche maintenance

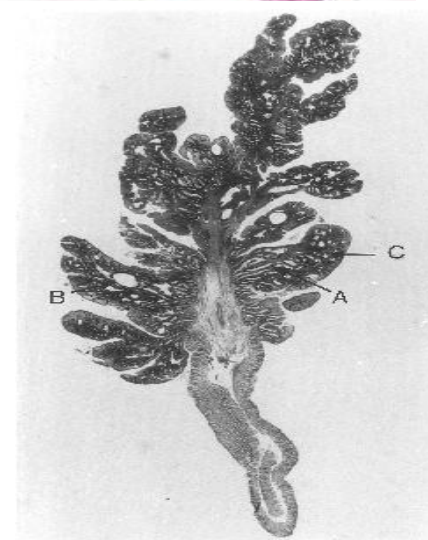
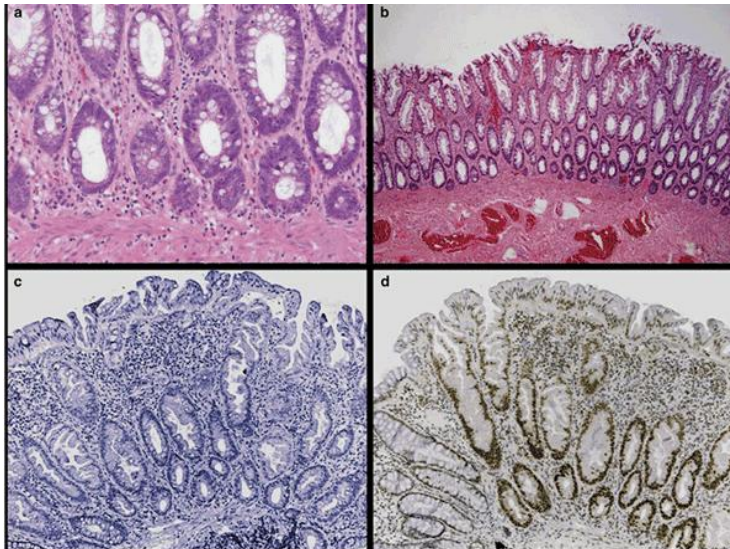
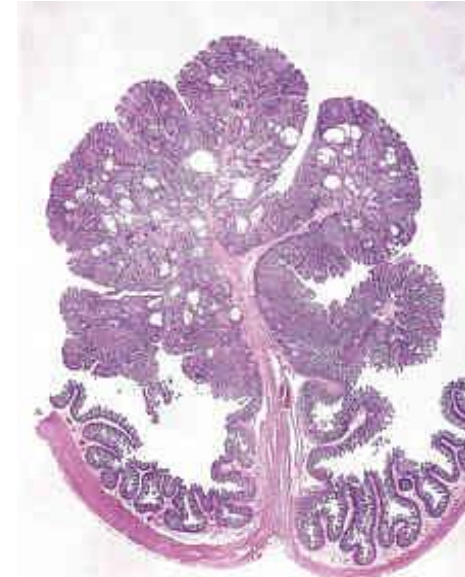
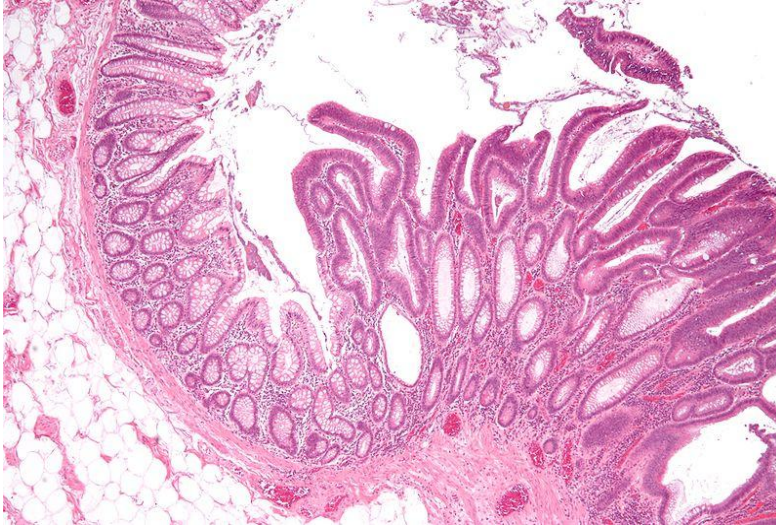


BMP pathway mutations were known to play a role in colorectal cancer

- BMPR1A and SMAD4 mutations cause juvenile polyposis (Mendelian, hamartomatous polyps + CRC)
- Somatic SMAD4 mutations in ~10% CRC

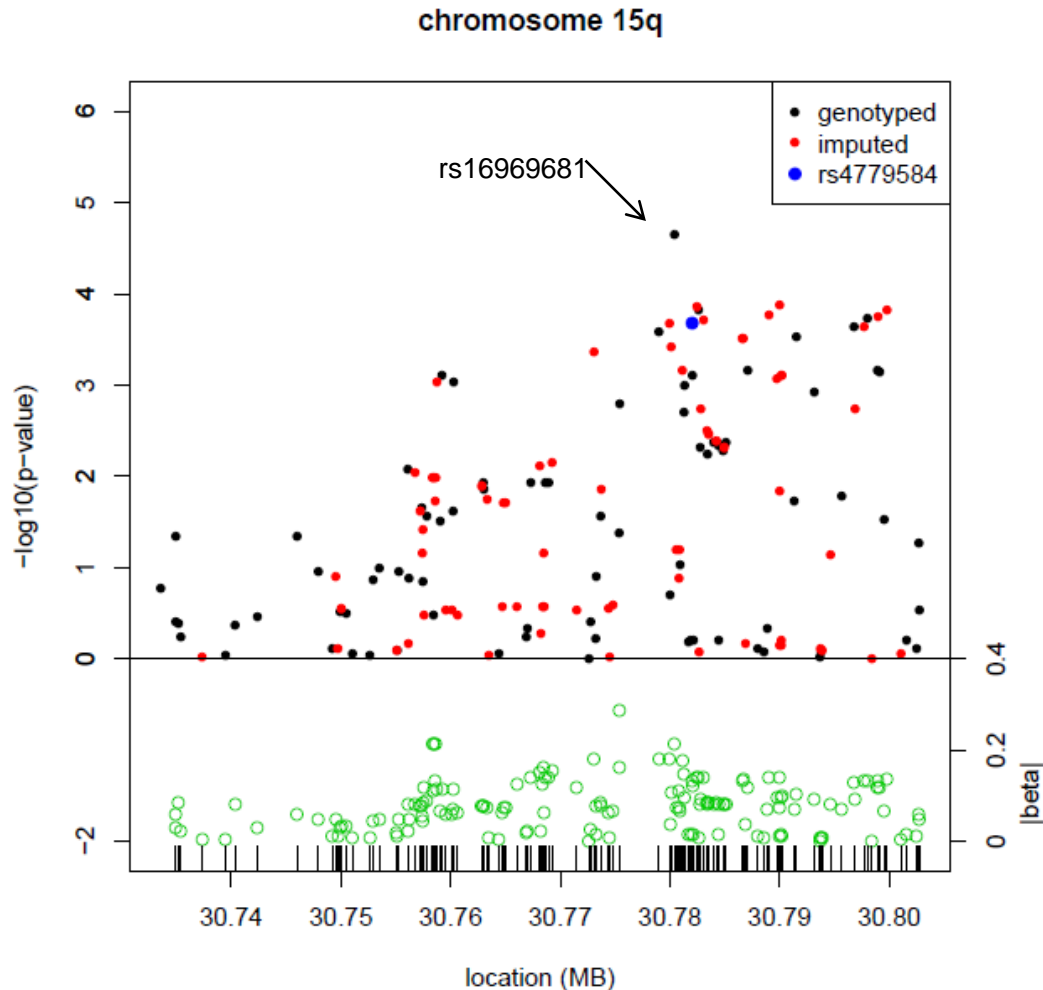


GREM1 is the locus for hereditary mixed polyposis syndrome (HMPS)

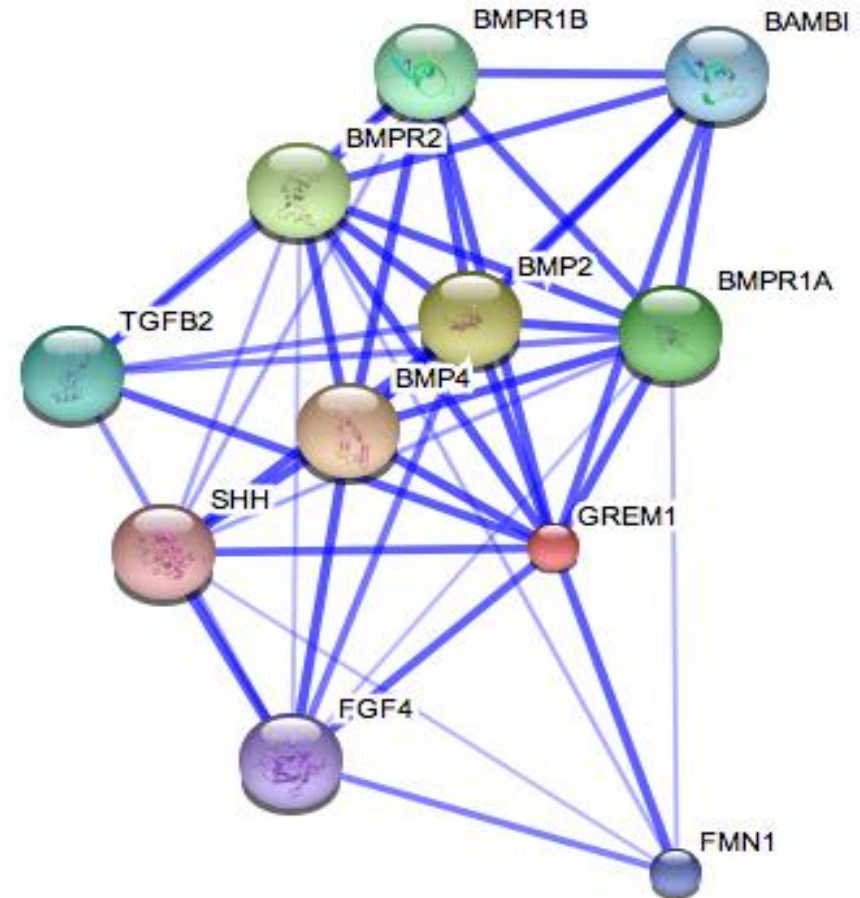
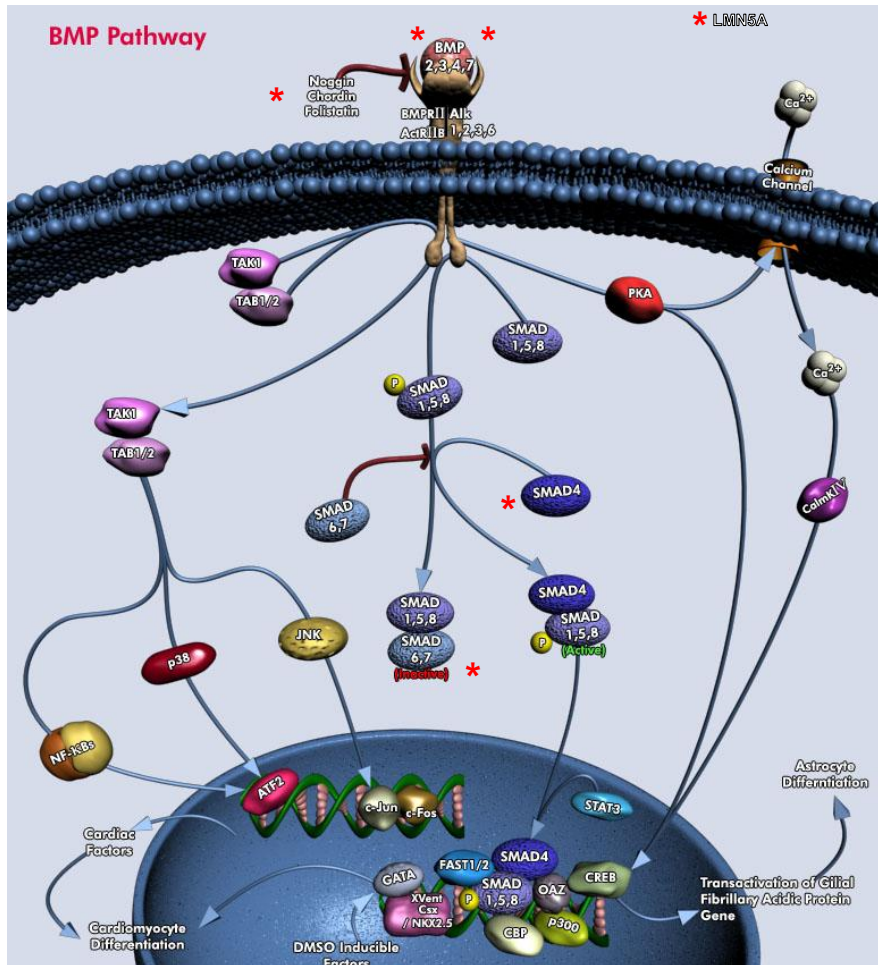


**Some BMP genes harbour
multiple, independent colorectal
cancer susceptibility variants**

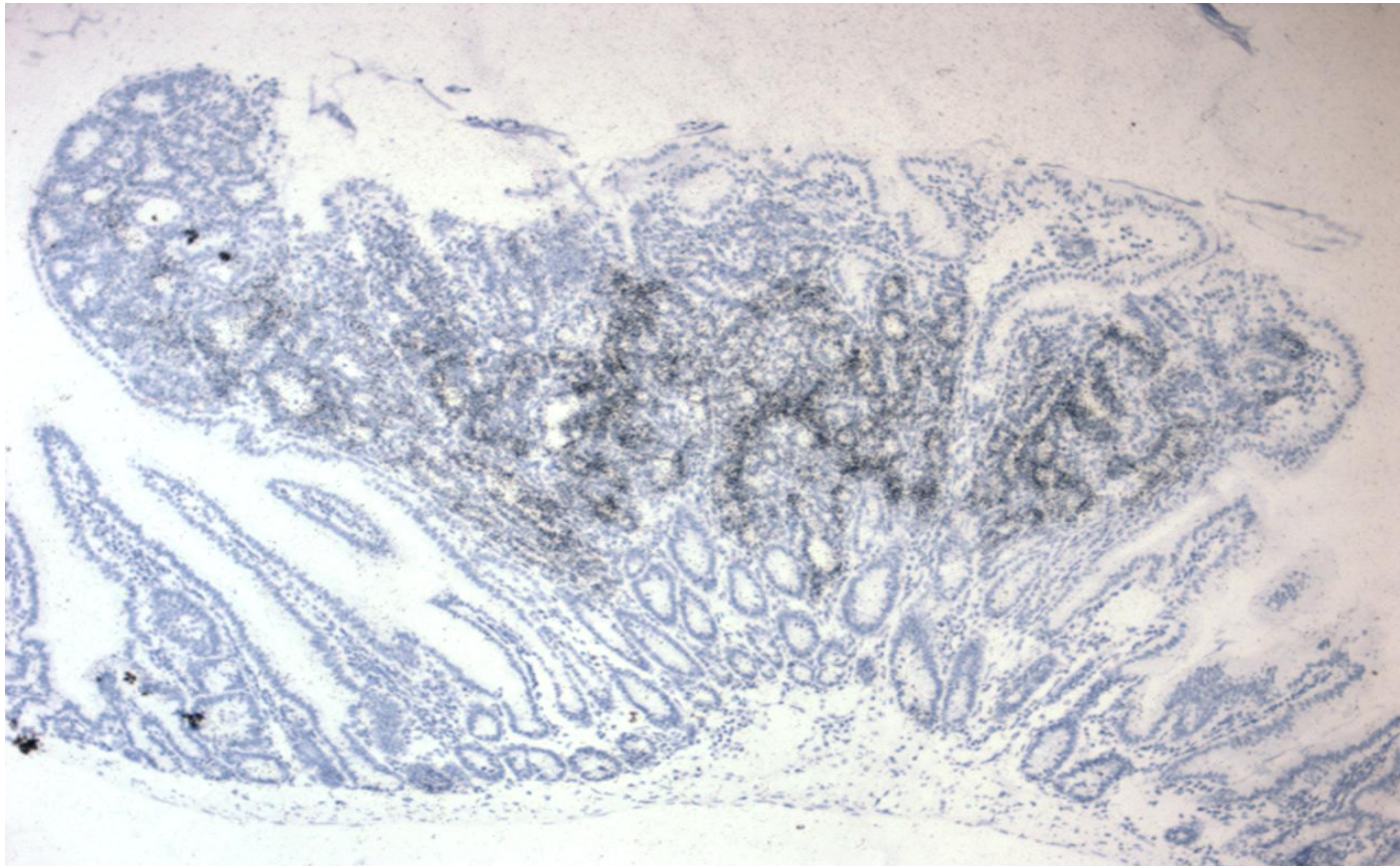
Imputation shows stronger signal at GREM1 than the original tagSNP



Several BMP pathway genes are involved in colorectal cancer predisposition



Can we test hypothesis that BMP variants are associated with different numbers of GI stem cells?



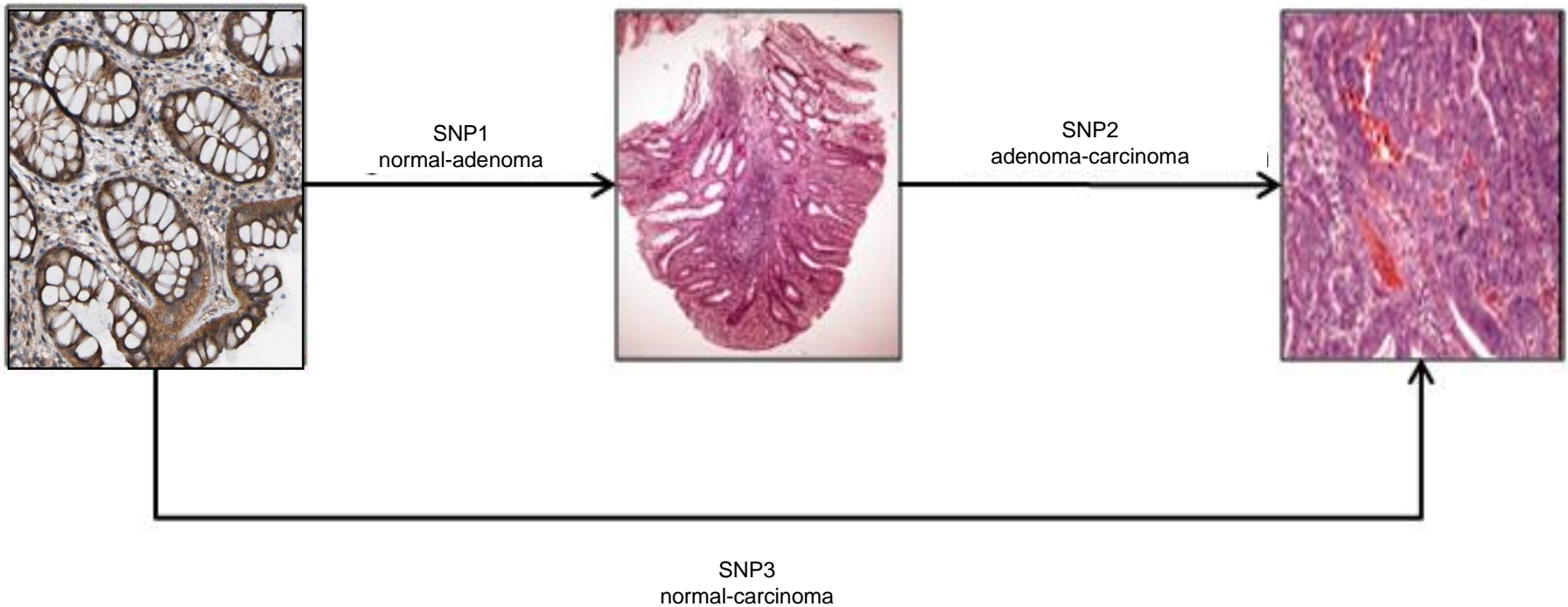
Two more independent colorectal cancer predisposition SNPs at BMP4 and BMP2

- For BMP pathway components, use lower threshold than GWAS norm ($P < 10^{-4}$) to take SNPs into validation phase
- rs1957636 (BMP4), $P = 2.09 \times 10^{-8}$
- rs4813802 (BMP2), $P = 3.93 \times 10^{-10}$
- Not in LD with other SNPs at these sites

GWA studies in colorectal cancer and the BMP pathway: conclusions

- BMP SNPs clearly over-represented in set of CRC-associated SNPs (8/17)
- More than one variant independently affecting disease at some loci. Is this one source of “missing heritability” in GWASs?
- Stem cell number most plausible model, but how to test?

Some CRC SNPs act on adenoma growth, some on carcinoma growth and some on both



Some colorectal cancer SNPs are associated with primary adenoma predisposition

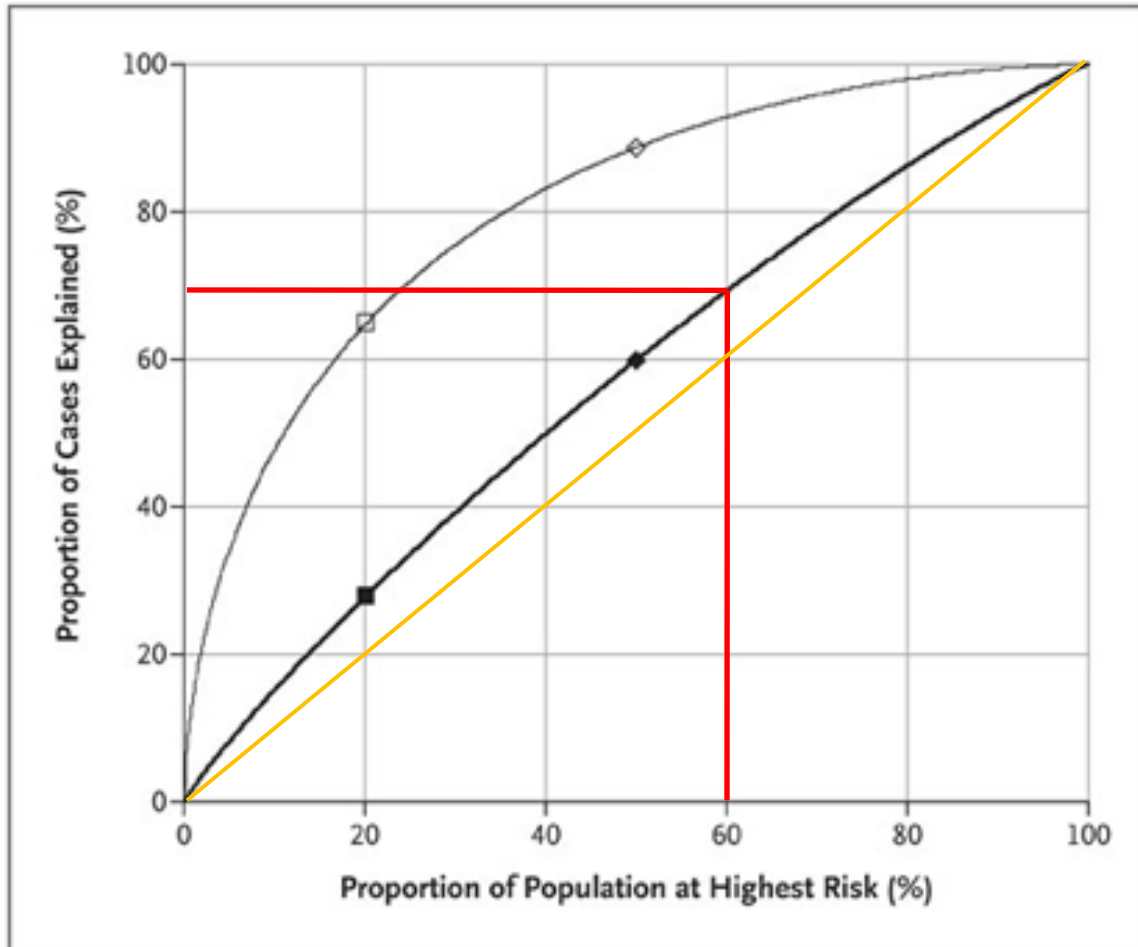
Chr	SNP	Gene?	Adenoma risk?	Carcinoma risk?
10p14	rs10795668	None	Yes	Probably not
20q13.33	rs4925386	LAMA5	Yes	Probably not
14q22.2	rs4444235	BMP4	Yes	Probably not
3q26.2	rs10936599	MYNN	Yes	Yes
16q22.1	rs9929218	CDH1	Yes	Yes
8q24.21	rs6983267	MYC	Yes	Yes
15q14	rs4779584	GREM1	Yes	Yes
18q21.1	rs4939827	SMAD7	Yes	Yes
8q23	rs16892766	EIF3H	Yes	Yes

The other 5 known CRC SNPs only influence carcinoma predisposition

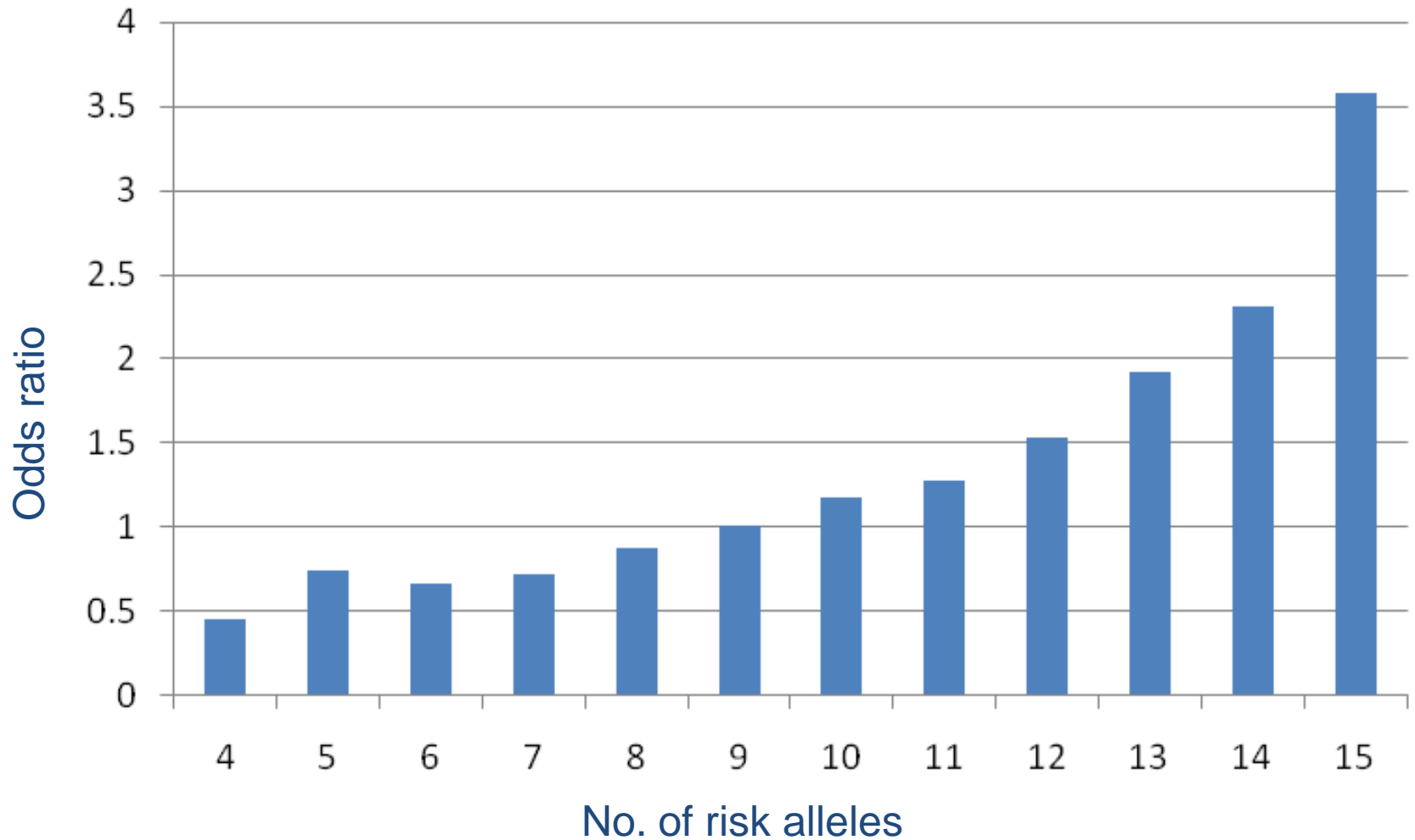
Lessons from cancer GWASs

- Common disease-common variant model applies to cancer
- Low risks (max ~1.3 per allele) and power
- ?multiplicative risks (?), no GxGs to date
- Candidate and/or nsSNPs generally absent, although our view of top candidates has changed
- Seem to be several different underlying mechanisms of increased risk
- Do most variants act via (long-range) changes in gene expression?

We cannot (yet) predict colorectal cancer risk for an individual with any degree of precision using common predisposition genes



But at the population level, relative risks are already quite impressive



Where do we go next? Rare alleles with low/moderate penetrance?

PERSPECTIVE

nature
genetics

Common and rare variants in multifactorial susceptibility to common diseases

Walter Bodmer & Carolina Bonilla

Yes, but issues of

- low power
- genotype calling
- pooling variants
- choice of candidates v genome-wide, cost
- utility for risk prediction

Are GWAS levels of significance needed?

Acknowledgements

Oxford lab.

- Luis Carvajal-Carmona
- Angela Jones
- Emma Jaeger
- Kimberley Howarth
- Enric Domingo
- Mike Churchman
- Ella Barclay
- Maggie Gorman
- Lynn Martin
- Jean-Baptiste Cazier

Others

- Houlston lab
- Dunlop lab
- All CORGI centres
- VICTOR and QUASAR2 trials
- COGENT consortium
- Paul Pharoah
- Jerry Cheadle, COIN
- APC and APPROVe trials

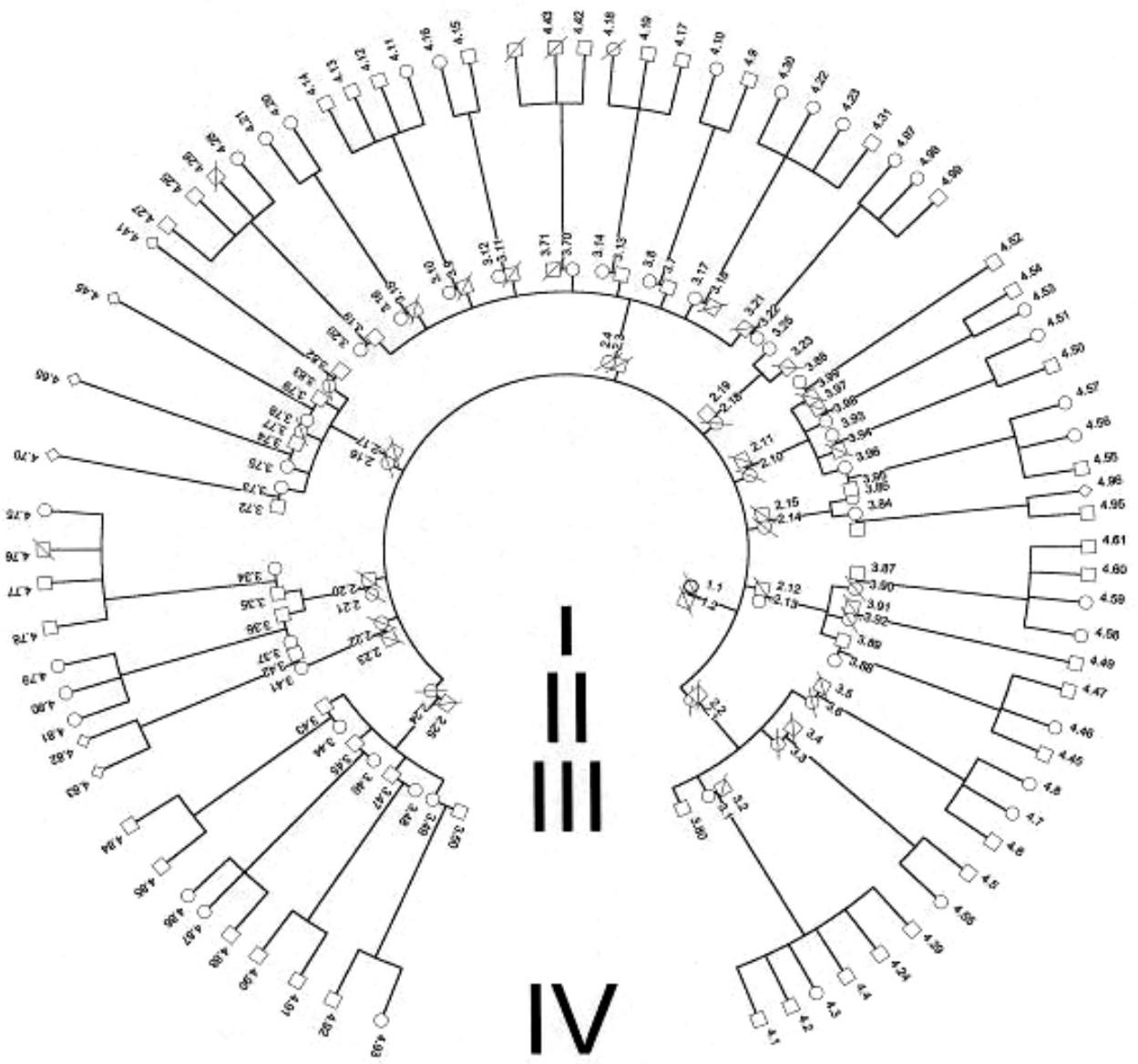
QC Issues

- <1% samples failed
- SNP call rate >99.5%
- All duplicate calls identical
- Eliminate sex mismatches
- Eliminate IBD predicted >6.25%
- Eliminate excess heterozygosity
- Eliminate samples clustering with non-White European



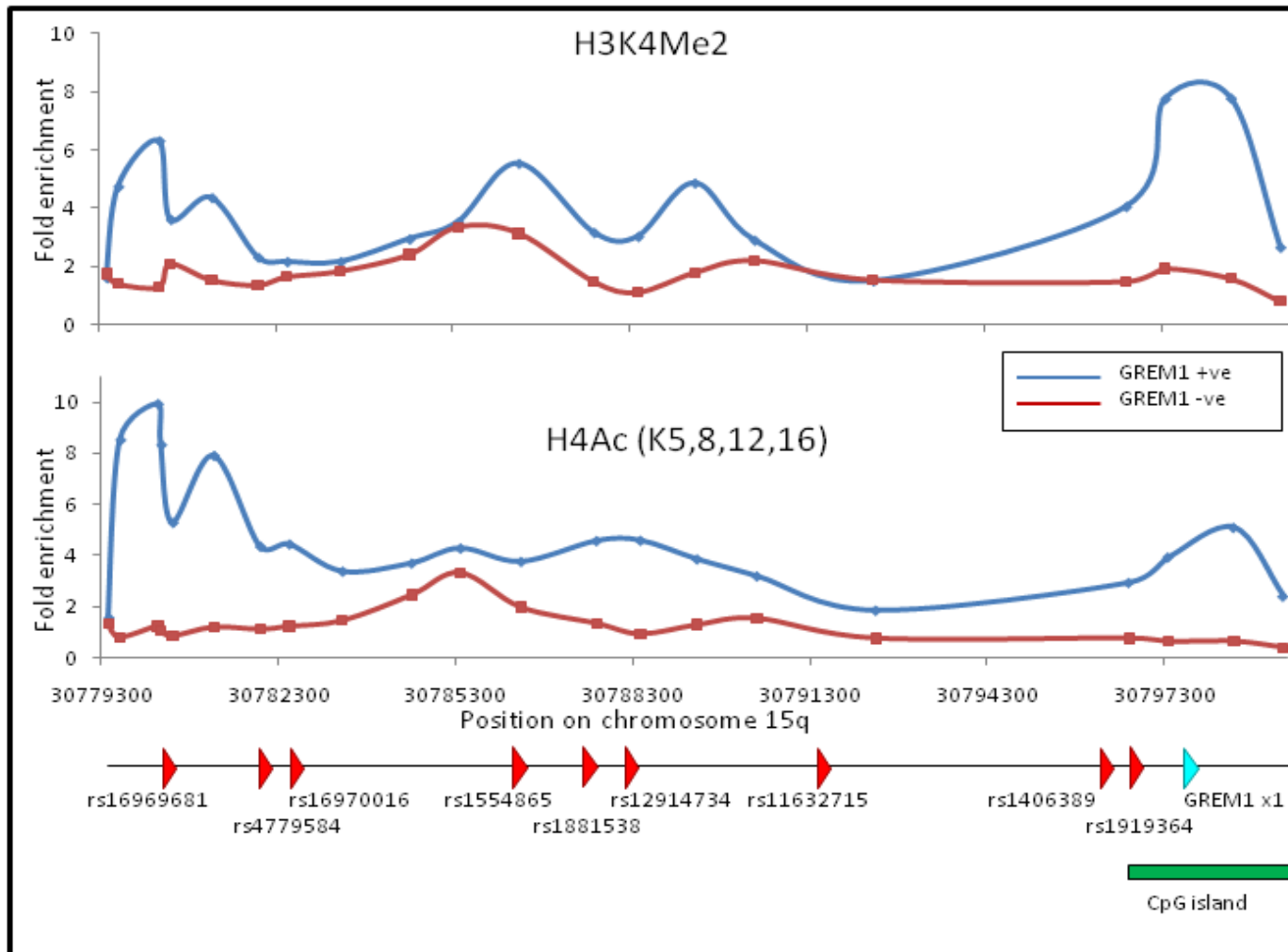
GREM1 is the locus for hereditary mixed polyposis syndrome (HMPS)

- Families of Ashkenazi origin
- Multiple (up to 40) polyps: essentially adenomas (some serrated)
- CRC
- No extra-colonic features
- Autosomal dominant

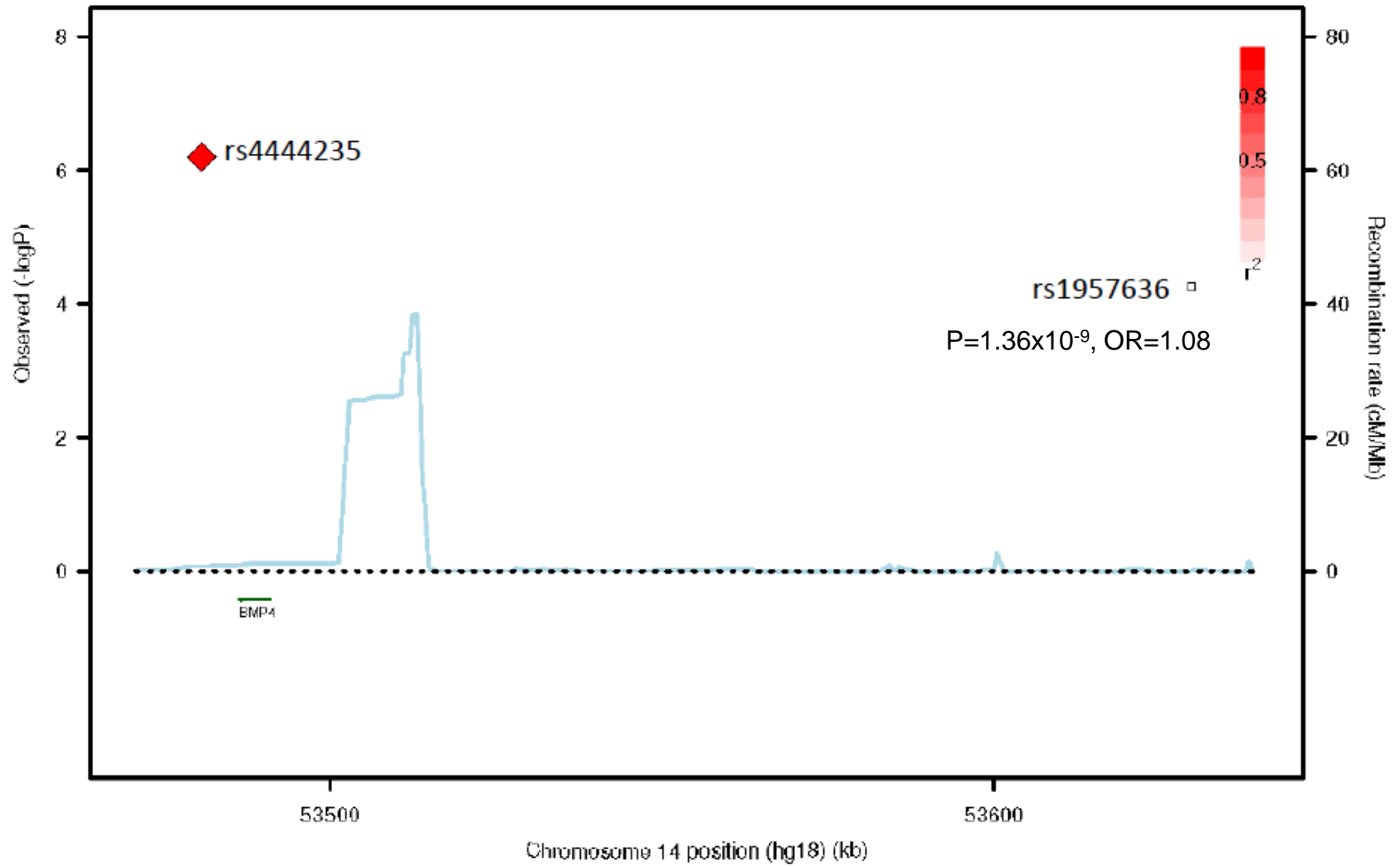


The original HMP5 Family SM96

rs16969681 may be functional



BMP4 region



BMP2 region

