

The genetic basis of human diseases

Professor Patricia Munroe
Clinical Pharmacology and Precision Medicine
William Harvey Research Institute, FMD



Genes and Environment in Disease

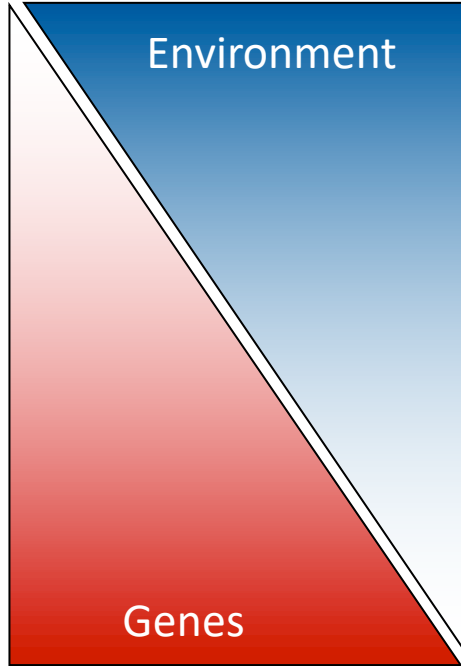
Infectious Disease

Environment

Hypertension
Obesity
Type 2 diabetes

Monogenic

Genes



Outline of lecture

- Introduction to monogenic and complex genetic diseases
- Methods to find genes for monogenic and complex traits
- Exemplar research project
 - Genetic basis of hypertension

Monogenic versus Complex

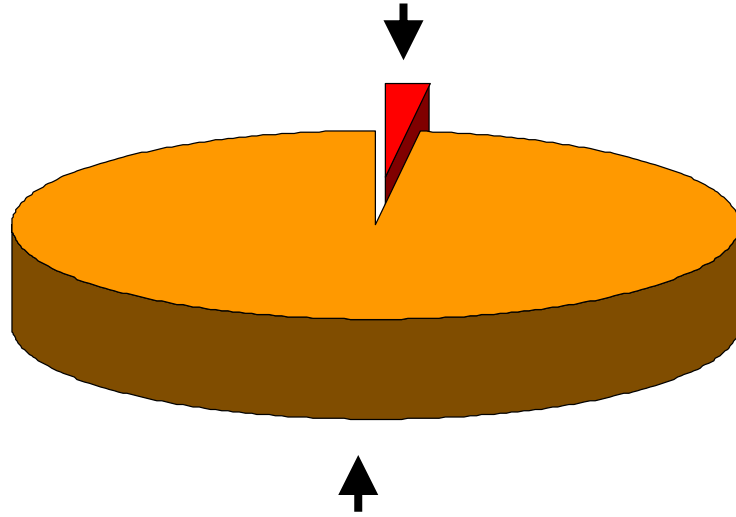
- **Mendelian Disease**

- Rare disease
- Single gene
- Mutations have low frequencies in the population
- Mutations have large effects on gene function
- High penetrance
- Autosomal dominant, recessive or X-linked
- Minimal to mild environmental influences

- **Complex Disease**

- Common disease
- Polygenic
- Gene variants have high frequencies in the population
- Gene variants have small effects on gene function
- Low penetrance
- Unclear modes of inheritance
- Significant environmental influences

There are 'monogenic' forms of most complex diseases but these only account for a small proportion of cases in the population



(complex) - polygenic disease

How do you determine if genes are important?

- Genetic factors
 - Familial aggregation
 - Twin studies
 - Adoption studies

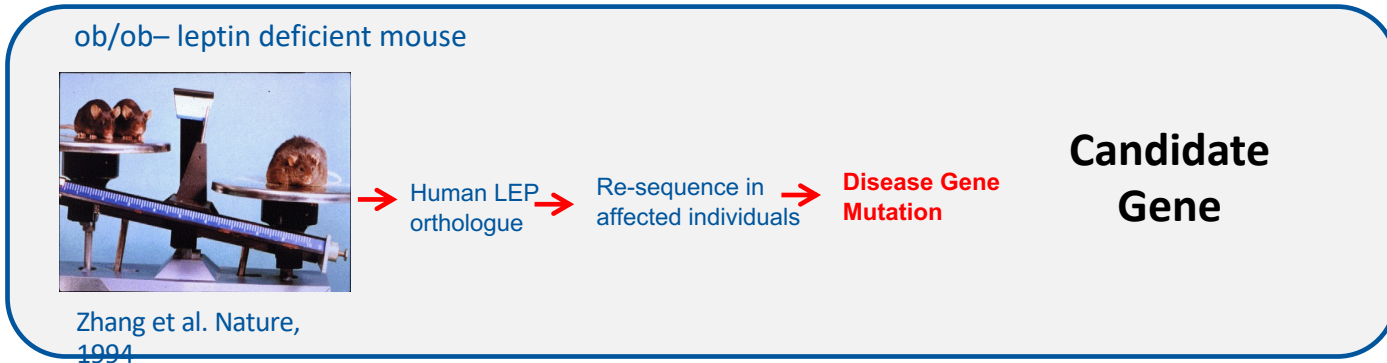
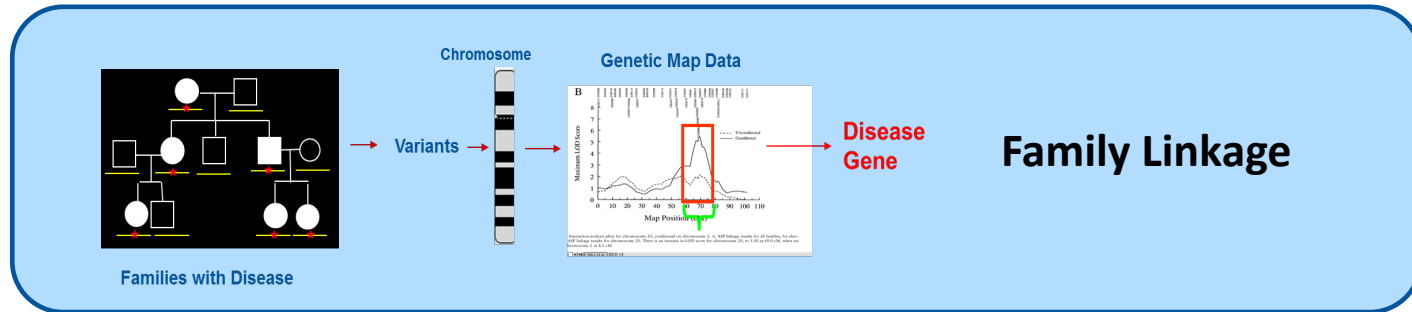
Benefits of gene identification for monogenic and complex diseases

- Better understanding of aetiology and pathogenesis of disease
- Identification of at-risk individuals for early diagnosis, prognosis and treatment
- Molecular targets for developing novel therapeutics

How do you find genes causing disease?

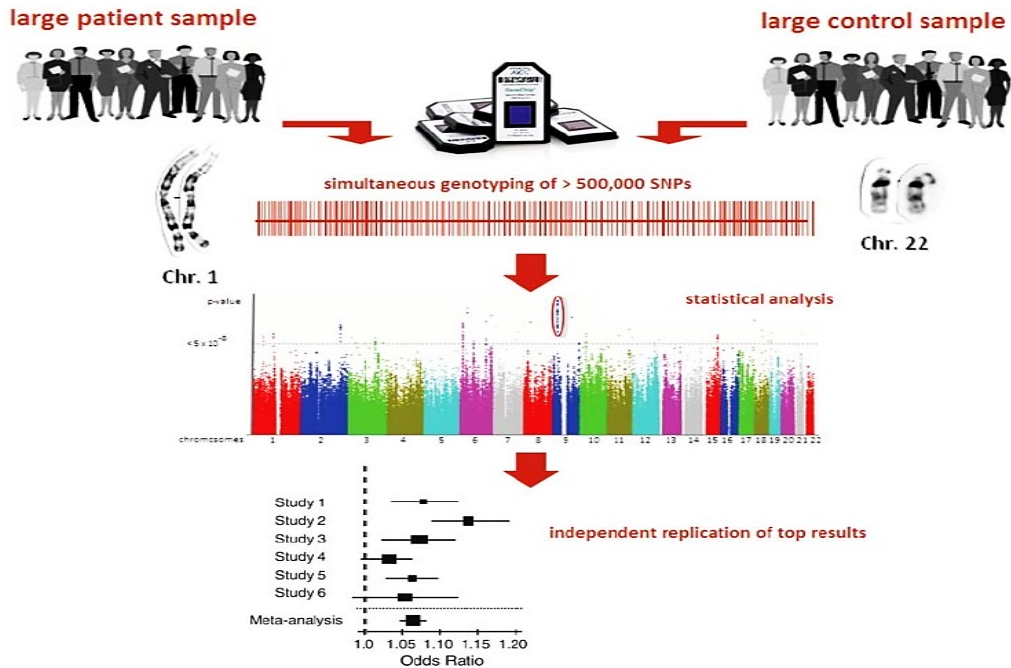
- **Monogenic disease**
 - Linkage analysis
 - Exome sequencing and now whole genome sequencing
- **Complex disease**
 - Historical - candidate gene approach
 - Physiology/Pharmacology, Rare single gene disorders, Experimental models
 - Contemporary - Genome-wide scanning
 - Genome-wide association studies (GWAS) in unrelated individuals (array-based genotyping)
 - Whole exomes and genomes

Genetic Approaches I

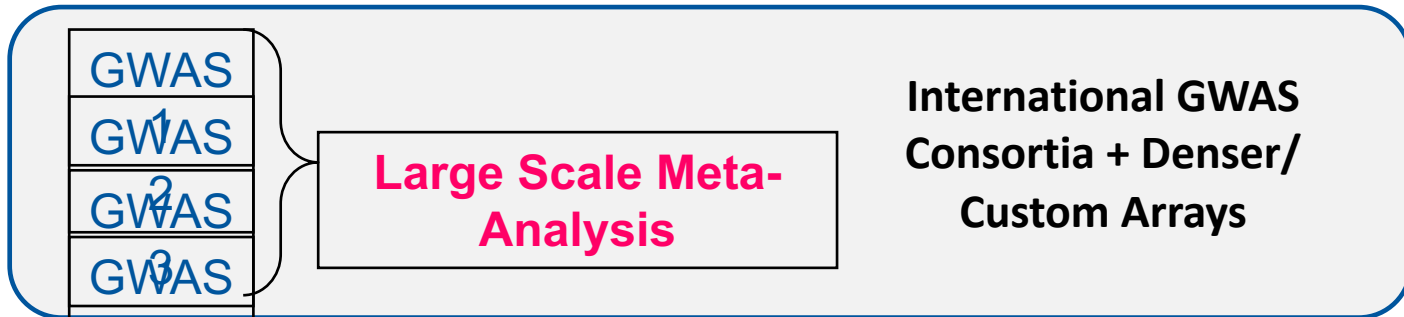
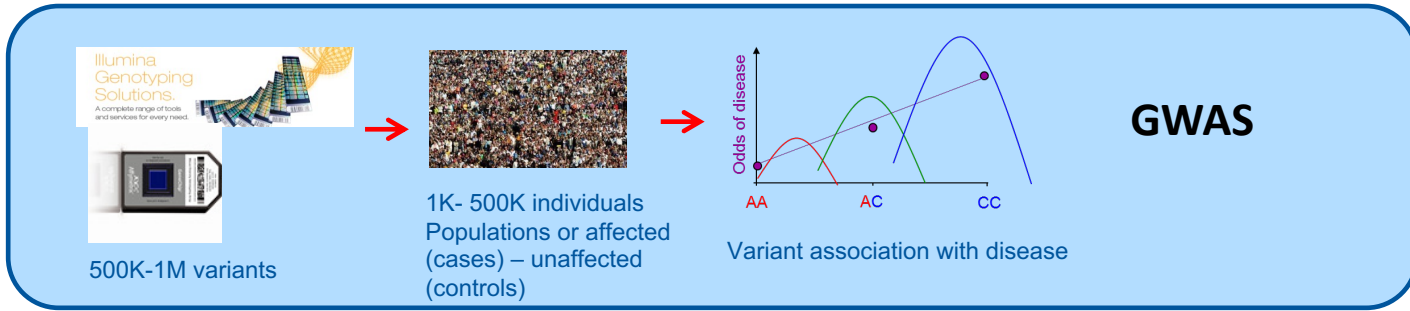


Slide adapted from a presentation by Ines Barroso, Sanger Institute

Genome-wide association study

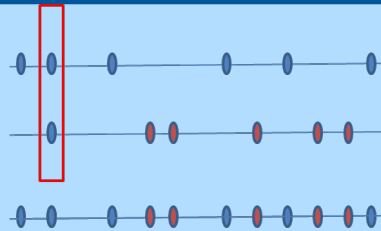


Genetic approaches II



Slide adapted from a presentation by Ines Barroso, Sanger Institute

Genetic Approaches III

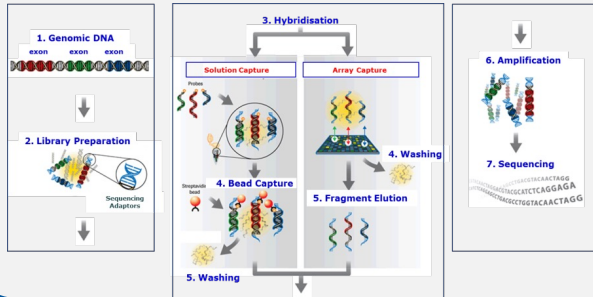


Study 1

Study 2

Study 1 with imputed missing SNPs

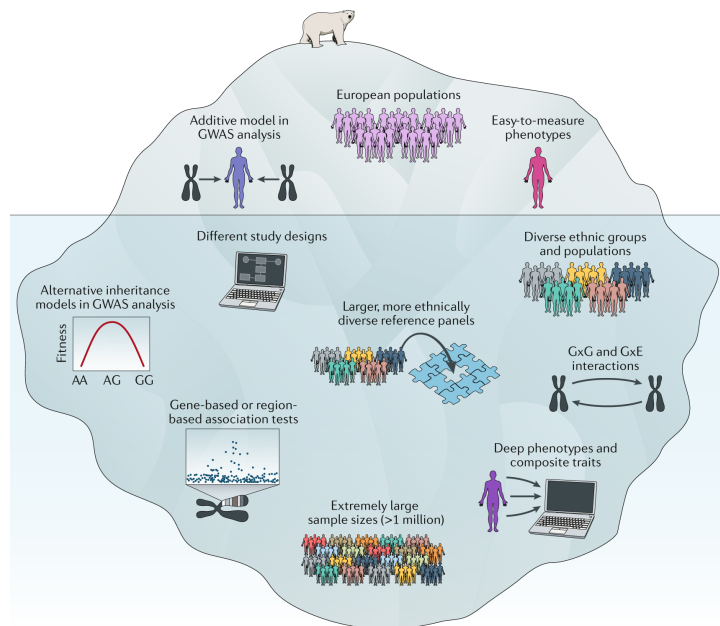
GWAS+ Imputation Dense Reference Panel



Whole-Exome and Whole-Genome Sequencing

Slide adapted from presentation by Ines Barroso, Sanger Institute

GWAS - Success story for gene discovery

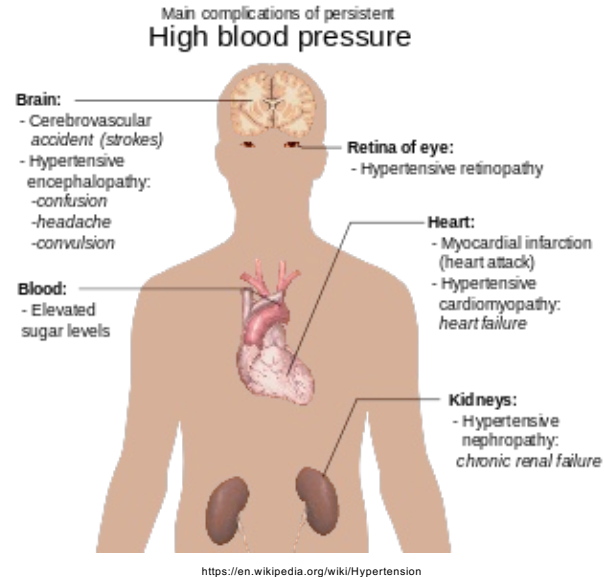


11/3/2023 - GWAS catalog latest figures on associations: 6306 publications and 493,105 associations.

Downstream applications: new biology, new therapeutic targets and polygenic risk scores

Hypertension

- Hypertension
 - SBP \geq 140mmHg and DBP \geq 90mmHg
- 1 billion hypertensives worldwide
 - Causes 4.5% of global disease burden
- Risk factor for CVD



Types of hypertension

- Primary hypertension (90%)
 - 30 - 50% - genes
 - Twins and family studies
 - 50 -70% lifestyle factors
 - Dietary salt intake, alcohol consumption, weight, inactivity etc
- Secondary hypertension (~10%)
 - Primary aldosteronism (Conns syndrome), phaeochromocytoma, thyroid disease, side effects of medications
 - **Rare monogenic syndromes (~1%)**

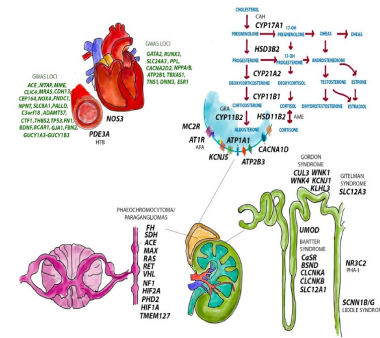
Rare monogenic forms of Hypertension

Monogenic HTN disorders

- 11 single gene hypertensive disorders known - seven now characterised at functional level
- Genetic linkage and exome sequencing - for discoveries over 20 years
- Rare conditions
- Important - insights into process of BP control

Monogenic forms of hypertension

Trait	Mode of Inheritance	Location	Gene
Glucocorticoid Remediable Aldosteronism	Autosomal dominant	8q24.3	11 β -hydroxylase/aldosterone synthase (CYP11B1/CYP11B2) chimera
Apparent Mineralocorticoid Excess	Autosomal recessive	16q22.1	11 β -hydroxysteroid dehydrogenase (HSD11B2)
Gordon's syndrome	Autosomal dominant and Autosomal recessive	1q, 2q, 5q, 12p, and 17q	Protein kinases, lysine deficient 1 and 4 (WNK 1, WNK 4)
			Kelch-like 3 (KLH3)
			Cullin 3 (CUL3)
Liddle's syndrome	Autosomal dominant	16q	Sodium channel non-voltage-gated 1, β and γ subunits (SCNN1B, SCNN1G)
Hypertension with brachydactyly	Autosomal dominant	12p	Phosphodiesterase 3A (PDE3A)



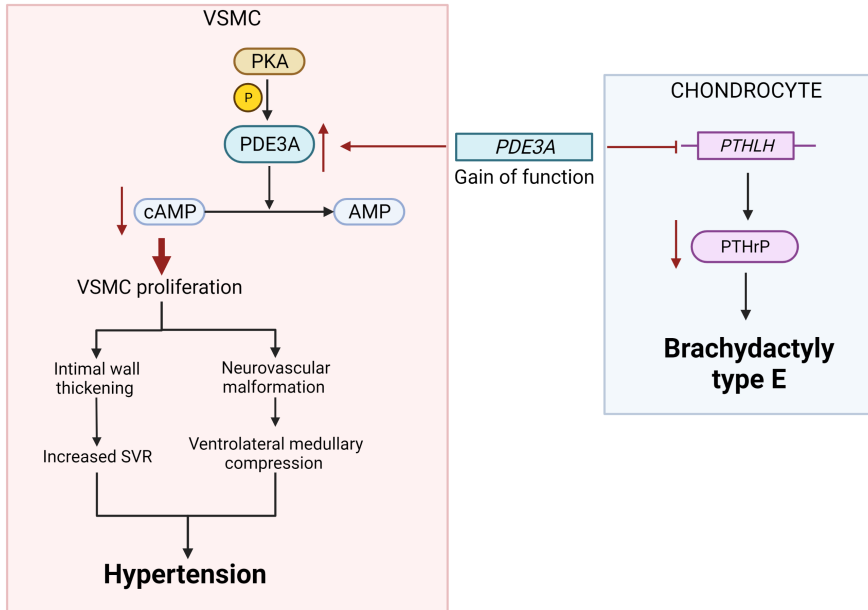
Hypertension with brachydactyly

- Also known as Bilginturan syndrome (HTNB; OMIM: 112410)
- Salt independent HTN increasing with age
- Short stature
- Autosomal dominant
- Impaired baroreflex sensitivity
- Very rare - prevalence <0.0001%

HTNB genetics

- 6 Turkish pedigrees - heterozygous mutation in PDE3A gene - gain of function
- Other families and mutations also described and sporadic mutations in gene
- Mutations cluster in exon 4

HTNB pathophysiology



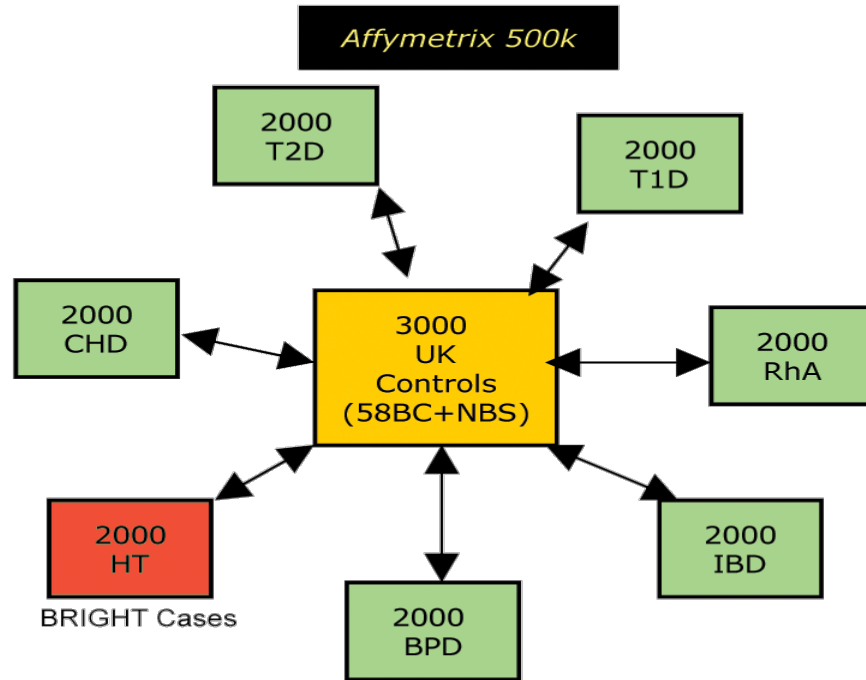
- Phenotype explanation
- In VSMC - PDE3A sensitivity to PKA, decrease in cAMP levels
- VSMC proliferation - sequelae - HTN
- In chondrocytes - lead to BDE - via downregulation of PTHLH gene decreasing PTHrP

HTNB Diagnosis and Management

- Resembles essential hypertension - diagnosis is difficult!
- There is the bracydactyly - but other symptoms than BP absent - neurovascular malformations - magnetic resonance angiography required.
- Genetic testing for mutations - NECESSARY
- Responsive to all BP medications

BP gene discovery using the *GWAS* approach

First GWAS for Hypertension



Wellcome Trust Case Control Consortium

WTCCC, Nature 2007.

WTCCC and hypertension

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

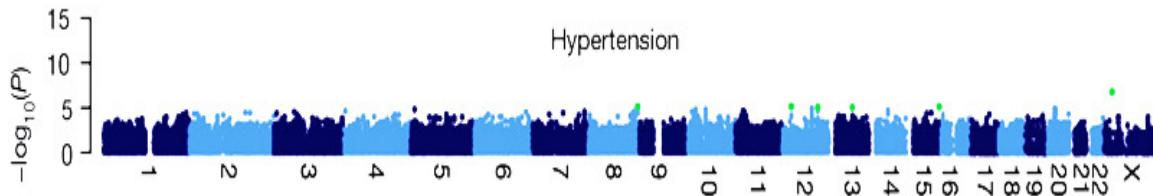
nature

ARTICLES

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

The Wellcome Trust Case Control Consortium*

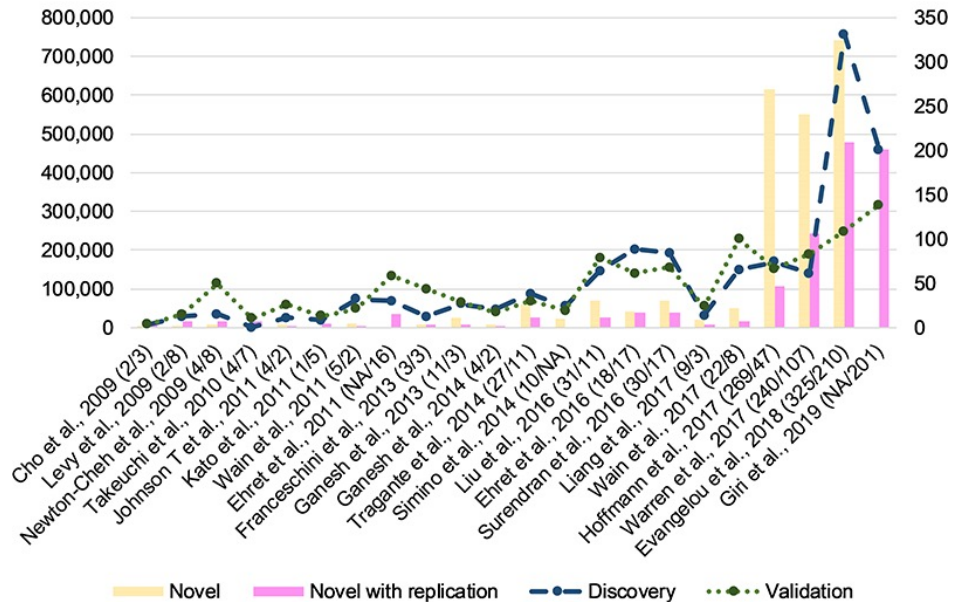
There is increasing evidence that genome-wide association (GWA) studies represent a powerful approach to the identification of genes involved in common human diseases. We describe a joint GWA study (using the Affymetrix GeneChip 500K Mapping Array Set) undertaken in the British population, which has 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 diabetes and 3 in type 2 diabetes. On the basis of prior findings and replication studies thus-far completed, almost all of these signals reflect genuine susceptibility effects. We observed association at many previously identified loci, and found compelling evidence that some loci confer risk for more than one of the diseases studied. Across all diseases, we identified a large number of further signals (including 58 loci with single-point P values between 10^{-5} and 5×10^{-7}) likely to yield



WTCCC, Nature 2007.

BP loci discovery 2009 - 2019

Blood Pressure Genetic Associations Timeline



2018 - the 1M GWAS



UK Biobank (UKB) → N=458,577
Samples excluded from QC of genetic & phenotypic data
Restricted to Europeans using PCA clustering
UKB GWAS of HRC-imputed SNPs using BOLT-LMM
BP ~ SNP + sex + age + age² + BMI + array
→ LD Score Regression → GC-adjustment

ICBP-GWAS → N=299,024 from 77 different cohorts
• N=150,134 (54 cohorts) from published ICBP-1000G
• N=148,890 from 23 new cohorts
Study-level QC and GC-adjustment
Fixed-effects inverse variance weighted meta-analysis
Stringent meta-level QC filtering of SNPs

3 GWAS:
SBP
DBP
PP
(med-adjusted)

Wain et al, *Hypertension*
(2017)



**UKB+ICBP GWAS Discovery
Meta-analysis (N=757,601)**
Analysing ~7 million SNPs
(MAF ≥ 1%) within both UKB & ICBP

Two stage design

Exclude all SNPs in 274 known BP loci, using SNPs previously reported

Follow-up sentinel SNPs with $P < 1 \times 10^{-6}$ for any BP trait

Independent Replication meta-analysis

in two studies MVP (N=220,520) and EGCUT (N=28,742)

→ combined meta-analysis (N=1,006,863)

- (i) genome-wide significant ($P < 5 \times 10^{-8}$) in combined meta
- (ii) $P < 0.01$ in replication meta-analysis
- (iii) concordant direction of effect

One-Stage Design

Consider any novel sentinel lookup SNPs which do not replicate from the 2-stage analysis



UKB-ICBP Internal Replication

- (i) $P < 5 \times 10^{-9}$ from UKB+ICBP discovery meta
- (ii) $P < 0.01$ in UKB GWAS
- (iii) $P < 0.01$ in ICBP GWAS meta-analysis
- (iv) concordant direction of effect UKB vs ICBP

535 novel BP loci

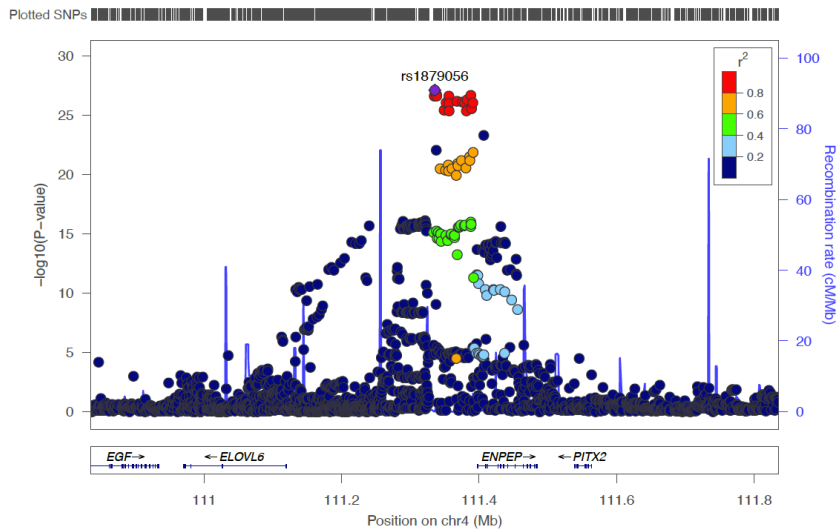
	<p>SBP</p> <p>(65 2-stage): ADORA1, APOLD1, ARHGAP29, ARIH2, ARL14EP, BCAR3, BCAS3, C11orf24, C1orf172, CCT6A, CITED2, CNTN3, CTC-228N24.3, DENND2A, DGKH, ERBB4, FAM193A, FAM208B, FOXC1, GABRA2, GDF2, GLIS3, HSPA12A, IER5L, IRF6, JAZF1, KAT2B, LINC00311, MCM9, MERTK, MLF1, NRXN1, PCCB, PLXNB2, POM121C, PRKD1, RARRES2, RBFOX1, RBM26, RBMS1, RNASEH2B, RP11-1055B8.6, RTN4, SEMA4A, SKI, SLC30A5, SOX5, SPIB, SREK1, ST5, SYT1, SZT2, TARS, TFCP2L1, THSD7B, TMEM108, TNKS, TOX, TRIP12, WDR7, WNT4, XPR1, ZBTB20, ZNF804A, ZSWIM2</p> <p>(32 1-stage): AC009120.4, AC010967.2, AGBL4, AHRR, APOH, BANP, CAND1, CBWD1, CLDN23, CLN8, DFNA5, DMRTA1, FBRSL1, FGR, FOXF1, FOXO6, GRIN2B, HTRA1, KANK1, LINC01091, LRBA, OLA1, PDE11A, PEPD, PKN2, PREX2, PRR20A, RGM8, RP11-122C21.1, RP11-428C19.4, snoU13, WASF3</p>	<p>(57 2-stage): ABCC9, AC004156.3, AC017083.3, AL163953.3, AP000721.4, APOE, ARHGEF25, CDKALI, CELF2, CENPP, COG5, COL15A1, CRB1, CTNND1, EBF1, EDN1, EPB41L2, ERAL1, FGF9, FZD2, GIPR, H1FN1, IGF1, INPP5A, KCNQ5, KDM4B, LCOLR, LNS4, LRCH1, LTBP2, MEIS1, MSRA, MYO1E, NCOA7, NEK6, NNT, NTS1B, OSBP17, PBX3, PDE8A, PHACTR4, PHTF2, PPP2R2D, PRR16, RADS2, RP11-158M2.4, RP11-89M16.1, SLC30A10, TGFBR2, THADA, TRHR, TROBP-NOL12, TSPAN14, WHSC1L1, ZCCHC2, ZMAT2, ZNF618</p> <p>(17 1-stage): CDYL2, CWC27, FAM46A, FBXO33, FOXO3, HHEX, MC4R, RNF130, RP11-227G15.6, SIRT1, SLC4A10, SPATS2L, TNS3, TTC28, UBAP1, WTI, YEATS2</p>	<p>(73 2-stage): AC005027.3, AC011294.3, AC069368.3, AC074391.1, ADAMTS1, AKR1A1, ALDH8A1, ANKUB1, AP000320.7-AP000318.2, BMP2, BNC2, BRD1, BUD13, C10TNF7, CDC30, CDKN1A, CLEC16A, CSM51, CNOT1, CYBRD1, DDAH1, DIP2A, DIRC1, DYNLRB1, FBXL17, FGD6, FOXD1, G6PC2, HAUS6, HSF2, IL6, ITGA1, ITGA9, KIF15, LCAL5, LIG3, LINC00521, LINC00536, LRRRC69, MALRD1, MAP2K2, MED13L, MN1, MS12, MTRNR1B, MXRA7, NDUFAF6, ODF2L, PAPP4, PDE3A, PHC2, PSMG2, OSOX1, RAMP2, RN7SKP15, RNF219, RP11-339B21.8, RP11-432I5.2, RREB1, SAMD4A, SCN10A, SGIP1, SHOX2, SMOG2, SNX19, STAM2, TRANK1, UBE2I, UQCRI10, WDR1, YAP1, YY1, ZNF385B</p> <p>(41 1-stage): AC007381.2, ADCY5, APOB, BCAT1, C10orf76, CDK14, CHD2, CHRM2, COL6A1, CTC-340I23.2, DAZAP1, DIO3, EEPD1, ELL, FHL2, FNDC3B, FOXN3, KLHL29, LIMK1, MAFK, MAK16-TT12, PPP4R2, PROM1, PRPF40A, RIN3, RN7SL89P, RNF144B, RP11-15824.5, RP1-130H16.18, RP11-373N2.3, RP11-497E19.2, RP11-95P2.1, SLC22A3, SNORA40, TBLX1R1, TBX18, TET1, TGFBR3, TMEM239, ZNF467, ZNF516</p>
<p>DBP</p> <p>(64 2-stage): AC011518.1, AC073218.1, AKR1B10, ANO1, ATAD5, AUTS2, BEND7, BTBD3, CCKBR, CD160, CDK17, CITED2, COLEC11, CTAGE1, CTBP2, CYP27A1, DNAJB4, EPN2, FOXK1, GRM7, GTF2I, GYPC, HSPA4, IGFBP7, LINC00211, MBNL2, MIR4421, MMP14, NACA, NCOR2, OR51E1, PCDH17, PDLIM5, PGR, PIAS1, PIEZO2, PIK3R3, PKD2L1, PLEC, POLD3, POLN, PPM1A, PRSS50, RERG, RP11-1038A11.3, RP11-20D14.4, RP11-34N19.1, RP4-655J12.4, RP4-71E24.1, RPS27P25, SCN2A, SFHML1, SLC03A1, STARDB6, STK38L, TMEM44, TRIM13, TRMT10C, UBE2E2, VEGFA, WDR90, ZAP70, ZNF462, ZSCAN2</p> <p>(88 1-stage): AC053503.11, AC068196.1, AC083949.1, AC097495.2, ACVR2A, AGPAT4, AP1B1P1, ARAP2, ASXL3, ATP10A, ATP12A, ATXN7, BCKDHB, BTRC, C1GALT1, C9orf170, CACNA1C, CAMTA1, CCDC33, CCDC68, CCM2, CCNT2-AS1, CDKSRAP1, CLNS1A, CTC-360G5.8, DACHI, DCDC1, DGKB, DPYSL2, DUSP1, EPC1, EXOC6B, FAM168A, FLJ00388, GAB2, GLI2, GRB10, HDAC4, KANK3, KB-1507C5.2, KCNBB1, KIF26A, KLF2, KLF7, LAMC1, LPHN3, MCPHI1, MIR3927, MLTK, MLXIP, MRPS31, MXD3, MXI1, NEDD4L, NTM4, OGFR1L, OSGIN2, PKFB2, PLCKD2, PLEKH11, PLK2, POU2F1, PPHL1, PRDM1, PTPRD, RBMS3, RGS17, RP11-125B21.2, RP11-145M9.2, RP11-264C15.2, RP11-302M6.4, RP11-506O24.1, RP11-61G23.1, RP11-65J21.4, SIK2, SLC7A2, SPRY2, SYNPO2, TBX20, TCF4, TDRD5, UBAP2L, UBASH3B, USP34, VP54, WAC, ZNF100, ZNF423</p>	<p>(53 2-stage): AC022431.2, ACTB2, ACVR1C, ALG62294.1, AQP1, ARHGAP15, ATP2B1, BANK1, BMPR1B, C12orf75, C1orf21, CAPRIN1, CTD-2260A17.2, CTD-234988.1, DLG1, FAM129B, FARP2, IRX1, KIF2A, KLF5, LINC01006, MAST4, MECR, MEF2A, MEX3C, MORC3, NFATC2, NRG4, NUDT3, OPRM1, PDP2, PLA2G12B, PPM1E, PTK2, PTPRD, RGS6, RP11-444A22.1, RP11-455F5.3, RP11-714L20.1, RP11-805L22.1, SNX6, SORBS3, STIM2, SWAP70, TGFBE2, TMEM107, TOP3A, TRIM48, TSNARE1, Y_RNA, YES1, ZEB2, ZFAND2A (28 1-stage): AC005592.2, AC019181.3, AC021218.2, AEBP2, ANK3, AQP4-AS1, ARHGEF26, C16orf97, CLPB, CPS1, EGF17, ICOS, LA16c-306E5.3, NAA16, NCALD, NKD1, PAX8, RAPGEF5, RBPMS, RNU6-192P, RP11-453O22.1, RP1-74B13.2, SMOX, SORCS3, STXB5P, TSHZ1, ZEB2, ZFP36L1</p>	<p>(8 2-stage): BCL2, KLF14, L2HGDH, LRP4, MARK3, PDGFC, RXFP2, TERT</p> <p>(4 1-stage): ARM4C, FGFR2, SNRN70, YTHDF3</p>	<p>(5 2-stage): RP11-95P2.1, SLC22A3, HAS3, KIAA1755, SLX4IP, STEAP2, UBE2D3</p>
			<p>PP</p>

1M GWAS results

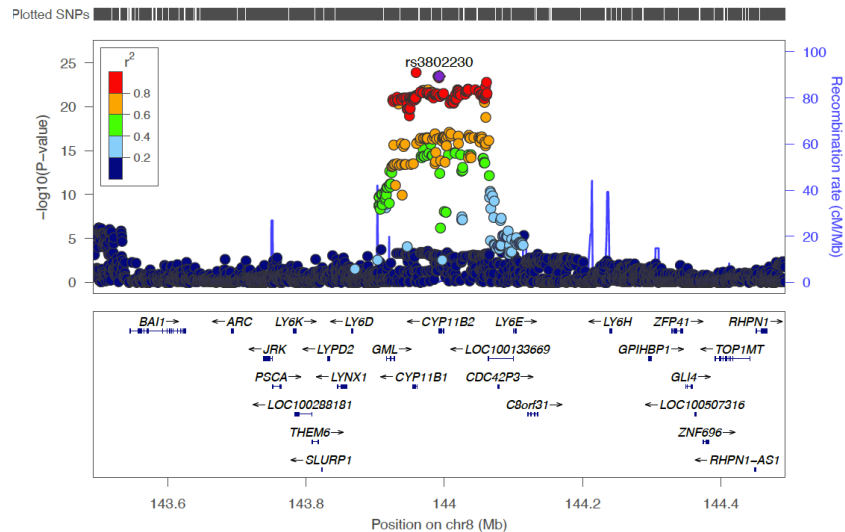
- **535 novel loci identified**
 - Mostly common variants (as per study design)
- **Support for all 274 previously published loci ($P < 0.01$)**
 - at least 95% of the exact SNPs covered in the GWAS reaching genome-wide significance
- Total 901 BP loci **tripling** the number of previously known loci

Association plots - challenges of identifying candidate genes

DBP-GWAS results for ENPEP Gene Region



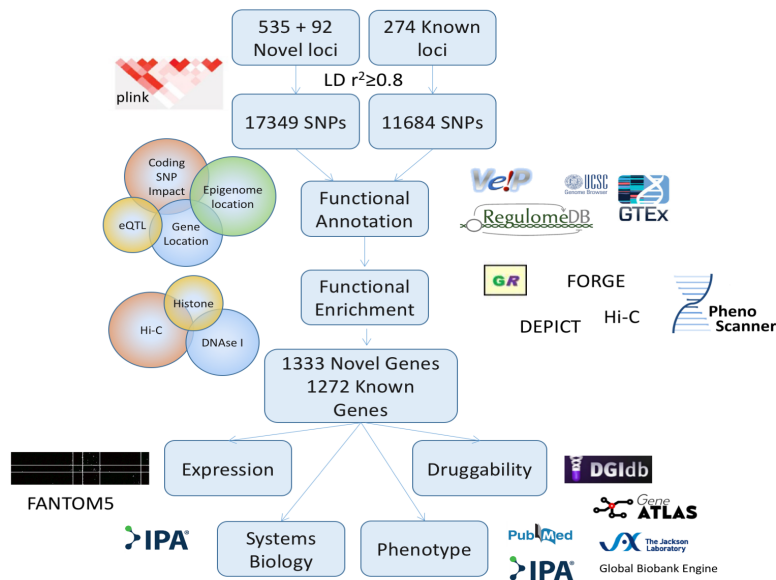
DBP-GWAS results for CYP11B2 Gene Region



Integrative bioinformatics - key first step to understanding genes and biology

Some questions you may wish to ask after performing a *GWAS*:

1. What is the candidate gene at a locus?
2. What is the causative SNP and mechanism?
3. What is the pathway affected?
4. Is there an opportunity to create new drug to the target - is the protein or pathway druggable?
5. Are there existing drugs which are safe and could be repositioned to treat hypertension?



Warren et al, 2017

Gene Expression & Enrichment

1. Lookups of SNPs & proxies ($r^2 \geq 0.8$) for association with expression of nearby genes **eQTLs** (*expression quantitative trait loci*)

44 tissues using the **Genotype-Tissue Expression (GTEx)** database

RESULT

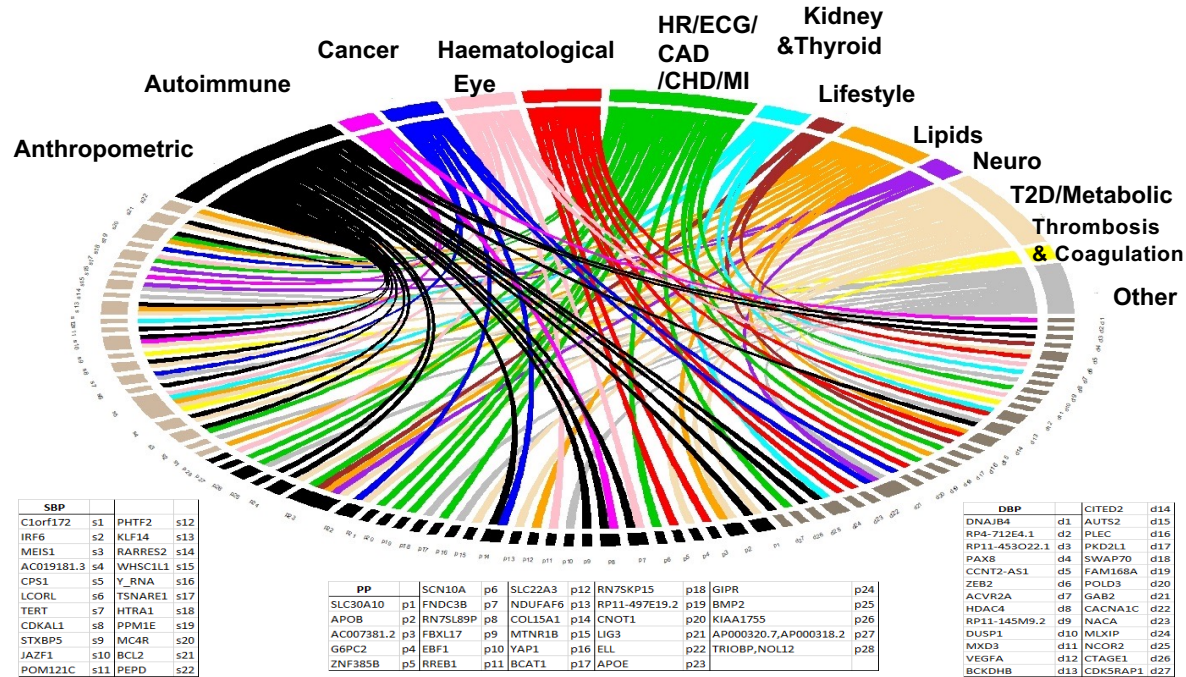
➤ Lists of candidate genes for functional exploration

2. Analysis looking for enrichment of BP SNPs with expression across 50 tissues and cells:

RESULTS

- **strongest enrichment in vasculature** (supporting prior work)
- increased enrichment in **adrenal tissue**
- new enrichment **in adipose tissues**

Check overlap of BP variants with Other Traits



Therapeutic targets?

Identified new GWAS signals at loci which are targeted by **established anti-hypertensive drugs**:

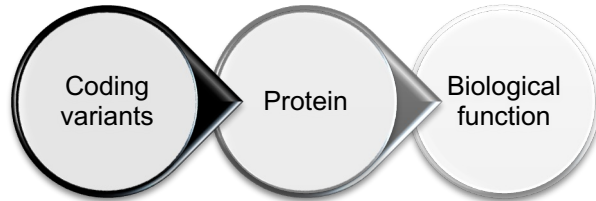
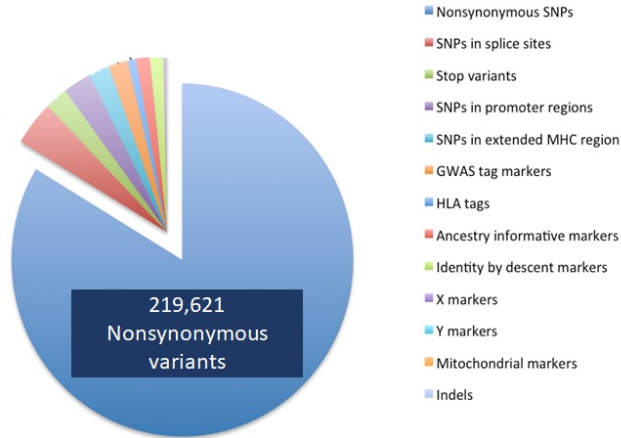
- *PKD2L1*: potassium-sparing diuretics (amiloride)
- *SLC12A2*: loop diuretics (bumetanide and furosemide)
- *CACNA1C*: calcium channel blockers (dihydropyridine)
- *CACNB4*: non-dihydropyridines
- *CA7*: thiazide-like diuretics (chlortalidone)

- Signals at loci not previously known to be associated with BP - protein target of an existing drug - potential repositioning opportunity



Role of rare variants in hypertension?

Rare variants and BP



- BP-ICE consortium
 - European and multi-ancestry analyses - >1M samples
 - 106 new BP associated regions
 - 87 rare variants at 58 loci
- Downstream analyses
 - Variant annotation & 'omics analysis for gene identification

Candidate genes from GWAS - functional follow up

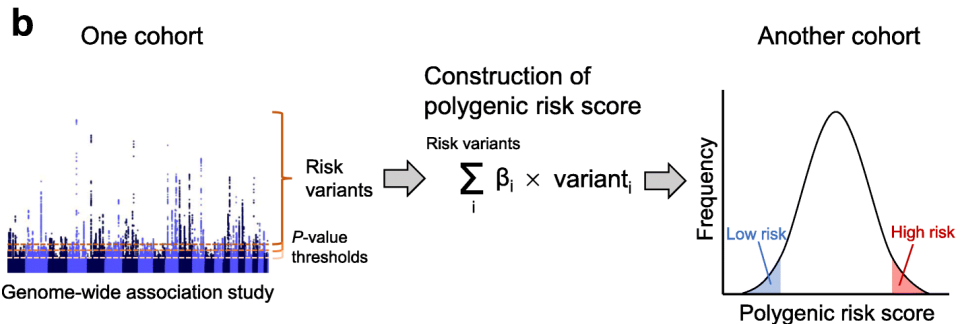
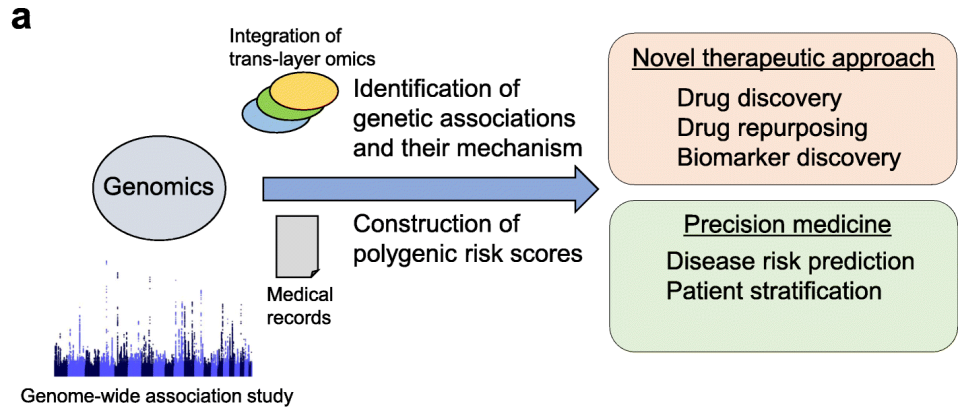
Locus/gene	GWAS discovery	SNP	Functional effect	Mechanism	PMID
MTHFR/NPPA/NPPB	2009	rs5068	Differential binding of mi-RNA425	Decreased expression of ANP, increase in BP	19219041
SLC39A8	2011	rs13107325	Missense, Ala391Trp	Affects cadmium transport and cell toxicity - lower cell viability	27466201
GUCY1A3	2011	rs7692387	intronic, affects promoter activity	Lower expression of soluble guanylate cyclase (sGC) key enzyme in NO/cGMP signalling	28487391
ARHGAP42	2011	rs604723	intronic, affects promoter activity	GTPase activating protein for RhoA in VSMC	28112683

BP genetics and clinical translational

Polygenic Risk Score (PRS)

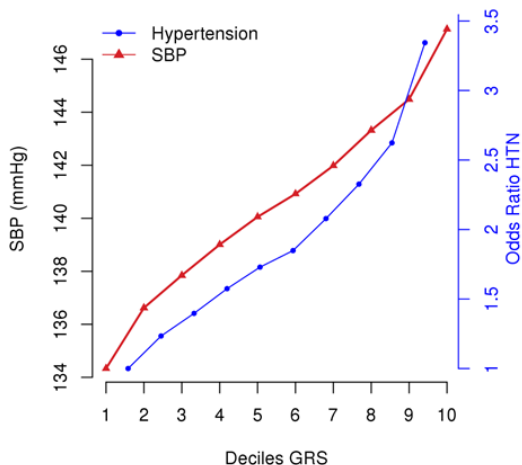
$$PRS_j = \sum_{geij} N_i \beta_i * dosa$$

where N is the number of SNPs in the score, β_i is the effect size (or beta) of variant i and dosage_{ij} is the number of copies of SNP i in the genotype of individual j.

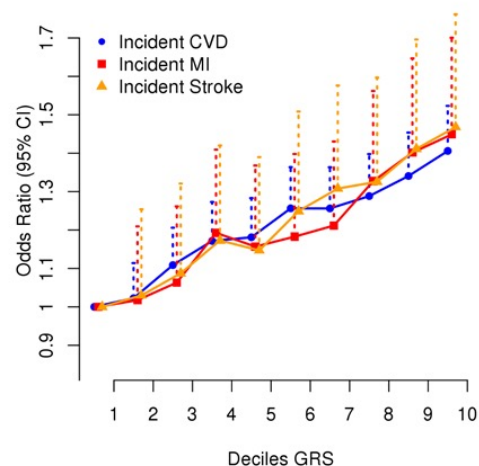


Comparing top vs bottom 10% of GRS

GRS
with
lead
variants
from 1M
GWAS

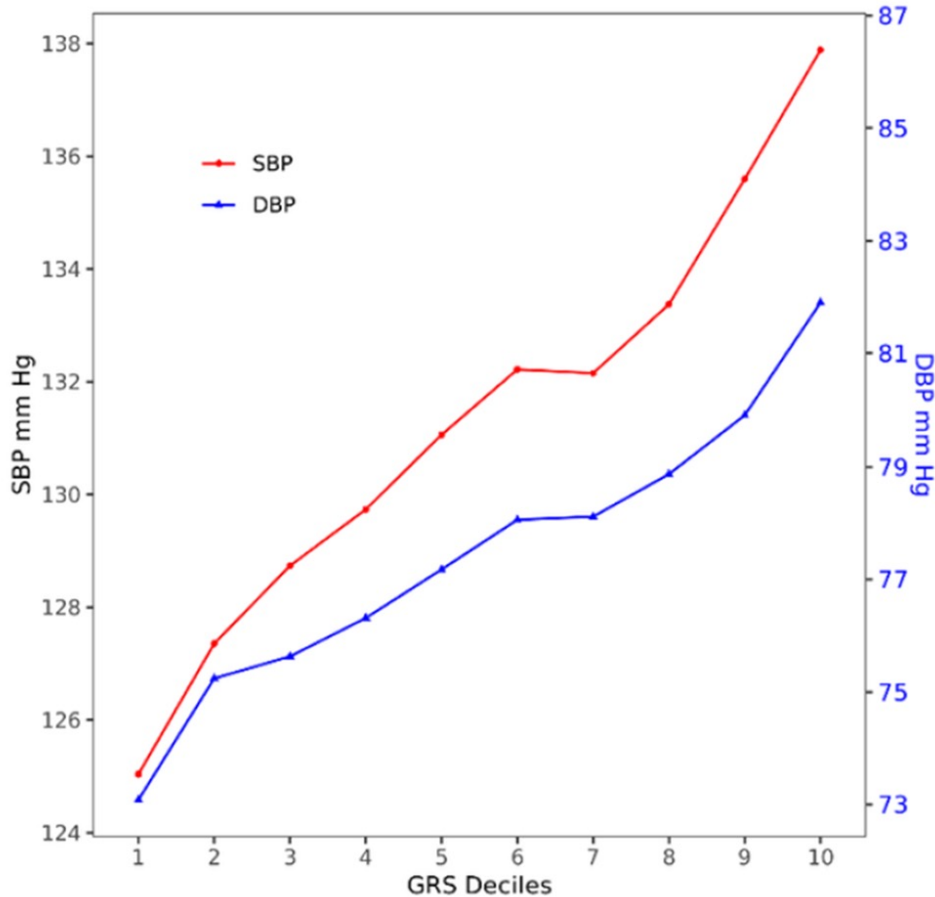


SBP increase of 12.85 mmHg
OR of 3.34 for HTN
($P < 1 \times 10^{-300}$)



**Increased risk of
all CVD outcomes**
OR = 1.52; $P = 1 \times 10^{-6}$

Polygenic Risk Score (more BP variants)



Mean difference of
top 10% vs bottom 10% of PRS

12.9 mmHg SBP
95% CI [11.5 - 14.2] ($P=9.08 \times 10^{-73}$)

8.8 mmHg DBP
95% CI [8.0 - 9.7] ($P=3.88 \times 10^{-93}$)

→ **5.4-fold increased
odds of HTN**
95% CI [4.12 - 7.10]
($P=9.71 \times 10^{-33}$)

Summary

- >2000 BP loci across all ancestries
 - Mostly common variants with small effect sizes
 - Some evidence for low and rare variants with larger effect sizes
- New insights into BP regulatory pathways (abnormal CV physiology, G protein signalling, embryonic growth retardation)
- Significant association of polygenic risk scores with HTN and CV diseases
- 1/6th of identified rare variants are located in or near potentially druggable genes (e.g. novel *NR3C2*);

GWAS accounting for gene-smoking interactions

ARTICLE

A Large-Scale Multi-ancestry Genome-wide Study Accounting for Smoking Behavior Identifies Multiple Significant Loci for Blood Pressure

Yun J. Sung,^{1,216,*} Thomas W. Winkler,^{2,216} Lisa de las Fuentes,^{3,216} Amy R. Bentley,^{4,216} Michael R. Brown,^{5,216} Aldi T. Kraja,^{6,216} Karen Schwander,^{1,216} Ioanna Ntalla,^{7,216} Xiuqing Guo,⁸ Nora Franceschini,⁹ Yingchang Lu,¹⁰ Ching-Yu Cheng,^{11,12,13} Xueling Sim,¹⁴ Dina Vojinovic,¹⁵ Jonathan Marten,¹⁶ Solomon K. Musani,¹⁷ Changwei Li,¹⁸ Mary F. Feitosa,⁶ Tuomas O. Kilpeläinen,^{19,20} Melissa A. Richard,²¹ Raymond Noordam,²² Stella Aslibekyan,²³ Hugues Aschard,^{24,25} Traci M. Bartz,²⁶ Rajkumar Deoajoo,²⁷ Yongmei Liu,²⁸ Alisa K. Manning,^{29,30} Tuomo Rankinen,³¹ Albert Vermeulen,^{32,33} Salman M. Tajuddin,³⁴ Bamidele O. Tayo,³⁵ Helen R. Warren,^{7,36} Wei Zhang,³⁷ Huihua Zhou,³⁸ Nana Matoba,³⁹ Tamar Sofer,⁴⁰ Maris Alver,⁴¹ Marzyeh Amini,⁴²

- 81 novel loci for SBP and DBP
- **10 loci showed significant interactions** ($P_{1df_INT} < 5 \times 10^{-8}$) in the **African ancestry** meta-analyses, not replicated in stage 2

- 38 novel loci for MAP and PP
- **5 loci showed significant interactions** ($P_{1df_INT} < 5 \times 10^{-8}$) in the **African ancestry** meta-analysis, not replicated in stage 2

A multi-ancestry genome-wide study incorporating gene–smoking interactions identifies multiple new loci for pulse pressure and mean arterial pressure

Yun Ju Sung ✉, Lisa de las Fuentes ✉, Thomas W Winkler ✉, Daniel I Chasman ✉, Amy R Bentley ✉, Aldi T Kraja ✉, Ioanna Ntalla ✉, Helen R Warren ✉, Xiuqing Guo ✉, Karen Schwander ... Show more

Human Molecular Genetics, ddz070, <https://doi.org/10.1093/hmg/ddz070>

Gene-smoking interactions in UK Biobank

- Genome-wide interaction analyses in UK Biobank participants to search for
 - **novel BP loci that can be identified by accounting for variant-smoking interactions**
 - **genetic associations that may differ by smoking status**
- To test for interactions we used
 - the **1df test of interaction**
 - the **2df joint test of main and interaction effects**
- BP traits
 - **Systolic blood pressure (SBP)**
 - **Diastolic blood pressure (DBP)**
 - **Pulse pressure (PP)**
- Smoking exposures
 - **Current Smoking (yes/no)**
 - **Ever Smoking (yes/no)**



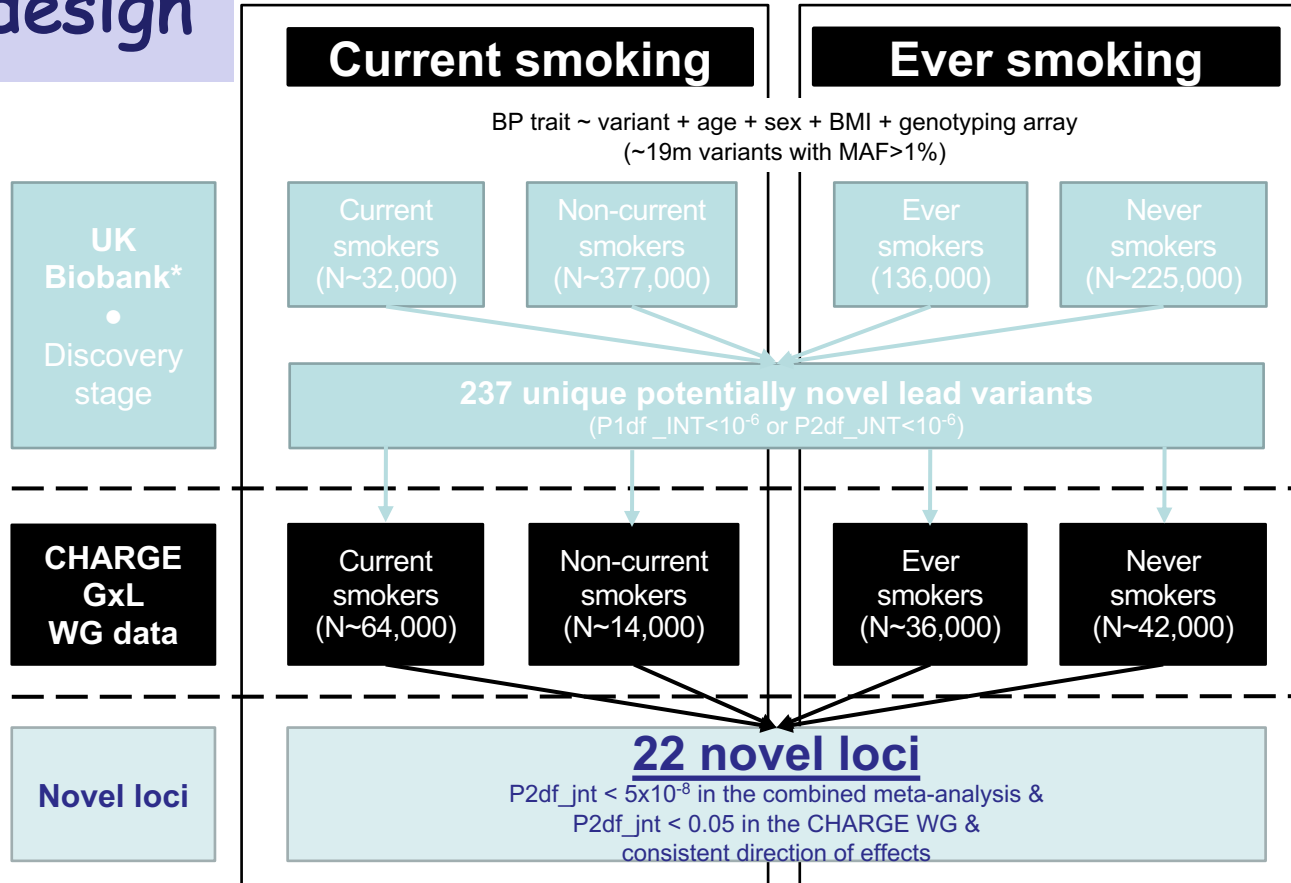
- large prospective cohort study of 0.5m participants
- outcomes and exposures measured with the same phenotypic tool during the same period
- well-powered for interaction analyses

Discovery dataset

- ~420,000 unrelated European UK Biobank participants
- 3x larger than previous discovery datasets

		Current smoking			Ever smoking	
		All participants	Current smokers	Non-current smokers	Ever smokers	Never smokers
SBP	N	420,919	32,080	377,385	136,665	225,532
	Mean	141.2	138.6	141.5	142.8	140.2
	SD	20.64	20.55	20.65	20.87	20.43
DBP	N	420,898	32,058	377,383	136,623	225,555
	Mean	84.32	83.45	84.41	84.9	83.98
	SD	11.25	11.34	11.24	11.25	11.23
PP	N	420,538	32,033	377,060	136,494	225,364
	Mean	56.84	55.2	57.06	57.87	56.27
	SD	14.13	13.94	14.15	14.48	13.9
Age (years)	Mean	56.82	55.12	57.03	57.8	56.13
	SD	7.96	7.98	7.92	7.74	8.03
BMI (kg/m²)	Mean	27.38	26.9	27.42	27.92	27.09
	SD	4.74	4.83	4.74	4.79	4.74
Sex	% males	45.96	52.12	45.06	54.09	40.35

Study design



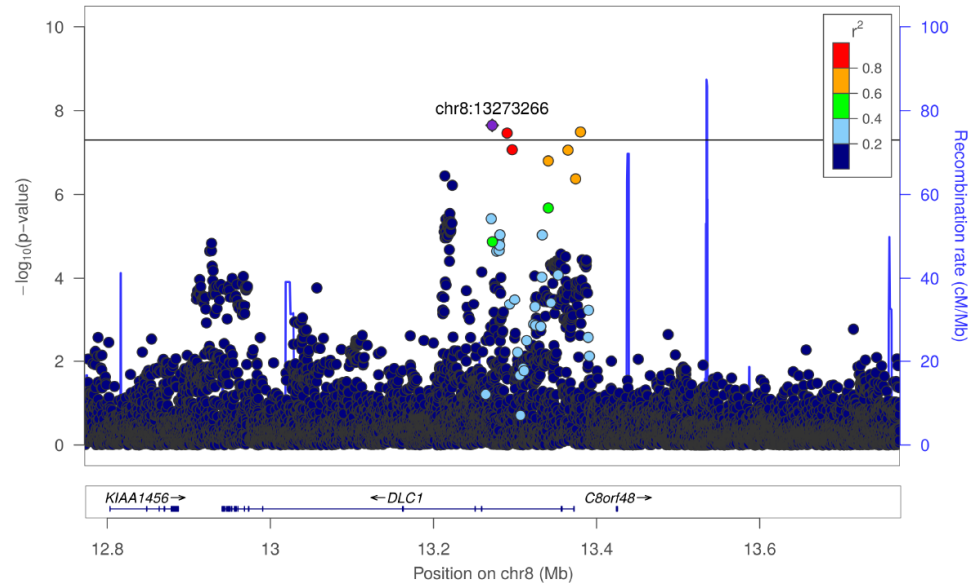
* We also performed variant main effects analyses in all participants (N~420,000)

Current smoking - 16 novel loci

Trait rsID	Nearest gene	Chr	Pos	EAO	EAF	UK Biobank				CHARGE WG				meta-analysis			
						beta	beta0	beta1	Pall	P1df_IN T	P2df_JN T	beta0	beta1	P1df_IN T	P2df_JN T	P1df_IN T	P2df_IN T
SBP rs59980837	<i>NGF</i>	1	115827266	G T	0.982	-1.31	-1.40	-0.46	2.66E-16	0.111	5.84E-17	-0.32	-2.14	0.058	0.039	0.591	2.04E-18
rs10928206	<i>GTDC1</i>	2	144965236	G C	0.482	0.22	0.22	0.22	2.42E-07	0.995	5.19E-07	0.26	0.24	0.806	0.019	0.998	1.95E-09
rs10950289#	<i>CALN1</i>	7	71429308	A G	0.838	0.29	0.32	0.02	3.67E-07	0.154	4.57E-07	0.32	0.17	0.711	0.042	0.152	5.65E-09
rs6601810	<i>KLF6</i>	10	4140086	C T	0.924	0.49	0.48	0.61	9.67E-10	0.633	3.22E-09	0.43	0.88	0.303	0.006	0.306	1.87E-12
rs11000060	<i>PRKG1</i>	10	53668890	T C	0.820	-0.29	-0.30	-0.21	1.24E-07	0.660	2.91E-07	-0.31	-0.15	0.695	0.042	0.483	2.56E-09
rs17245822	<i>SNORA9</i>	13	73131694	A C	0.627	-0.25	-0.25	-0.11	2.48E-08	0.360	7.18E-08	-0.31	-0.29	0.846	0.002	0.491	6.66E-11
rs3761287	<i>XRN2</i>	20	21281988	G A	0.380	0.22	0.22	0.30	3.59E-07	0.600	7.48E-07	0.33	0.17	0.474	0.016	0.810	4.34E-09
DBP rs1213404	<i>NREP</i>	5	111103018	G A	0.582	0.13	0.13	0.08	1.72E-07	0.535	3.16E-07	0.17	0.08	0.717	0.017	0.440	1.66E-09
rs9496614	<i>LOC105377911</i>	6	100613551	T C	0.755	0.17	0.17	0.18	6.95E-10	0.911	1.09E-09	0.26	0.09	0.510	0.002	0.773	6.72E-13
rs1178947#	<i>FZD9</i>	7	72850178	T C	0.796	0.15	0.14	0.26	1.91E-07	0.285	5.43E-07	0.21	0.14	0.550	0.041	0.419	6.1E-09
8:49445257:ID	<i>EFCAB1</i>	8	49445257	I D	0.875	0.22	0.23	0.07	9.60E-10	0.202	1.39E-09	0.20	0.35	0.834	0.029	0.483	7.9E-12
rs78550103	<i>RHCG</i>	15	90027296	G A	0.857	0.18	0.18	0.17	1.10E-07	0.947	4.37E-07	0.24	0.13	0.529	0.029	0.793	3.13E-09
rs11083857*	<i>TMEM160</i>	19	47549454	C T	0.766	0.14	0.15	0.07	3.46E-07	0.413	7.31E-07	0.84	0.31	0.357	0.004	0.412	2.74E-08
PP rs144370193	<i>DLC1</i>	8	13273266	I D	0.798	-0.21	-0.22	-0.11	1.17E-08	0.385	2.24E-08	-0.31	-0.28	0.816	0.005	0.506	2.42E-11
rs7099368	<i>CCDC6</i>	10	61566323	T C	0.590	0.15	0.16	0.02	2.39E-07	0.180	4.27E-07	0.06	0.39	0.050	0.013	0.992	3.96E-08
rs78799967	<i>UBE4A</i>	11	118266523	C T	0.974	0.66	0.66	0.92	9.05E-13	0.431	5.36E-13	0.51	1.14	0.444	0.045	0.307	2.95E-15

PP - Current Smoking - *DLC1*

- intronic common variant in *DLC1*
- the effect allele was associated with lower PP with no major differences between the two smoking strata
- *DLC1* encodes a small GTPase involved in RhoA signaling - a known BP homeostasis pathway
- *Dlc1* mouse models show abnormal heart development (MGI:3529997)



Ever smoking - 6 novel loci

Trait rsID	Nearest gene	Chr	Pos	EA	OA	EAF	UK Biobank						CHARGE WG				meta-analysis	
							beta	beta0	beta1	Pall	P1df_INT	P2df_JN T	beta0	beta1	P1df_INT	P2df_JN T	P1df_INT	P2df_JN T
SBP rs1229984*	ADH1B	4	100239319	T	C	0.026	-0.93	-0.58	-1.56	1.35E-11	1.0E-04	6.75E-14	-0.64	-0.88	0.459	0.005	0.001	1.10E-16
rs767717	ISL1	5	50880788	T	C	0.373	0.25	0.18	0.35	9.35E-08	0.055	1.27E-07	0.12	0.32	0.332	0.028	0.037	9.50E-10
rs17085414#		13	27880846	T	C	0.885	0.36	0.34	0.41	2.25E-07	0.633	7.35E-07	0.21	0.38	0.442	0.044	0.532	1.77E-08
DBP rs1950500	NFATC4	14	24830850	T	C	0.292	0.15	0.17	0.09	1.55E-07	0.094	5.11E-07	0.22	0.11	0.386	0.018	0.075	2.68E-09
PP rs1523475	BCL6	3	187444210	T	C	0.193	-0.19	-0.26	-0.07	1.53E-06	0.010	1.62E-07	-0.29	-0.13	0.277	0.014	0.008	8.81E-10
rs4143530	ZFHX4	8	77593246	A	C	0.863	0.31	0.37	0.22	2.12E-12	0.076	1.84E-12	0.25	0.23	0.990	0.027	0.097	1.21E-14

*Missense variant; #Proxy of the lead variant in the locus.

beta: effect estimate from the main effects analyses; beta0: effect estimate from the non-exposed group; beta1: effect estimate from the exposed group.

Pall: P-value from main effects analyses in all participants of UK Biobank; P1df_INT: p-value from the 1df interaction test; P2df_JNT: p-value from the 2df joint test

Learning objectives

- Understanding of monogenic and complex genetic disorders
- Methods for analyzing the genetic basis of genetic disorders
- Hypertension - exemplar
 - Rare monogenic forms
 - BP gene discovery research programme
 - GWAS, rare variants, bioinformatics, functional studies
 - Examples of the utility of genetics for clinical translation
 - Gene x environment analyses

Thank you

- Any questions or if you want to know more, please do email me
 - p.b.munroe@qmul.ac.uk