Population Genetics & Drug Discovery

examples from Finland

Mark J. Daly
Chief, Analytic and Translational Genetics Unit
Massachusetts General Hospital
Co-director, Medical and Population Genetics
Broad Institute
Visiting Professor, Institute for Molecular Medicine, Finland (FIMM)

FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

EARLY SETTLEMENT

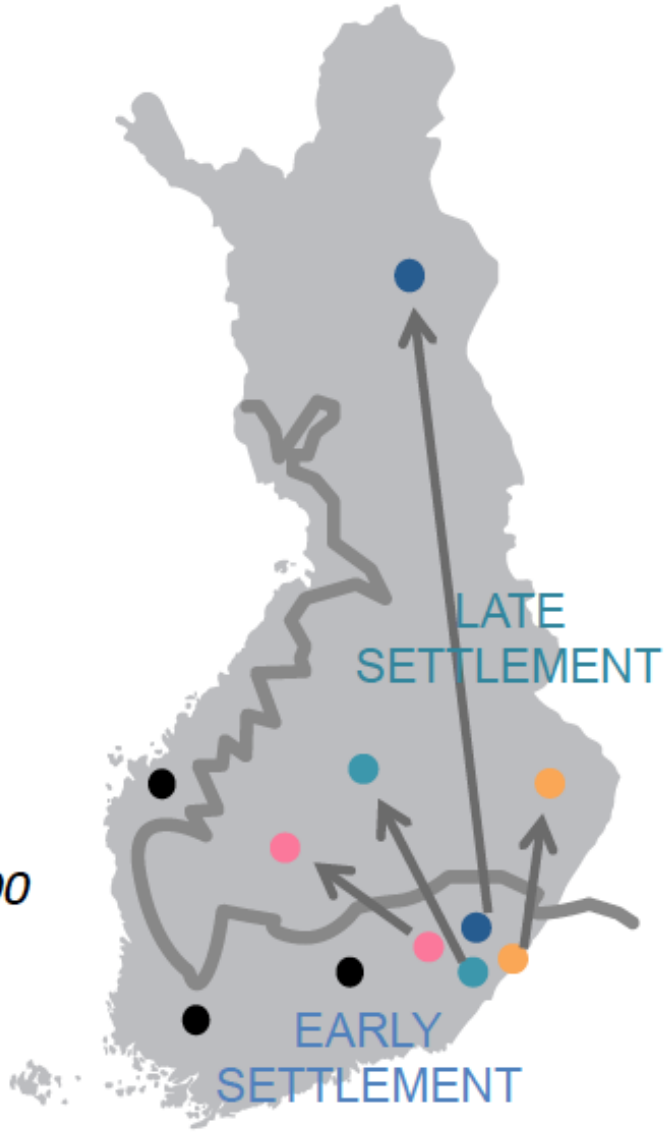
- 2000-10 000 years ago
- South and Coast

LATE SETTLEMENT

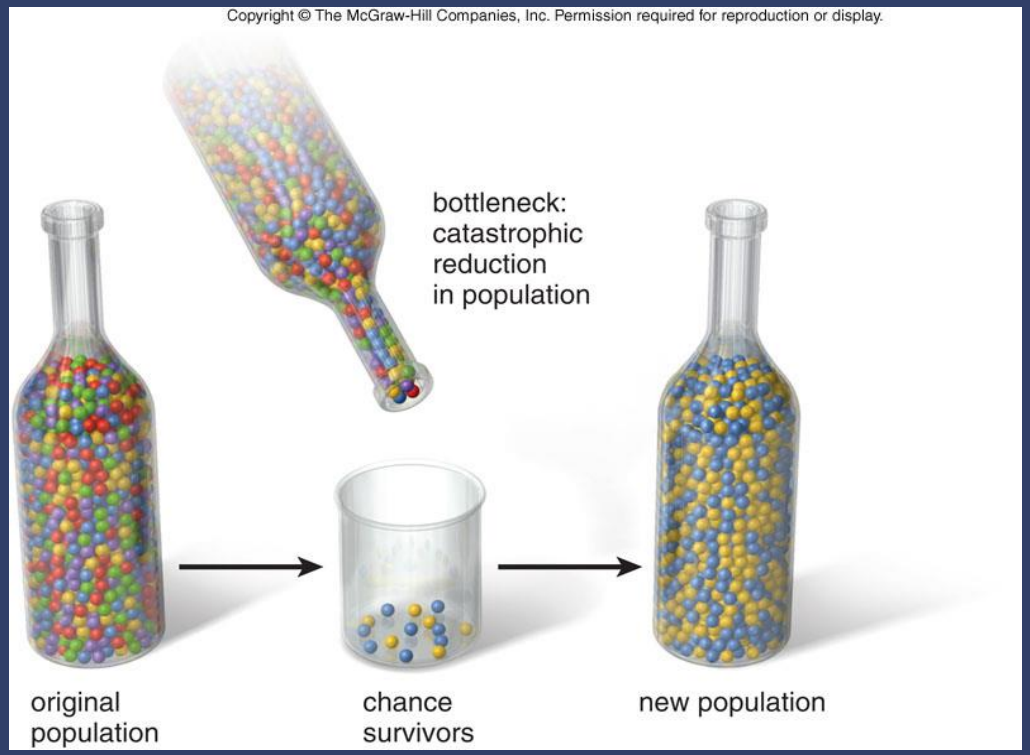
- 16th century
- multiple bottle necks

EXPANSION

- 18th century – *population 250 000*
- Today – ***population 5.4 million***



Effect of population bottleneck



Finns have dramatically fewer rare DNA variants than other Europeans

Some rare DNA variants jump to unusually high frequencies:

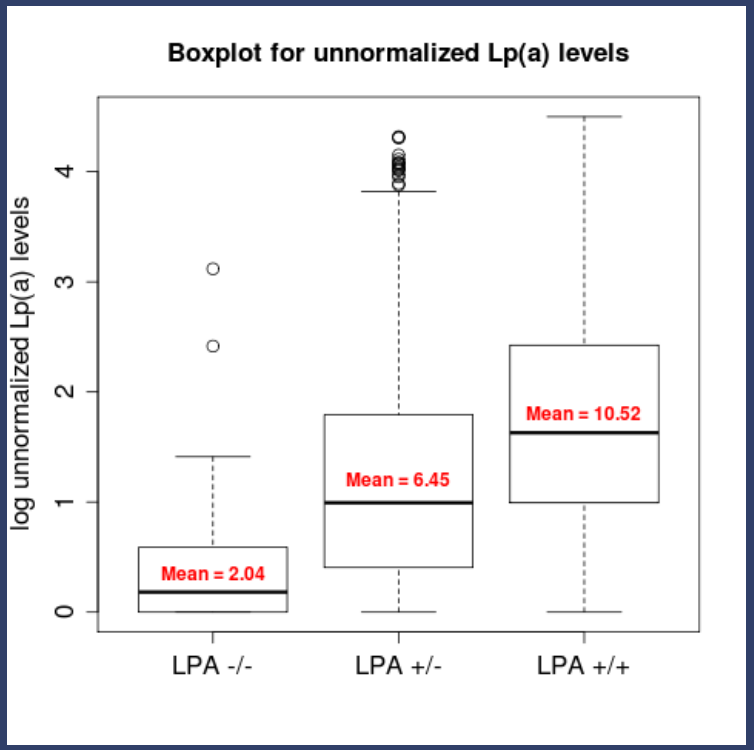
- Some are known causes of Finnish enriched diseases
- Thousands more alter or truncate proteins – some may improve health – unique opportunity for discovery!

FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

Loss-of-function variant in LPA protects from Coronary Heart Disease



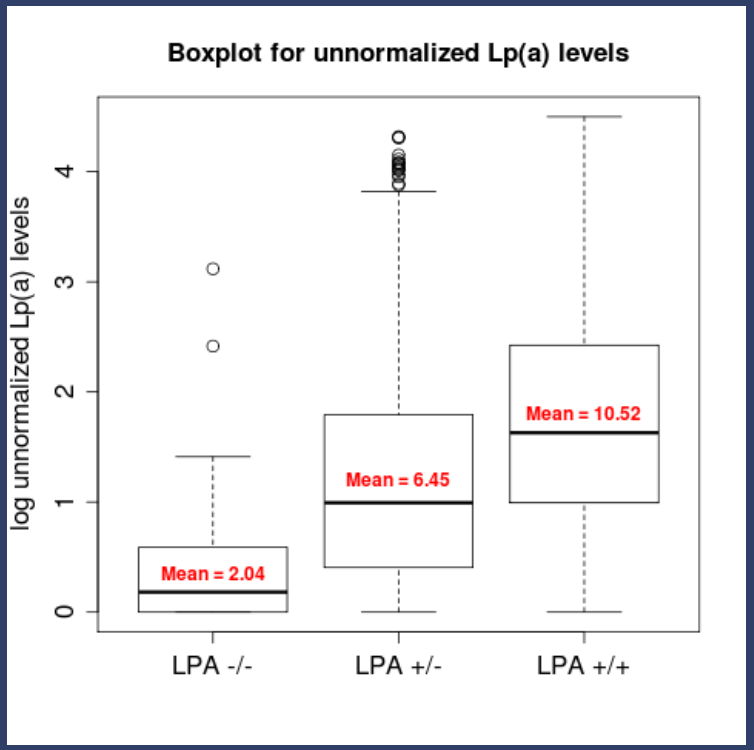
Finnish enriched DNA variant lowers LPA levels

FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

Loss-of-function variant in LPA protects from Coronary Heart Disease



Association for LPA splice variants with cardiovascular diseases

Study	Ncases n/N	Ncontrols n/N	Odds Ratio (95% CrI)	Odds Ratio (95% CrI)
FINRISK (CHD)	1076/25020	23944/25020		0.79 (0.72 to 0.86)
Estonian ExomeChip (IHD+HF)	768/4600	3832/4600		0.69 (0.31 to 1.50)
Estonian Imputed (IHD+HF)	853/7953	7100/7953		0.83 (0.51 to 1.36)
MIGEN ExA (MI)	8890/18176	9286/18176		0.88 (0.78 to 0.99)
Total	11587/55749	44162/55749		0.84 (0.80 to 0.88)

0.5 1 2

Each copy of LPA lowering variant lowers CHD risk by 15-20% !!!

Finnish enriched DNA variant lowers LPA levels

FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

Such examples provide precise clues to interventions that could be safe and effective

Key elements for drug discovery efforts

- Extensive health and registry data
  - Fully available to research
  - Collected consistently over a large population
  - Longitudinal data over decades
- Population participation and recall studies
- Integration with disease consortia

FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

Registry data available for research purposes

FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

NATIONAL REGISTERS

NAME OF REGISTER	DATA INCLUDED IN REGISTER	KEEPER	ESTABLISHED IN COMPUTER FORMAT
Hospital Discharge Register (HILMO)	Homes and institutions for the mentally disabled, including information on treatment	THL	1967
Finnish Registry for Kidney Diseases	Diseases, type of treatment, and laboratory tests	ETK	1964
Cancer Register	Cancer patient information from hospitals, pathology, laboratory measurements etc.	THL	1953
Finnish Register of Visual Impairment	Patient's visual ability	THL	1983
National Infectious Diseases Register	Detailed information on cases in infectious diseases	THL	1989
Register of Congenital Malformations	Infants and foetuses	THL	1963
Drug Reimbursement Registers	Disease that is being treated and medication used	KELA	1967
Medical Birth Register	Information on all births in Finland, from gestation week 22+0 or birthweight 500g	THL	1987
Cause-of-Death Register	Intermediate case of death and contributing causes of death	STAT	1969
Register on Occupational Disease	Diagnosis of occupational disease	FIOH	1964
Drug Surveillance Register		FIMEA	1982
National Sickness Insurance	Social benefit information	KELA	1967
Register on Pensions	Work pension information, age of individual, type of pension	ETK	1962
Finnish Employment Register	Work in private sector, work as entrepreneur and work without pay	ETK	
Central Population Register	Relations (stillbirths are not registered)	VRK	1973
Register on Social Assistance		THL	1985
Child Welfare Register	Individual-level information on children taken into custody	THL	1991

THL = National Institute of Health and Welfare
ETK = The Finnish Kidney and Liver Association
KELA = Social Insurance Institution

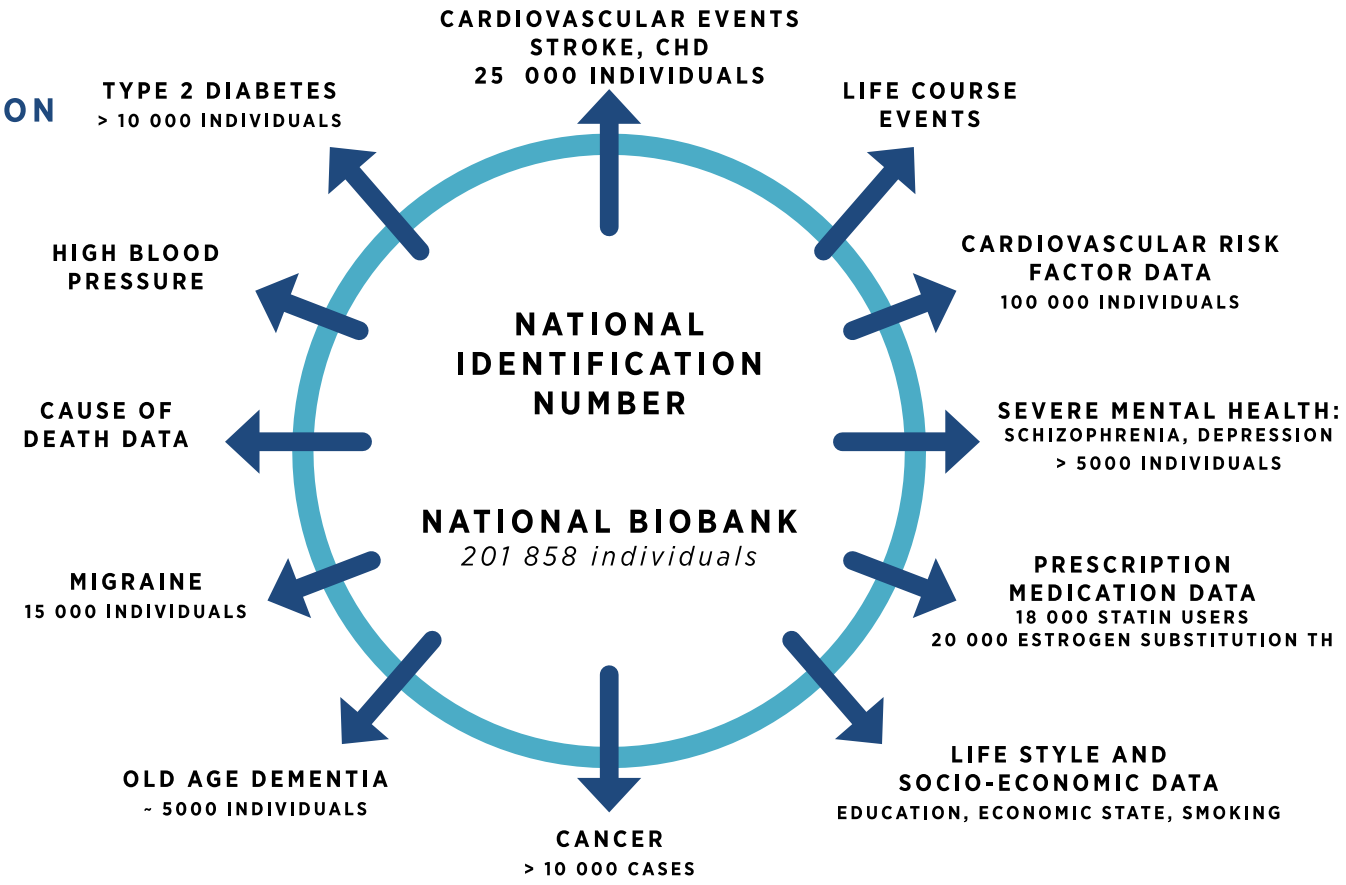
STAT = Statistics Finland
FIOH = Finnish Institute of Occupational Health
FIMEA = National Agency for Medicine

ETK = Finnish Centre for Pensions
VRK = Central Population Register

THL Biobank - example of one Finnish Biobank

130 000 INDIVIDUALS FROM POPULATION COHORTS

70 000 INDIVIDUALS FROM DISEASE COLLECTIONS



To date:

73,000 GWAS

26,000 WGS/WES

Example of health histories from **two persons** from the national biobanks **with a 40 year follow-up**

SNP CTA**A**GTA

GENOME SEQUENCE

HAPLOTYPE A
Neoplasm of unspecified nature
Benign neoplasm of breast
Spontaneous vertex delivery
Cerebral aneurysm, nonruptured
Spontaneous vertex delivery
Maternal care for poor fetal growth
Sterilisation
Cerebral aneurysm, nonruptured
Other dystonia
Severe depressive episode without psychosis
Recurrent depressive disorder without psychosis
Acute cholecystitis

DISEASE HISTORY

20

30

40

50

AGE

GENOME SEQUENCE

SNP CTA**C**GTA

HAPLOTYPE B
Diarhea of infectious origin
Other bursitis
Internal haemorrhoids with complications
Essential (primary) hypertension
Polymyalgia rheumatica
Guillain-Barré syndrome
Sleep apnea
Other spondylosis with radiculopathy
Nonallergic asthma
Pilonidal cyst without abscess
Pain localised to upper abdomen
Acute pancreatitis
Calculus of gallbladder without cholecystitis
Tear of meniscus
Inguinal hernia, without obstruction or gangrene

Photos are used for illustratrative purposes only. Photos by IStock photo. Posed by model.

New study designs facilitated

- Search for adverse outcomes before targeting protective mechanism (excess diagnoses, prescriptions, etc.)
- Use of prescription registry to enhance genetic studies of common outcomes

FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

Protective LPA LoF variants show no elevated disease risk

LoF variant	Frequency in Finns	Frequency in non Finns
LPA1(4974)	2.8%	0.47%
LPA2(4289)	4.8%	3.6%

227 Finns homozygous/compound het

Linking to National Health Records

No sign of increased morbidity

FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

Drug registry genetics proof of concept: Statin use GWAS

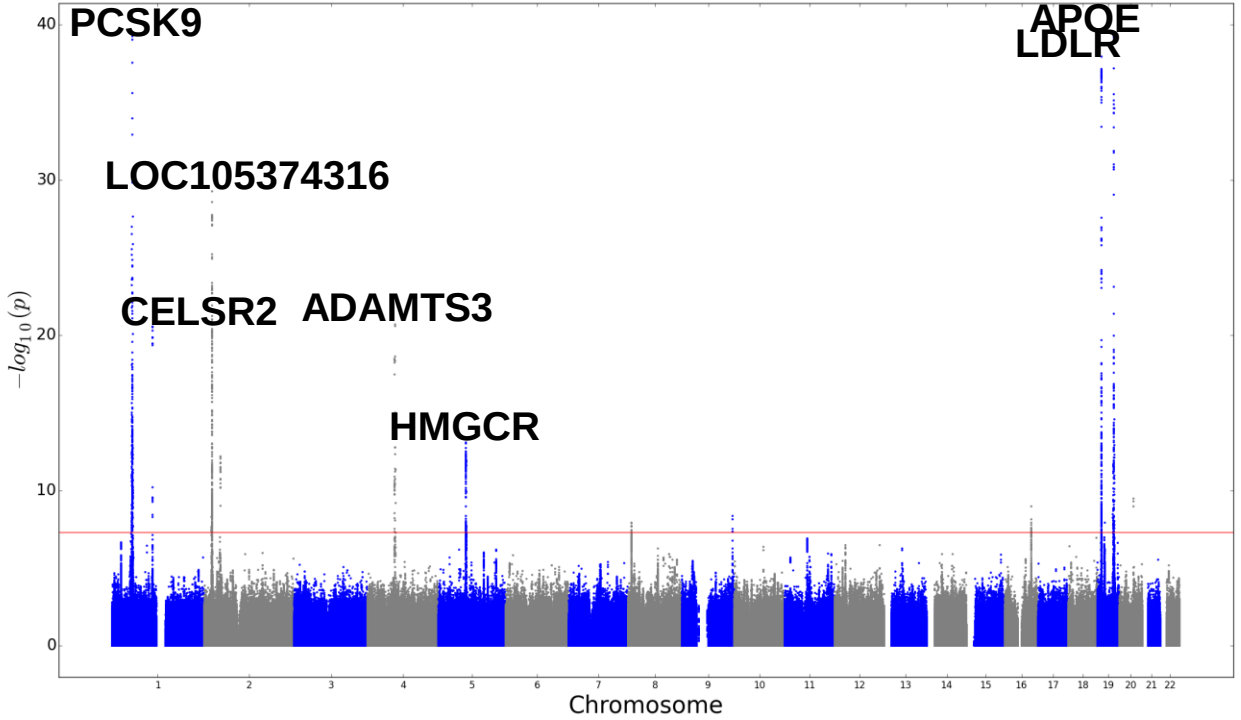
- Finrisk genotyped cohort N=25000
- Statin use data obtained from KELA
  - Months of use, starting age
- Roughly 7000 statin users versus 18000 who have never been prescribed statins

FIMM

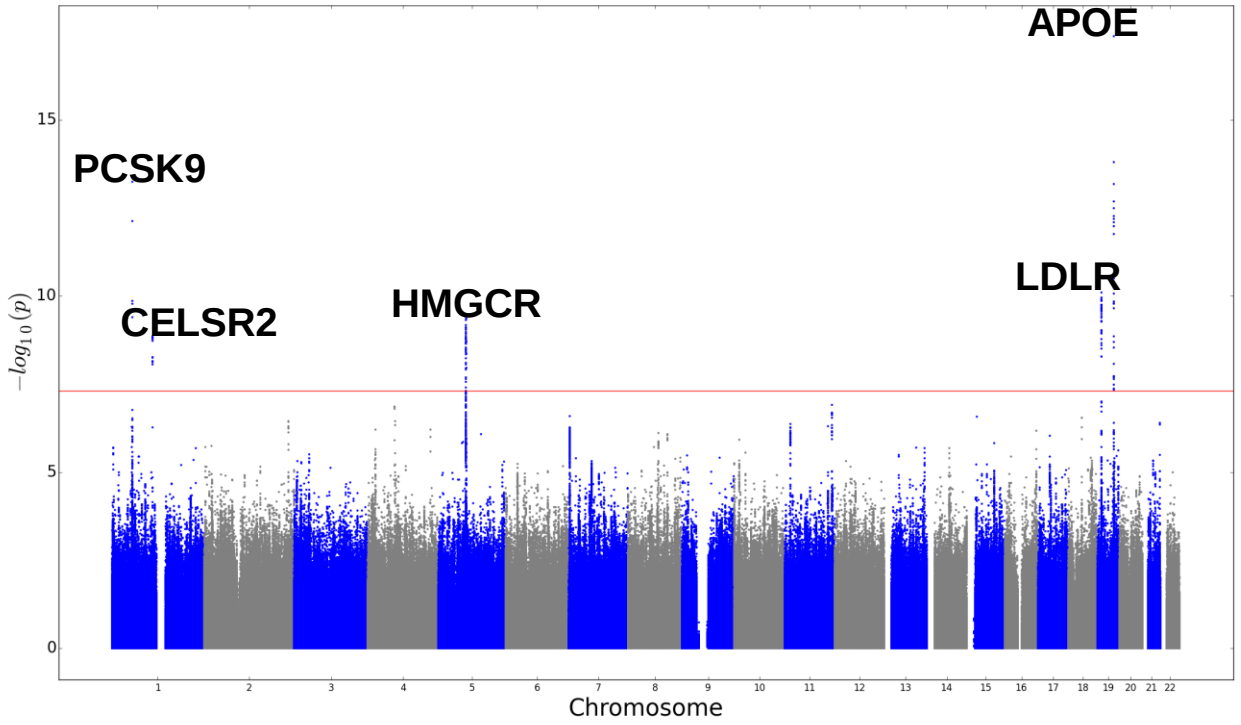
BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

LDL as quantitative trait; N=18000 non-Statin users



7000 Statin users versus 18000 non-Statin users



Population participation & genotype-based recall

FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

THE PEOPLE

HIGH LEVEL EDUCATION

ACCESS TO THEIR OWN DATA (PHR)

HIGH WILLINGNESS TO PARTICIPATE

TRUST IN AUTHORITIES

TECH SAVVY

"You can count on Finns to participate in studies where they don't stand to gain anything themselves."

The Biobank Act of 2013
- Recalling made easy

FAST & EASY RECALLING

BIOBANK ACT
+
NATIONAL IDENTIFICATION NUMBER

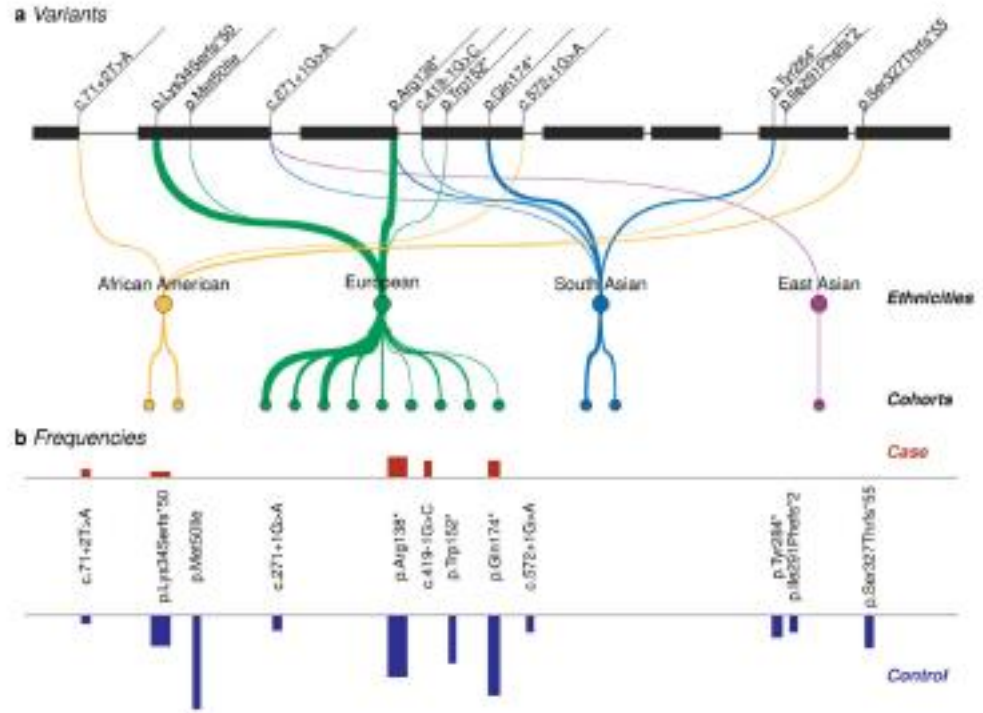
FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

Nat Genet. 2014 April ; 46(4): 357–363. doi:10.1038/ng.2915.

Loss-of-function mutations in *SLC30A8* protect against type 2 diabetes



Truncating mutations in SLC30A8 seen 3x more often in healthy controls than diabetics

STRONG protection against diabetes !!!

FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL


50% of all known Finnish R138X carriers are around Jakobstad
And 12, i.e. 50% of them, seem to have a connection to an island called Larsmo
Kokkola
Karleby
Kälviä
Kelviä
Prästasholmen
Larsmo
Luoto
Kronoby
Kruunupyy
Nedervetil
Alaveteli
Jakobstad
Pietarsaari
E8
E8
E8

Recalling SLC30A8 carriers

- More than 75% of individuals recalled have returned for extensive phenotyping involving a test meal and numerous blood draws over three hours
- 73 individuals already successfully studied under this protocol
- Almost all are healthy controls with no self-interest in the research

FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

Integration with international disease consortium studies

FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

No man is an island…

- Fun fact: Entire PMI will have fewer cases of IBD, early onset MI than have already been exome sequenced
- Most primary disease discoveries will continue to emerge collaboratively
- Biobanks allow deeper exploration of these findings, not unique access to them

FIMM BROAD INSTITUTE MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

RNF186 harbors UC protective truncating variant

Table 1 Association of p.R179X in *RNF186* with ulcerative colitis

Study	Data type	UC		Controls		Control MAF	*P*	OR
		179X	R179	179X	R179			
GWASseq	Sequence (targeted)	0	1774	6	1828	0.33%		
Finland	Sequence (exome)	0	1016	23	16223	0.14%		
Screen	-	0	2790	29	18051		0.022	0
US+Canada	Exome Chip	4	6354	21	12883	0.16%		
UK	Sequencing	2	3854	10	7294	0.14%		
Sweden	Exome Chip	2	1518	45	10813	0.41%		
Belgium	Genotyping	0	1696	0	1764	0.00%		
Germany	Genotyping	1	2035	7	4399	0.16%		
Dutch	Genotyping	1	1133	8	4164	0.19%		
Italy	Genotyping	0	0	2	1914	0.10%		
Iceland	Sequencing + Imputation	7	2899	4130	525358	0.78%		
		N cases 1453, N controls 264,744; MAF = 0.78%						
Replication							8.69×10^{-6}	0.33 (0.20-0.55)
Combined (screen +replication)							6.89×10^{-7}	0.30 (0.19-0.50)

UC, ulcerative colitis; OR, odds ratio; *P*, p-value. Screen + replication *P* value is computed using Mantel-Haenszel chi-squared test with continuity correction.

FIMM

BROAD INSTITUTE

MGH 1811 MASSACHUSETTS GENERAL HOSPITAL

Collaborating Labs:

Aarno Palotie

Samuli Ripatti

Ben Neale

Dan MacArthur

International IBD Genetics Consortium