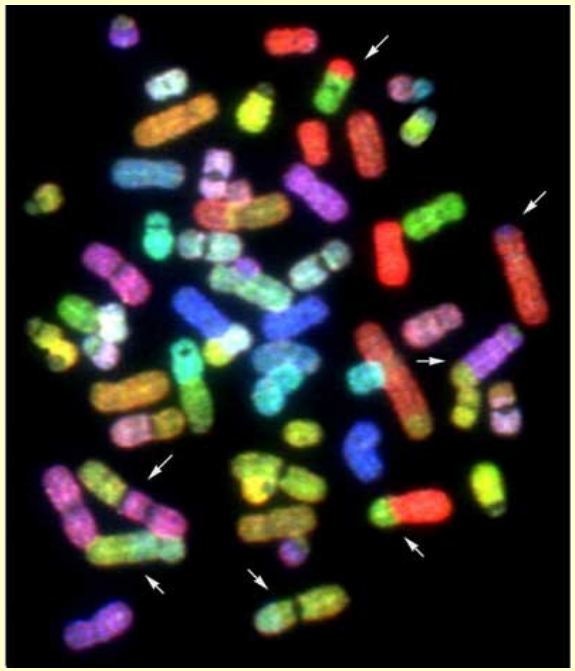
Computational Molecular Biology
Biochem 218 – BioMedical Informatics 231

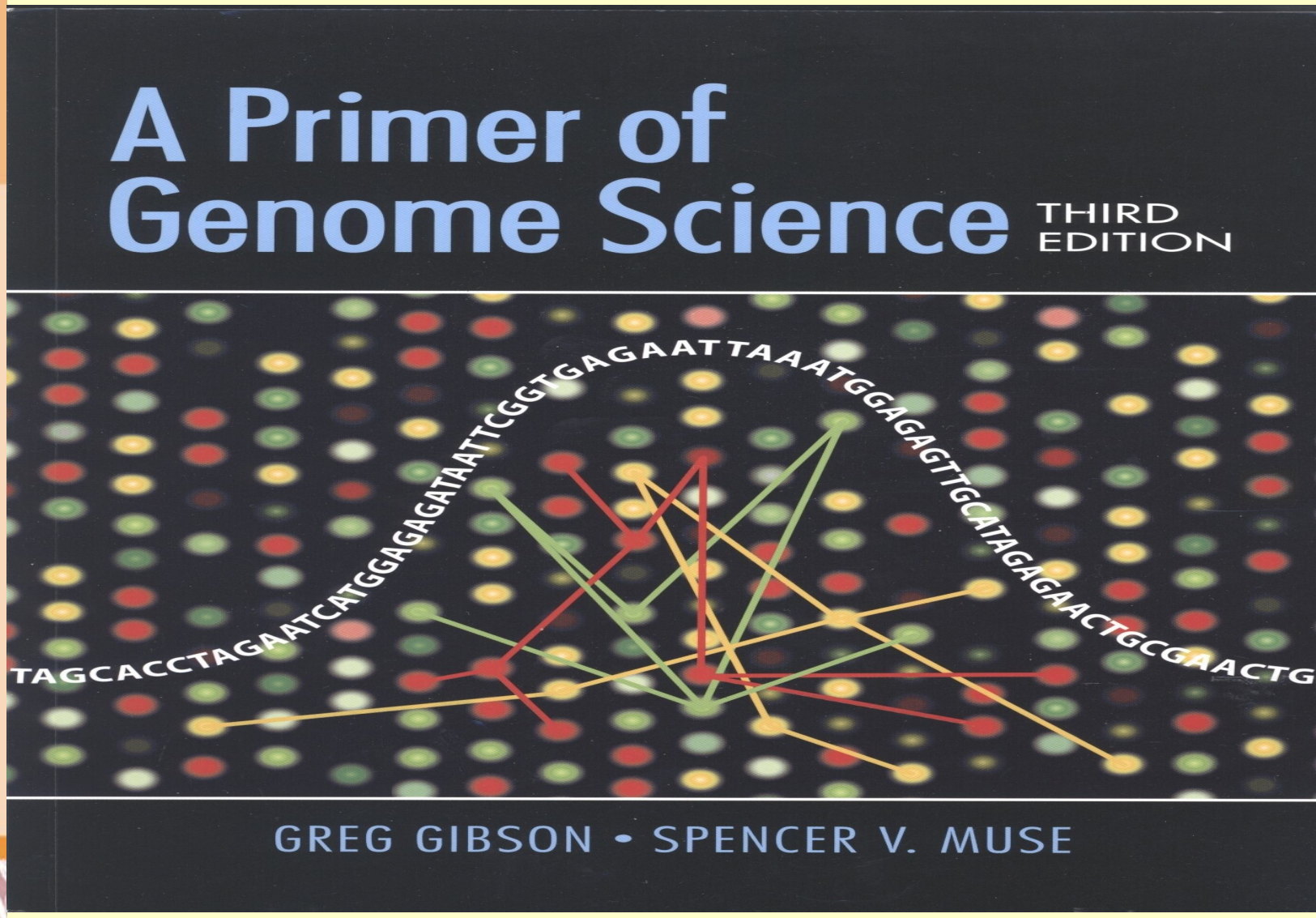
http://biochem218.stanford.edu/

Genome-Wide Association Studies:
Linking Genes to Disease

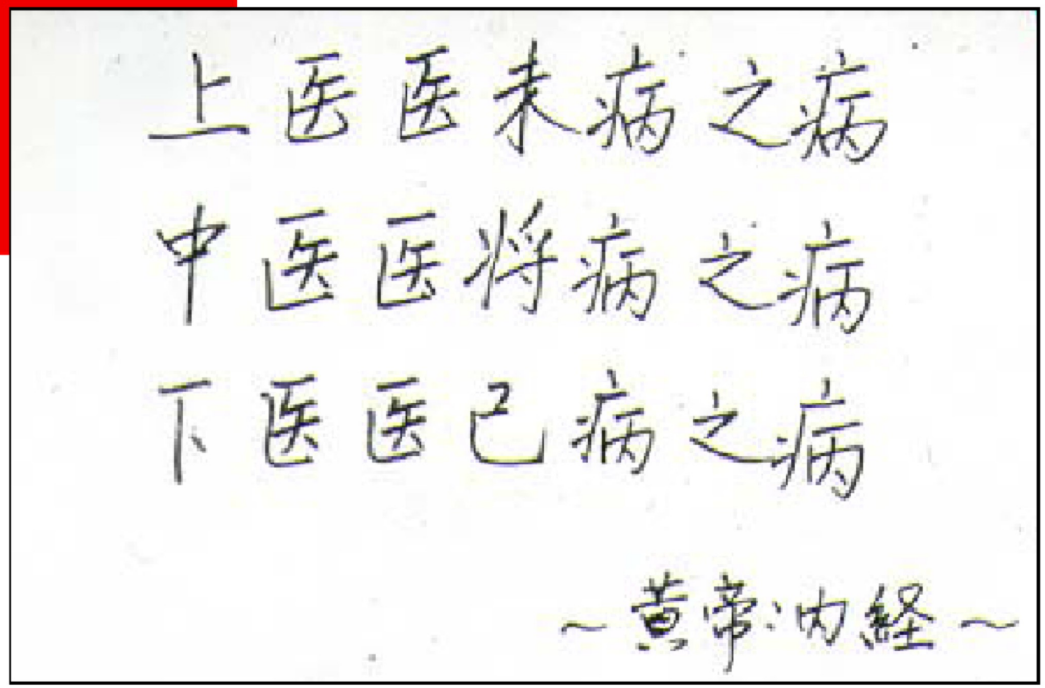


Doug Brutlag
Professor Emeritus
Biochemistry & Medicine (by courtesy)

A Primer of Genome ScienceGibson and M



Preventive Medicine



"Superior Doctors Prevent the Disease.
Mediocre Doctors Treat the Disease Before Evident.
Inferior Doctors Treat the Full Blown Disease."

-Huang Dee: Nai - Ching (2600 B.C. 1st Chinese Medical Text

Preventive Medicine

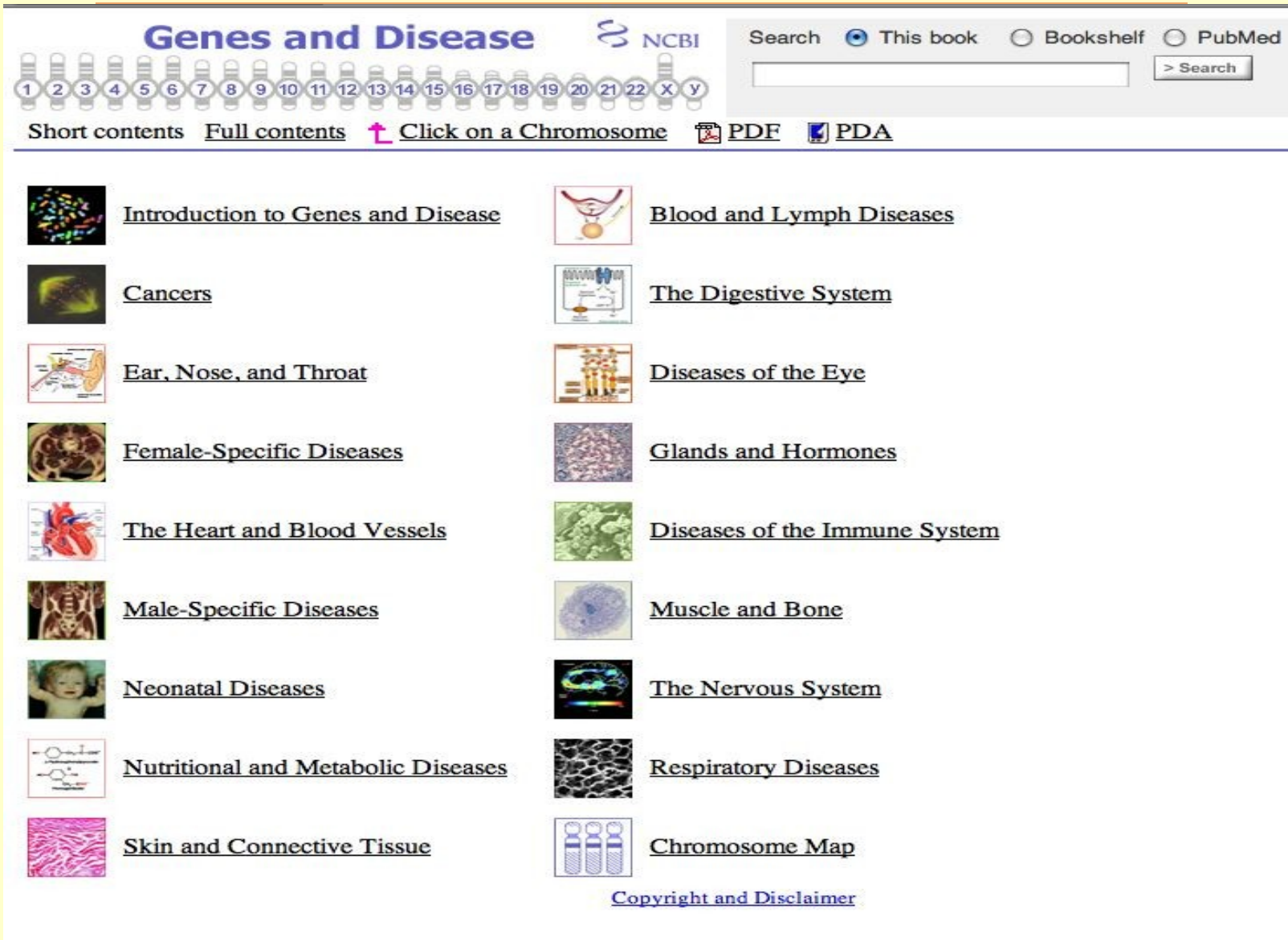
- Prevent disease from occurring
- Identify the cause of the disease
- Treat the cause of the disease rather than the symptoms
  - Example peptic ulcers
  - Pyrogens
- Genomics identifies the cause of disease
- "All medicine may become pediatrics"
  Paul Wise, Professor of Pediatrics, Stanford Medical School, 2008
- Effects of environment, accidents, aging, penetrance …
- Health care costs can be greatly reduced if
  - invests in preventive medicine
  - one targets the cause of disease rather than symptoms

Penetrance and Environmental Factors

- Highly penetrant Mendelian single gene diseases
  - Huntington's Disease caused by excess CAG repeats in huntingtin's protein gene
  - Autosomal dominant, 100% penetrant, invariably lethal
- Reduced penetrance, some genes lead to a predisposition to a disease
  - BRCA1 & BRCA2 genes can lead to a familial breast or ovarian cancer
  - Disease alleles lead to 80% overall lifetime chance of a cancer, but 20% of patients with the rare defective genes show no cancers
- Complex diseases requiring alleles in multiple genes
  - Many cancers (solid tumors) require somatic mutations that induce cell proliferation, mutations that inhibit apoptosis, mutations that induce angiogenesis, and mutations that cause metastasis
  - Cancers are also influenced by environment (smoking, carcinogens, exposure to UV)
  - Atherosclerosis (obesity, genetic and nutritional cholesterol)
- Some complex diseases have multiple causes
  - Genetic vs. spontaneous vs. environment vs. behavior
- Some complex diseases can be caused by multiple pathways
  - Type 2 Diabetes can be caused by reduced beta-cells in pancreas, reduced production of insulin, reduced sensitivity to insulin (insulin resistance) as well as environmental conditions (obesity, sedentary lifestyle, smoking etc.).

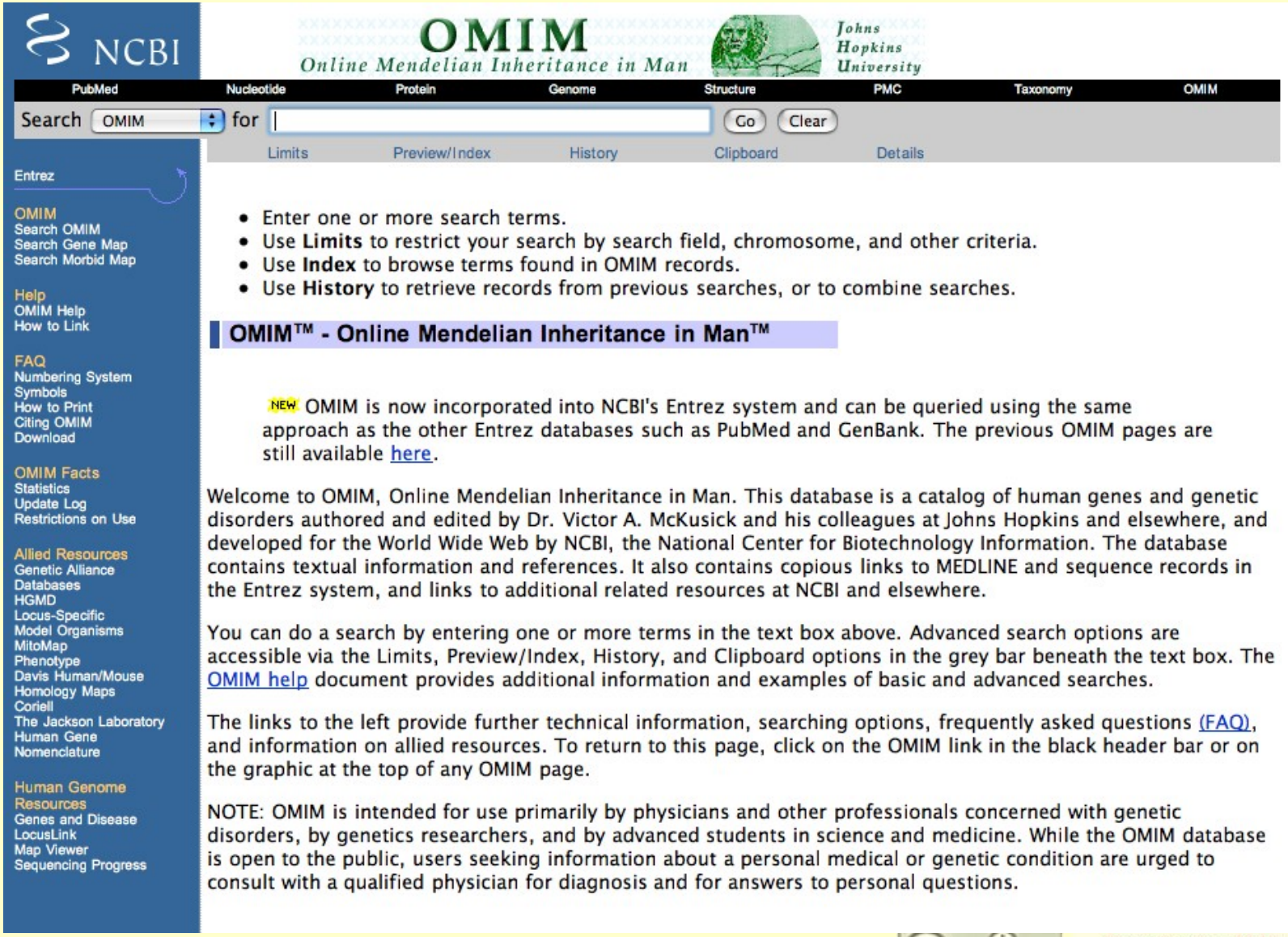
Genes & Disease

http://www.ncbi.nlm.nih.gov/books/bookres.fcgi/gnd/tocstatic.html



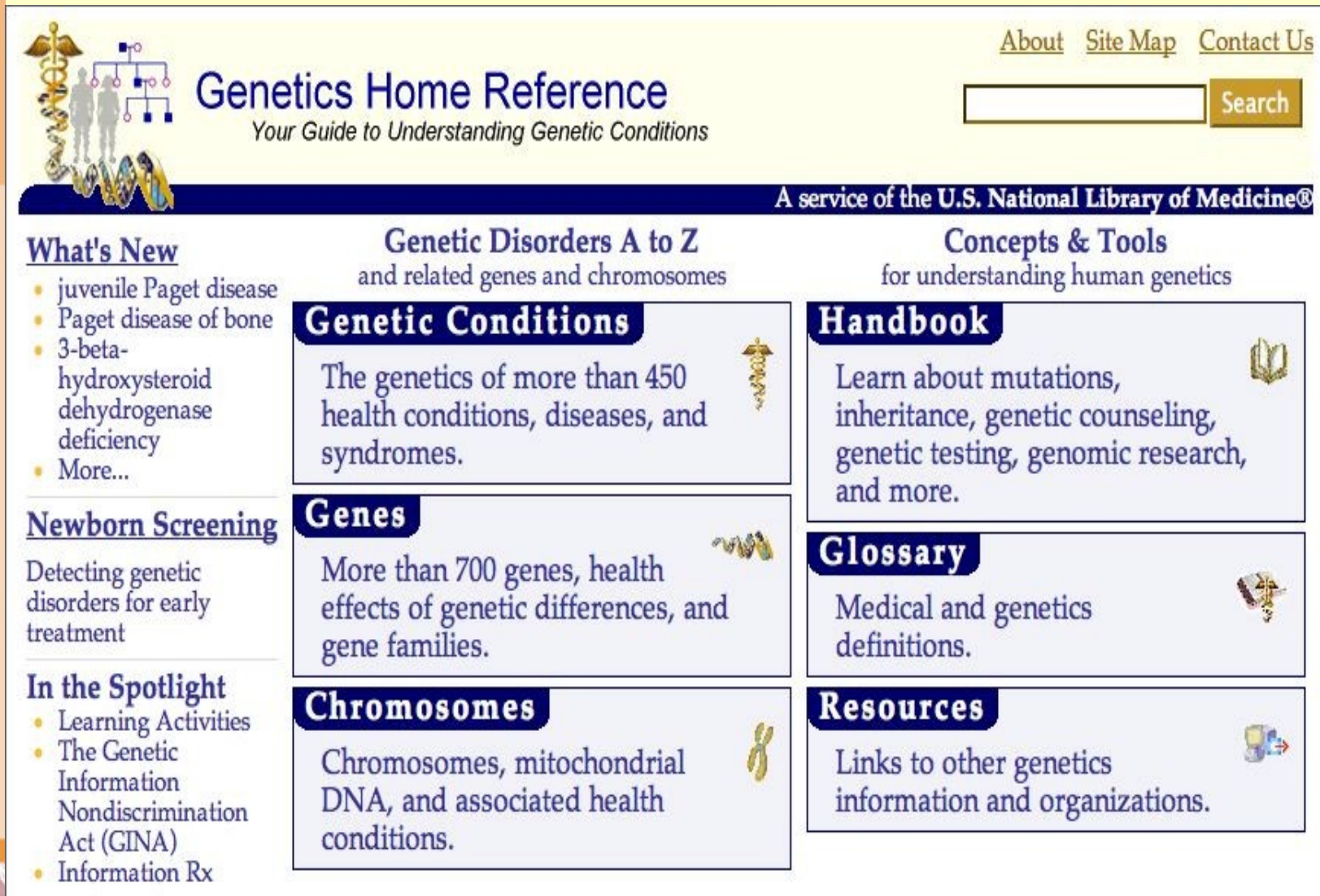
OMIM Home Page

http://www.ncbi.nlm.nih.gov/omim/



Genetics Home Reference

http://ghr.nlm.nih.gov/



Medline Plus

http://medlineplus.gov/

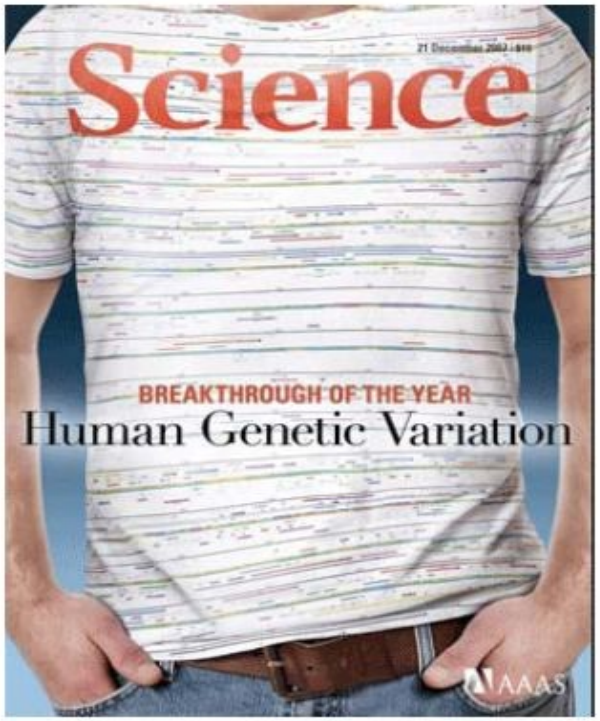
Common Gene Variation in Complex Disease

Case-control studies, comparing the frequencies of common gene variants can identify susceptibility and protective alleles. Many have multiple identified genes (*)

Phenotype	Gene	Variant
Peptic ulcer	ABO	B
IDDM*	HLA	DR3,4
Alzheimer dementia	APOE	E4
Deep venous thrombosis*	F5	Leiden
*Falciparum malaria**	HBB	β^S
AIDS*	CCR5	Δ32
Colorectal cancer*	APC	3920A
NIDDM*	PPARγ	12A

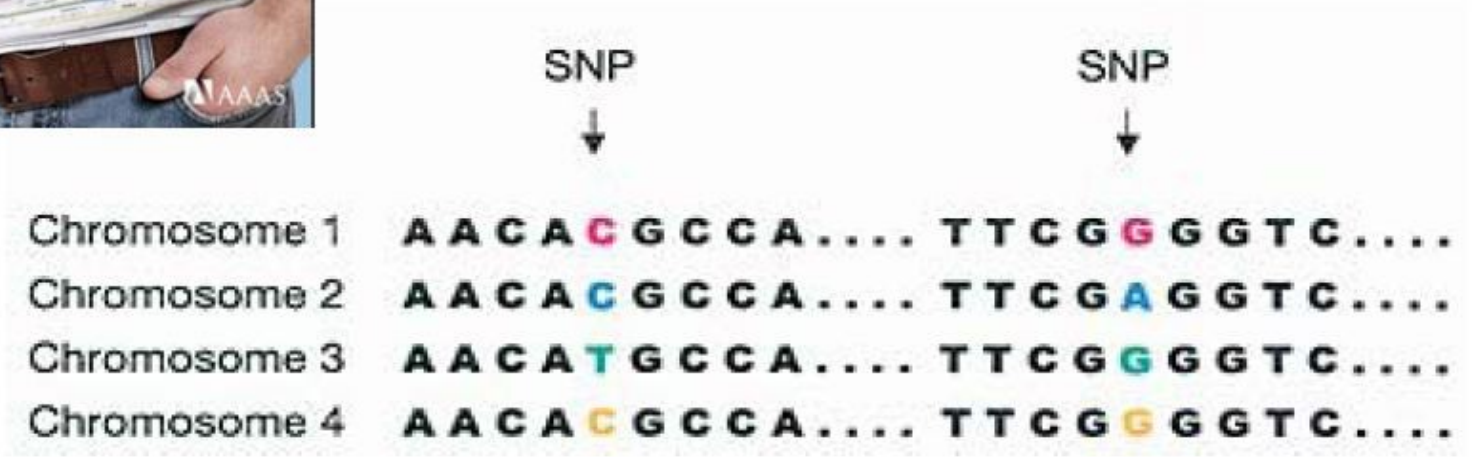
2007 SCIENTIFIC BREAKTHROUGH OF THE YEAR

Science Magazine, December 21, 2007



"It's all about me!"

Single Nucleotide Polymorphisms (SNPs)



http://www.sciencemag.org/cgi/content/full/318/5858/1842

International HapMap Project

http://www.hapmap.org/

中文 | English | Français | 日本語 | Yoruba

The International HapMap Project is a partnership of scientists and funding agencies from Canada, China, Japan, Nigeria, the United Kingdom and the United States to develop a public resource that will help researchers find genes associated with human disease and response to pharmaceuticals. See "**About the International HapMap Project**" for more information.

Project Information

About the Project
HapMap Publications
HapMap Tutorial
HapMap Mailing List
HapMap Project Participants
HapMap Mirror Site in Japan

Project Data

HapMap Genome Browser (Phase 1, 2 & 3 - merged genotypes & frequencies)
HapMap Genome Browser (Phase 3 - genotypes, frequencies & LD)
HapMap Genome Browser (Phase 1 & 2 - full dataset)
GWAs Karyogram
HapMart
HapMap FTP
Bulk Data Download
Data Freezes for Publication
ENCODE Project
Guidelines For Data Use

News

- 2009-12-14: **Notice to Haploview users**

  Recently, there are several questions about Haploview data format errors, and these errors were observed when users tried to analyze HapMap release 27 data dumped from HapMap. The current Haploview version (4.1) does not work with release 27 data. Haploview will generate a software error similar to "Hapmap data format error: NA06984" when trying to open the data.

  The r27 data format will be supported by next Haploview version. There is a beta test version that is supposed to work and it can be obtained from http://www.broadinstitute.org/haploview/haploview-downloads. But since it is NOT an official release version, please use it base on your own judgment.

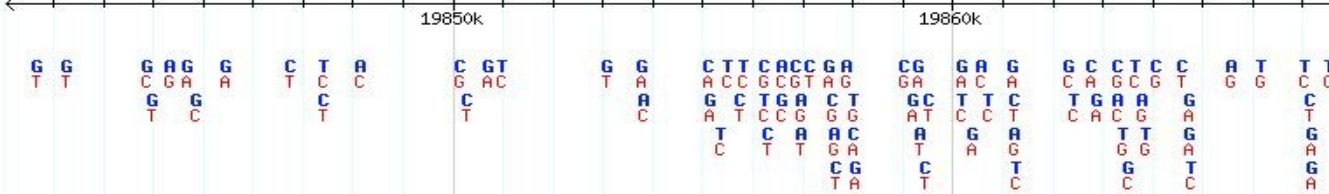
- 2009-12-10: **Corrected HapMap3 phased haplotypes available for chromosome X**

  Phased haplotypes for consensus HapMap3 release 2 data for chromosome X has been corrected and the new data are now **available for bulk download.** Sorry for any inconvenience this might have caused.

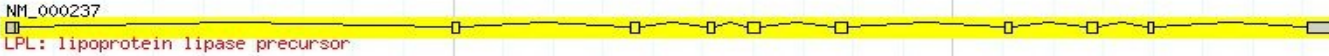
- 2009-12-02: **HapMap3 phased haplotypes available for chromosome X**

  Phased haplotypes for consensus HapMap3 release 2 data has been phased for chromosome X and are now available for bulk download. [Update: The downloading was disabled because several users have found that there are repeating data in some of the chrX phasing data files. The data source is being contacted and the downloading will be enabled as soon as the problem is cleared.]

Genotyped SNPs



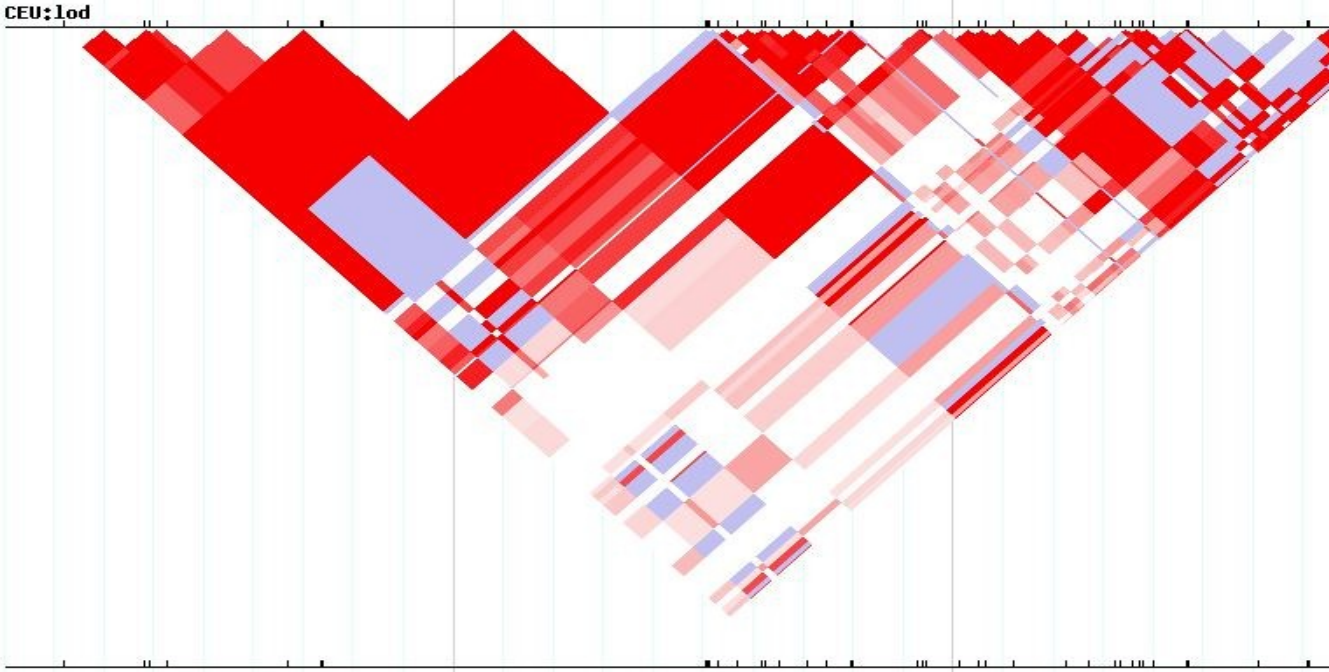
Entrez genes



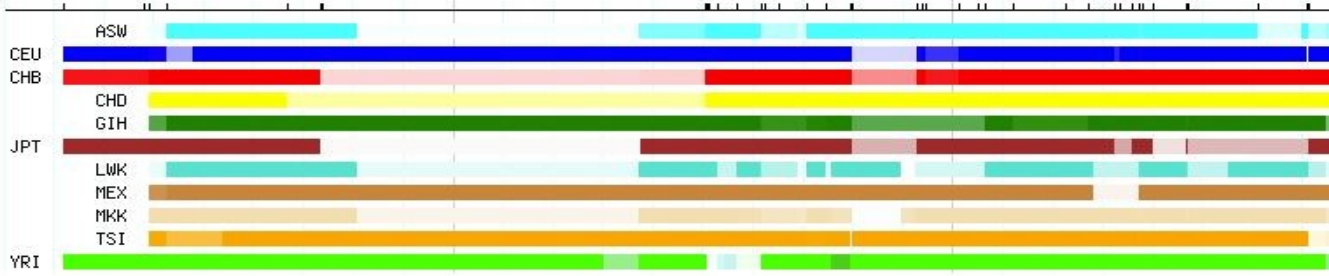
Reactome pathways

Genome-Wide Association studies (NHGRI Catalog)

LD Plot

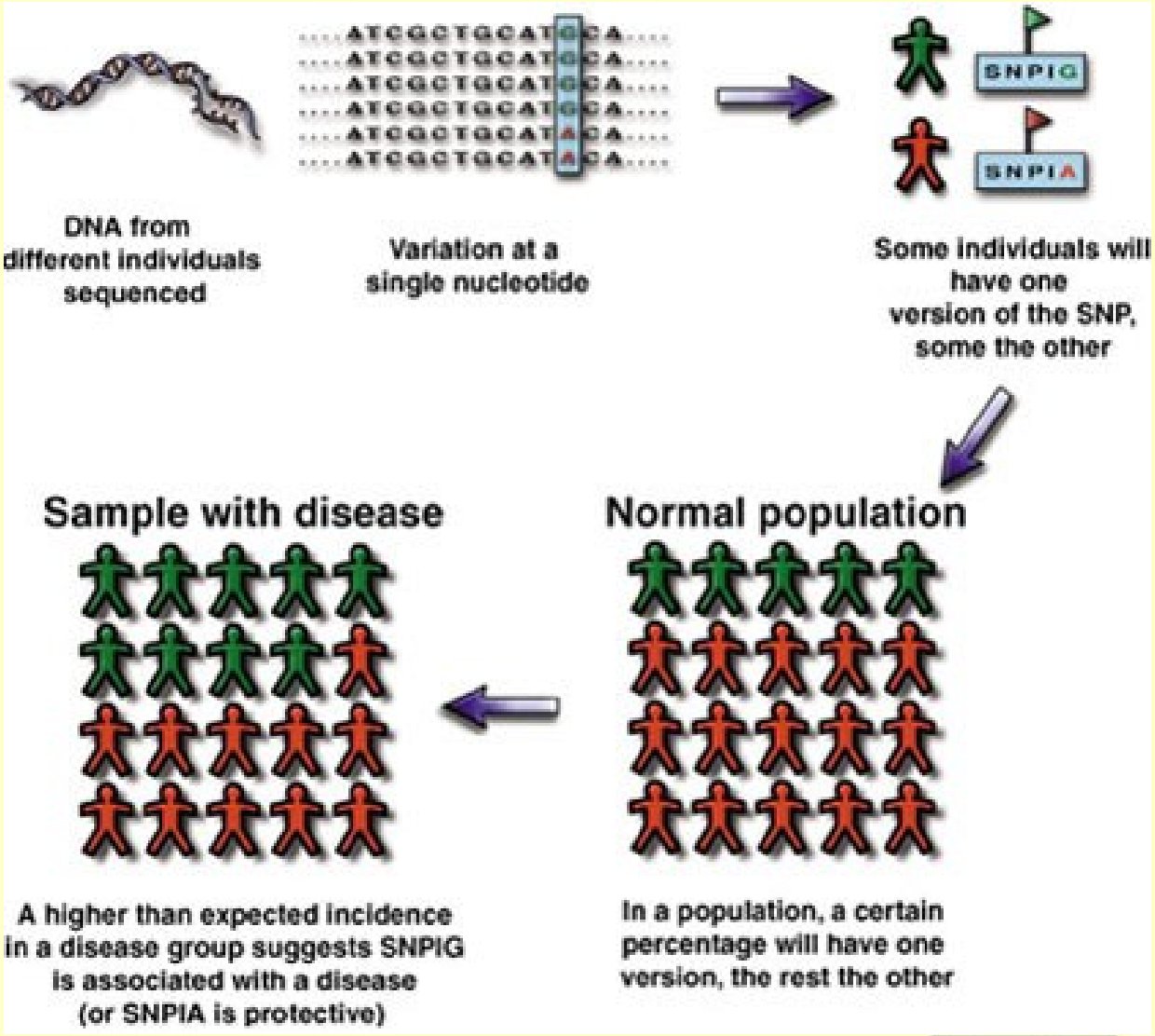


LD Heat Plot



Clear highlighting

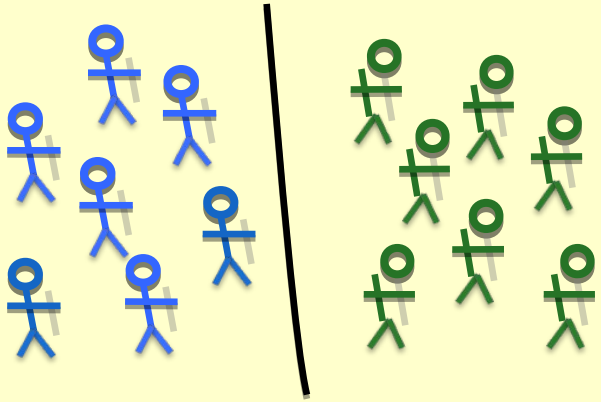
Using SNPs to Track Predisposition to Disease and other Genetic Traits



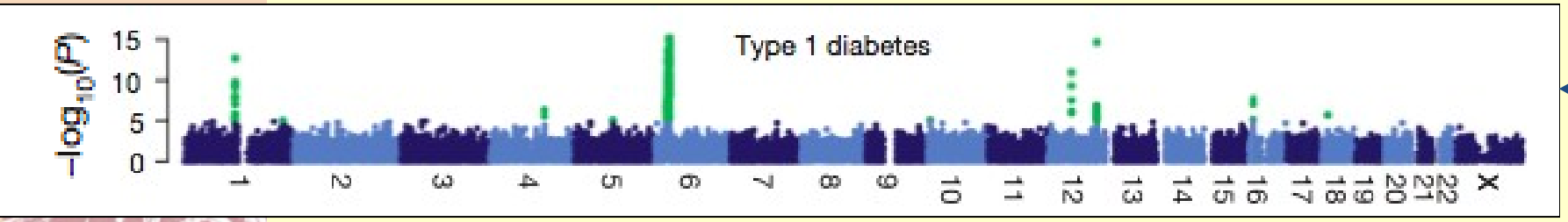
GWAS: Genome-Wide Association Study
A Brief Primer

Control Population

Disease Population



SNP chip

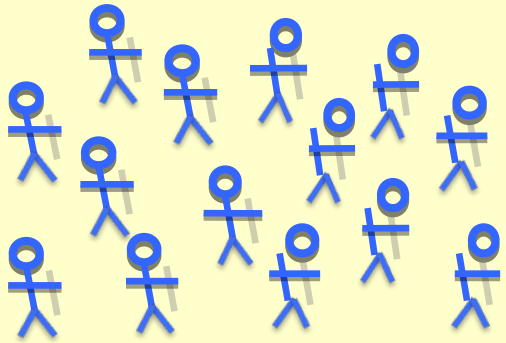


WTCCC, Nature 2007

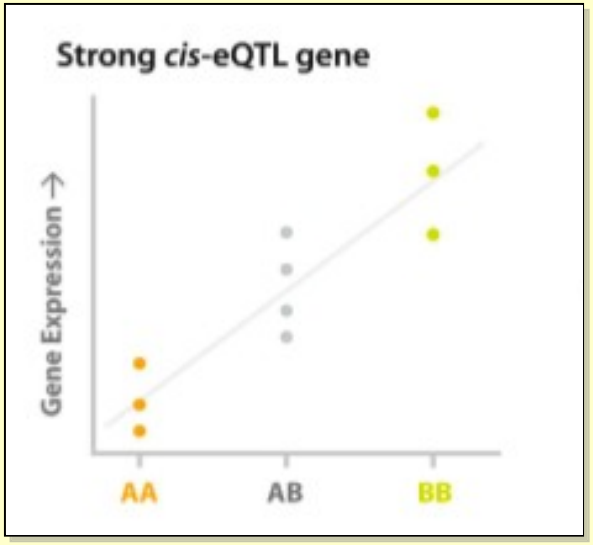
Thanks to Daniel Newburger

A Quantitative Gene-Expression Association

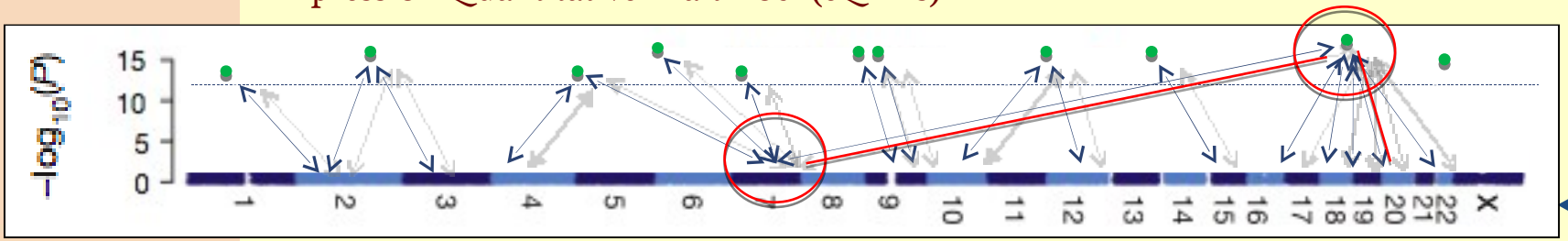
Sample Population



cDNA Levels



Expression Quantitative Trait Loci (eQTLs)



Modified from WTCCC, Nature 2007

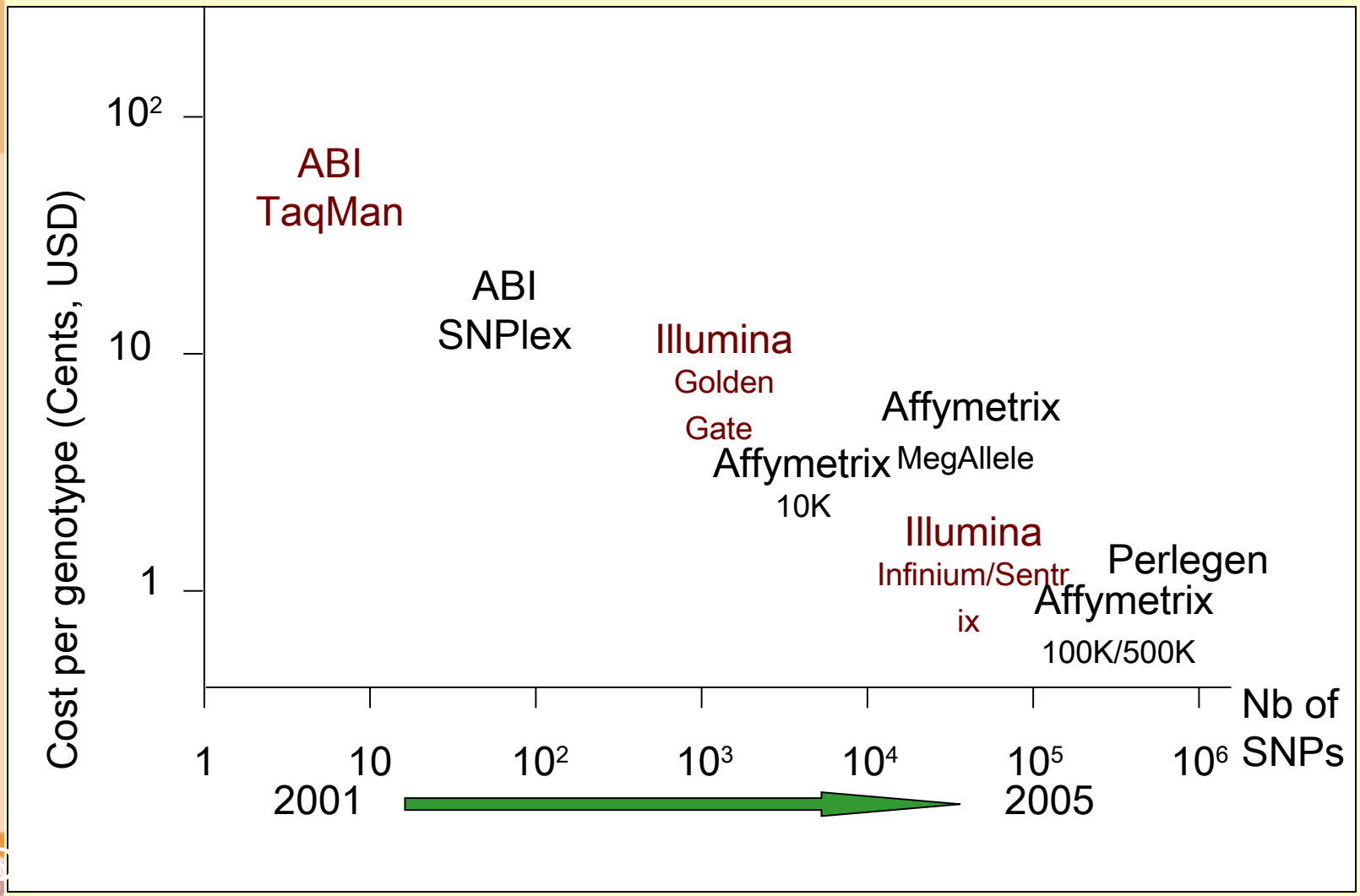
Thanks to Daniel Newburger

Genome-Wide Association Approach to Common and Complex Diseases

- Identify all 10 million common SNPs
- Collect 1,000 cases and 1,000 controls
- Genotype all DNAs for all SNPs
- That adds up to 20 billion genotypes

- In 2002, this approach cost 50 cents a genotype.
- That's $10 billion for each disease – completely out of the question

Progress in Genotype Technology



Genome-Wide Association Approach to Common and Complex Diseases

- Identify an optimum set of 300,000 tag SNPs
- Collect 1,000 cases and 1,000 controls
- Genotype all DNAs for all SNPs
- That adds up to 600 million genotypes

- In 2008, genotyping dropped to $0.0010, amounting to $600,000 for each disease

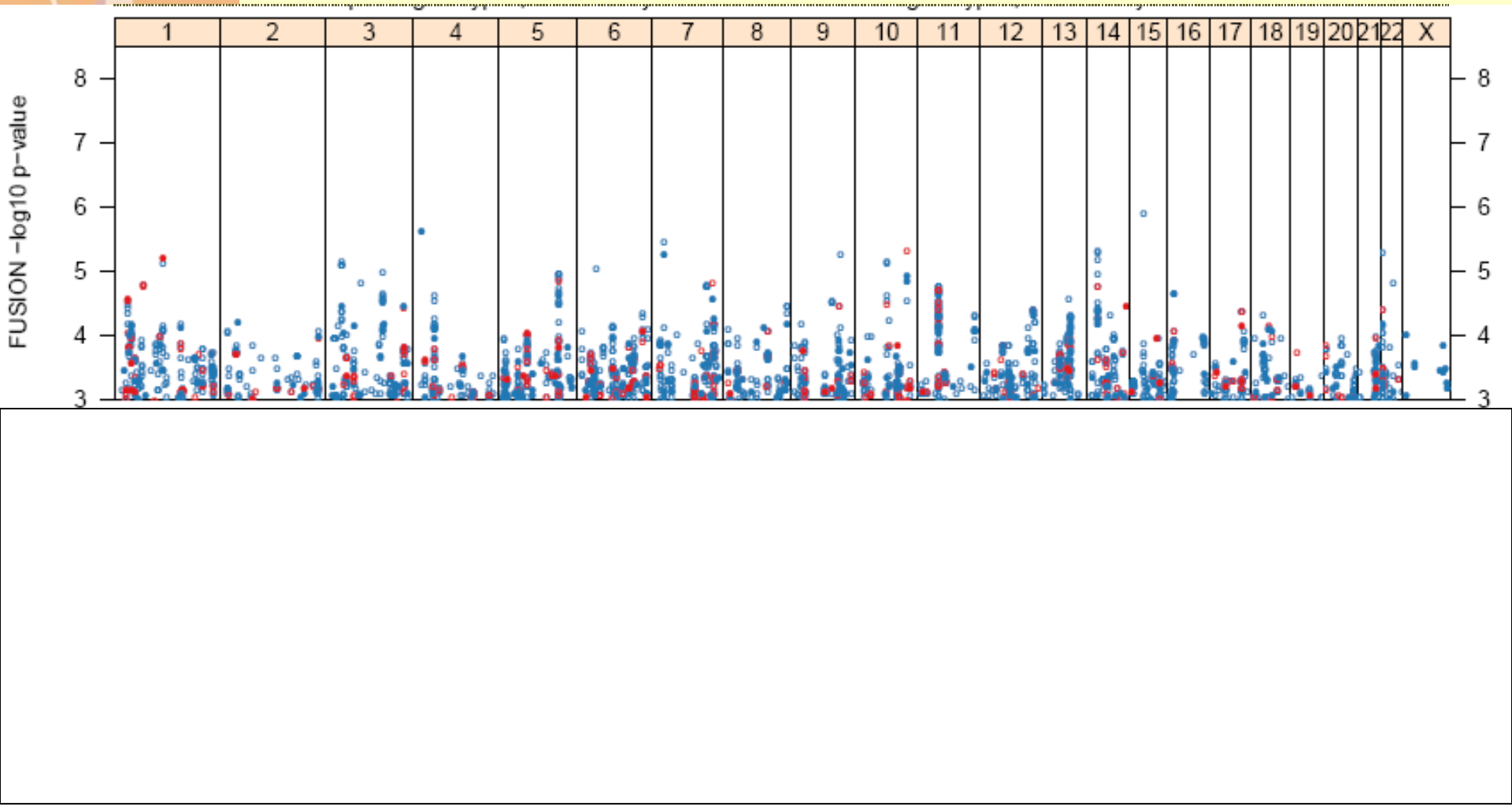
The FUSION Study
Finland-United States Investigation of NIDDM

n = 10,068

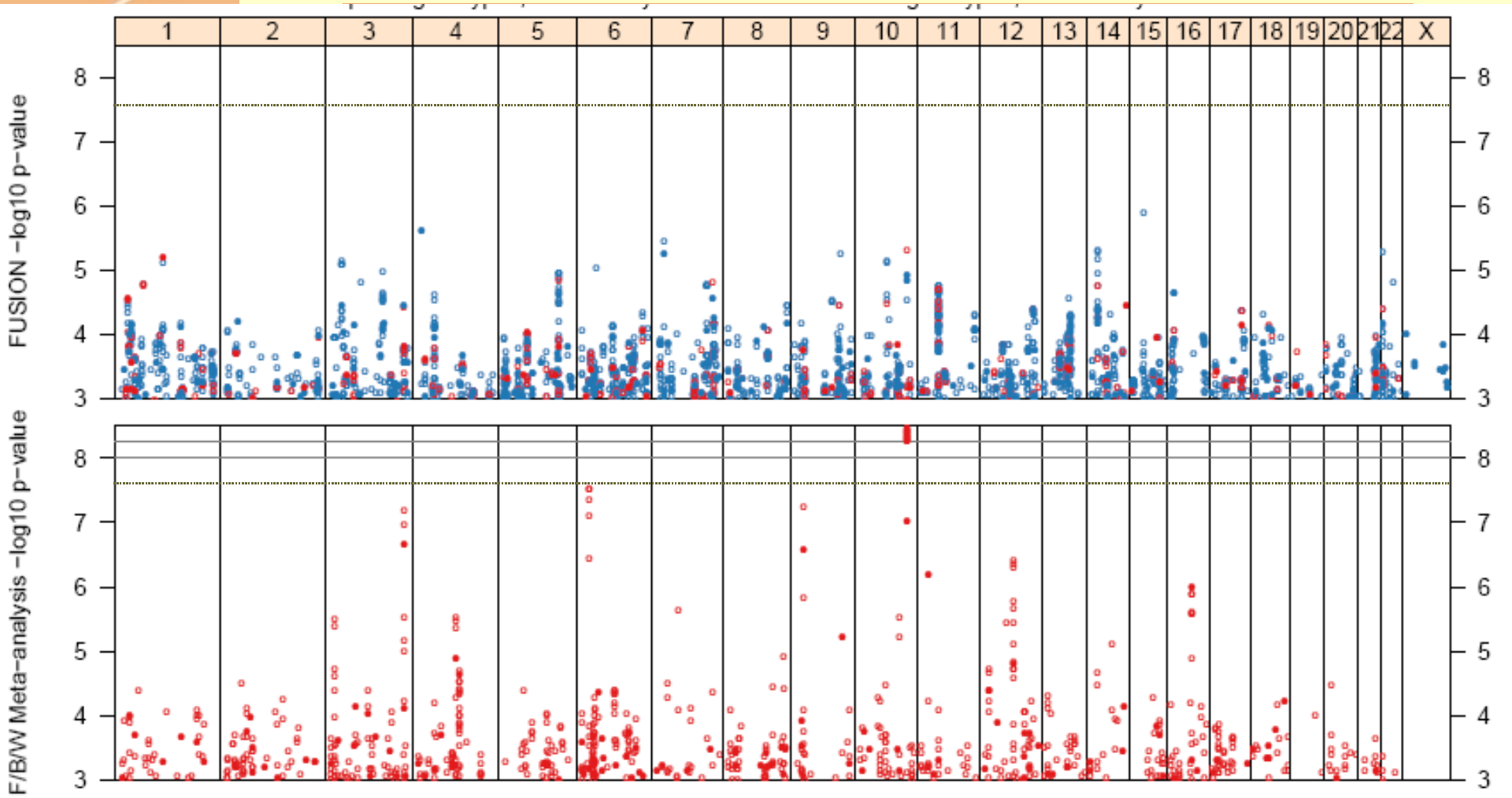
- Subject Recruitment and Clinical Testing
  - National Public Health Institute
  - Helsinki, Finland
- Molecular Genetics
  - National Human Genome Research Institute, Bethesda, MD
  - University of North Carolina, Chapel Hill, NC
- Biochemical Measurements
  - USC Keck School of Medicine, Los Angeles, CA
- Statistical Analysis
  - University of Michigan School of Public Health, Ann Arbor, MI

Results of Genome-Wide Association of Type 2 Diabetes with 317,503 SNPs

Stage 1: FUSION only (1161 cases + 1174 controls)

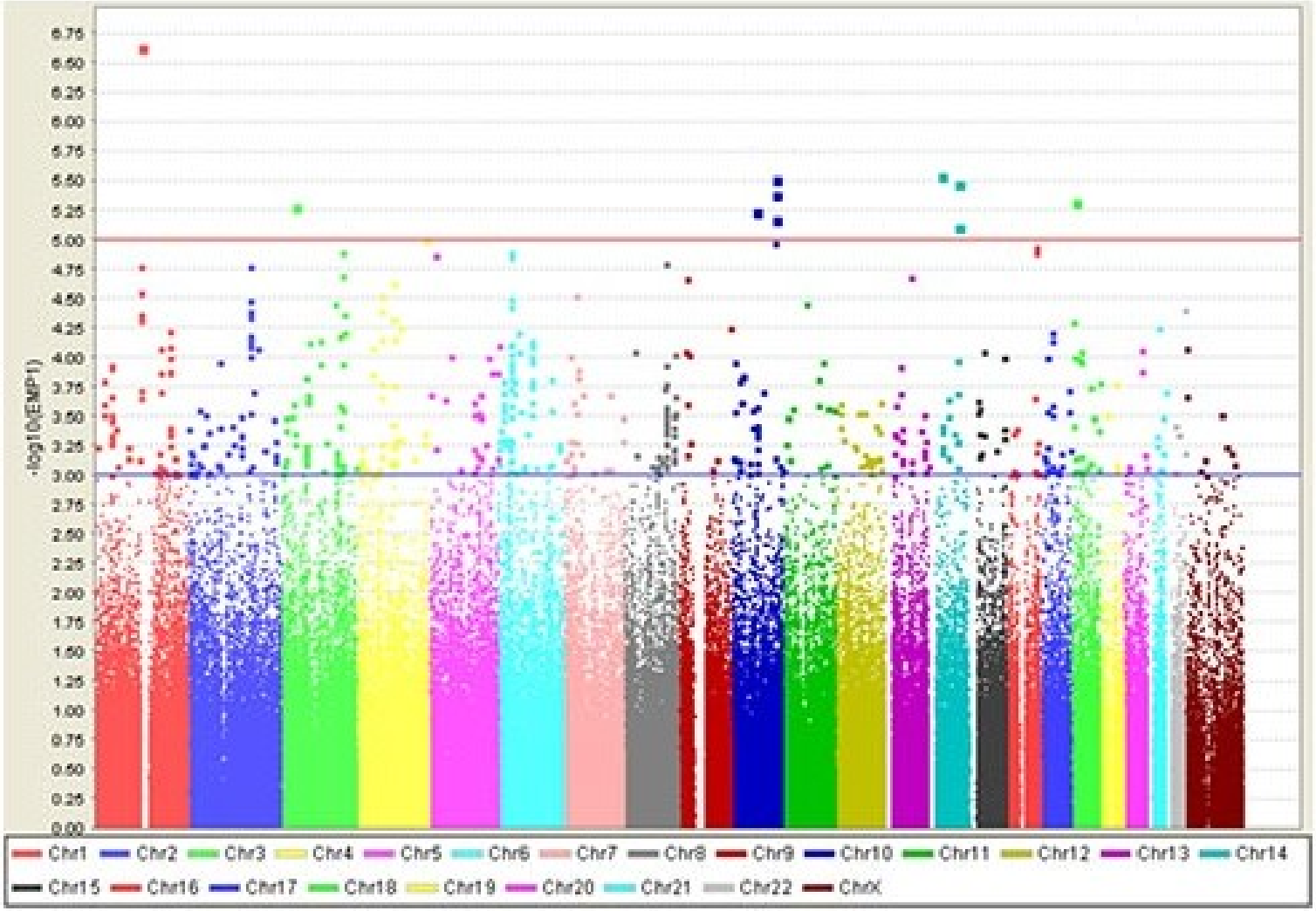


Results of Genome-Wide Association of Type 2 Diabetes with 317,503 SNPs



Stage 2 – FUSION + DGI + WTCCC (4549 cases + 5579 controls)

Genome-Wide Scan for Type 2 Diabetes in a Scandinavian Cohort

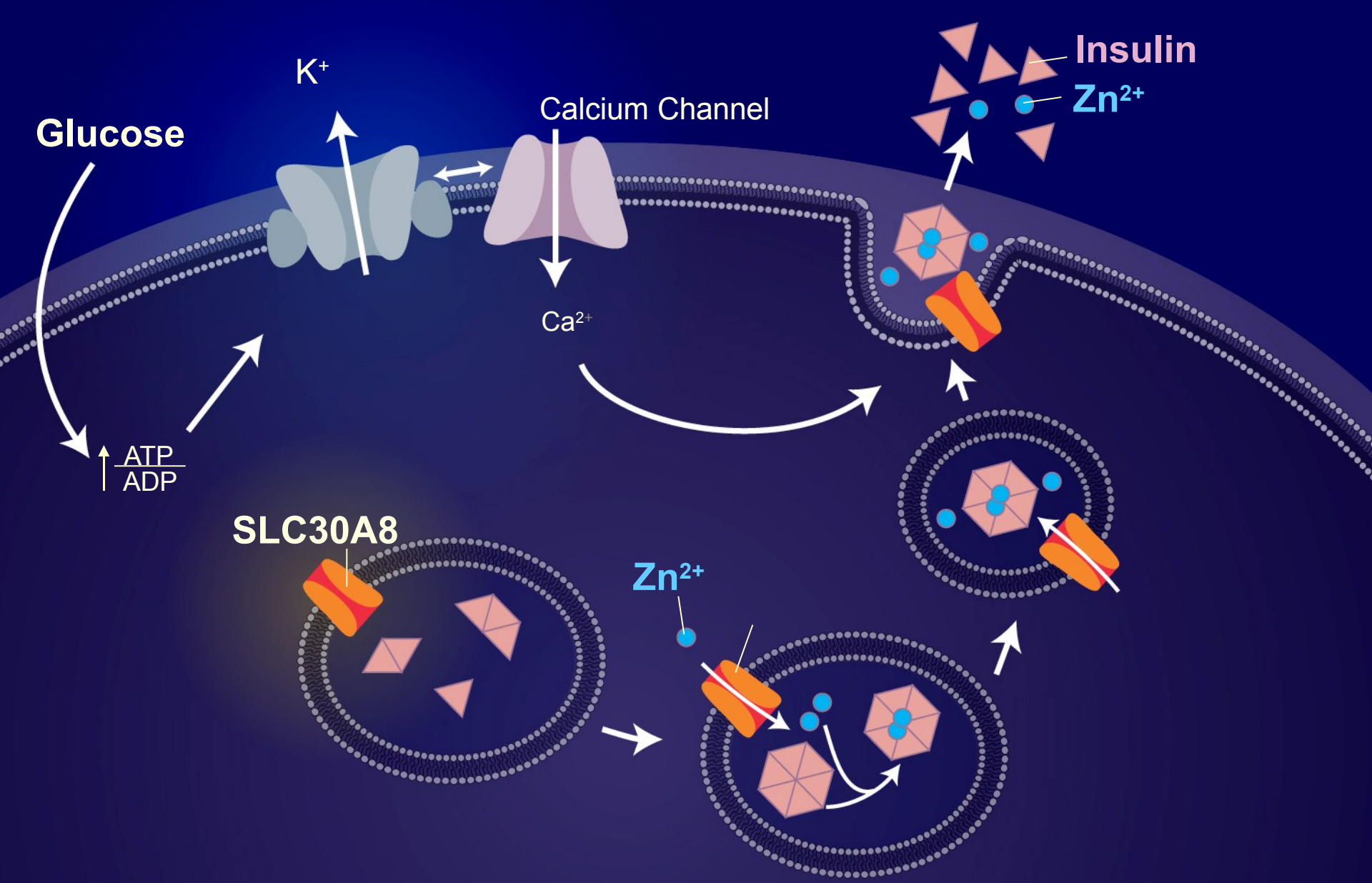


Top 10 Results From Combined Analysis

	FUSION		DGI		WTCCC/UKT2D		All Samples	
Gene	OR	p-value	OR	p-value	OR	p-value	OR	p-value
TCF7L2	1.34	1.3×10^{-8}	1.38	2.3×10^{-31}	1.37	6.7×10^{-13}	1.37	1.0×10^{-48}
IGF2BP2	1.18	2.1×10^{-4}	1.17	1.7×10^{-9}	1.11	1.6×10^{-4}	1.14	8.9×10^{-16}
CDKN2A/B	1.20	.0022	1.20	5.4×10^{-8}	1.19	4.9×10^{-7}	1.20	7.8×10^{-15}
FTO	1.11	0.016	1.03	0.25	1.23	7.3×10^{-14}	1.17	1.3×10^{-12}
CDKAL1	1.12	0.0095	1.08	0.0024	1.16	1.3×10^{-8}	1.12	4.1×10^{-11}
KCNJ11	1.11	0.013	1.15	1.0×10^{-7}	1.15	0.0013	1.14	6.7×10^{-11}
HHEX	1.10	0.026	1.14	1.7×10^{-4}	1.13	4.6×10^{-6}	1.13	5.7×10^{-10}
SLC30A8	1.18	7.0×10^{-5}	1.07	0.047	1.12	7.0×10^{-5}	1.12	5.3×10^{-8}
Chr 11	1.48	5.7×10^{-8}	1.16	0.12	1.13	0.068	1.23	4.3×10^{-7}
PPARG	1.20	0.0014	1.09	0.019	1.23	0.0013	1.14	1.7×10^{-6}

Top 10 Results From Combined Analysis

	FUSION		DGI		WTCCC/UKT2D		All Samples	
Gene	OR	p-value	OR	p-value	OR	p-value	OR	p-value
TCF7L2	1.34	1.3×10^{-8}	1.38	2.3×10^{-31}	1.37	6.7×10^{-13}	1.37	1.0×10^{-48}
IGF2BP2	1.18	2.1×10^{-4}	1.17	1.7×10^{-9}	1.11	1.6×10^{-4}	1.14	8.9×10^{-16}
CDKN2A/B	1.20	.0022	1.20	5.4×10^{-8}	1.19	4.9×10^{-7}	1.20	7.8×10^{-15}
FTO	1.11	0.016	1.03	0.25	1.23	7.3×10^{-14}	1.17	1.3×10^{-12}
CDKAL1	1.12	0.0095	1.08	0.0024	1.16	1.3×10^{-8}	1.12	4.1×10^{-11}
KCNJ11	1.11	0.013	1.15	1.0×10^{-7}	1.15	0.0013	1.14	6.7×10^{-11}
HHEX	1.10	0.026	1.14	1.7×10^{-4}	1.13	4.6×10^{-6}	1.13	5.7×10^{-10}
SLC30A8	1.18	7.0×10^{-5}	1.07	0.047	1.12	7.0×10^{-5}	1.12	5.3×10^{-8}
Chr 11	1.48	5.7×10^{-8}	1.16	0.12	1.13	0.068	1.23	4.3×10^{-7}
PPARG	1.20	0.0014	1.09	0.019	1.23	0.0013	1.14	1.7×10^{-6}

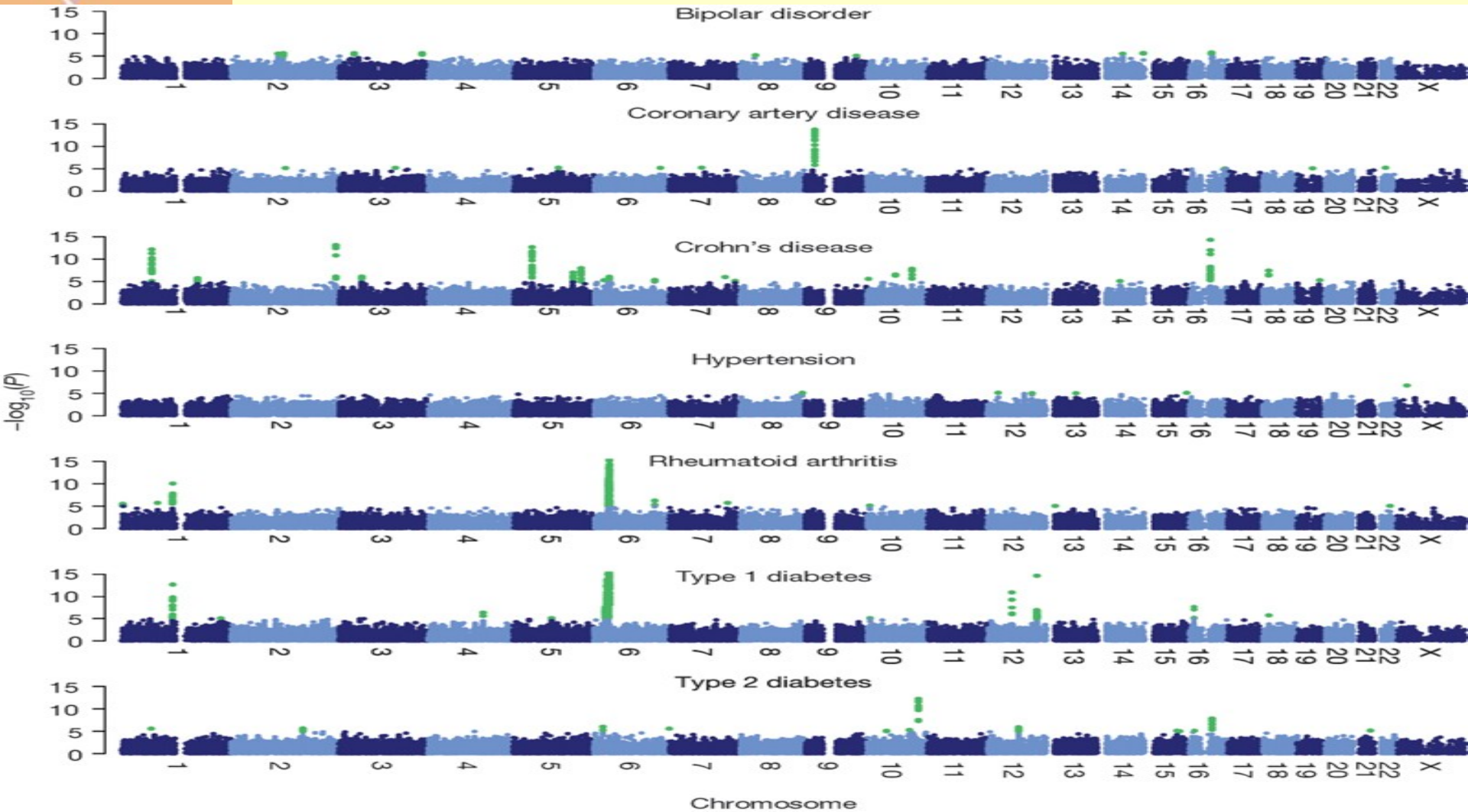


SLC30A8 – A Beta Cell Zinc Transporter

The Wellcome Trust Case Control Consortium

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

Nature 447, 661-678 (7 June 2007)



2007: The Year of GWA Studies?

Hokusai

The Genomics Gold Rush

Disease	Gene or Loci	Date Reported
Prostate Cancer	8q24	April 1, 2007
Acute Lymphoblastic Leukemia	PAX 5 and others	April 12, 2007
Obesity	FTO	April 12, 2007
Multiple Solid Tumors	CASP8	April 22, 2007
Diabetes, Type II	CDKAL1 and 6 others	April 26,2007
Myocardial Infarction, Coronary Artery Disease	9p21	May 3, 2007
Breast Cancer	FGFR2, TNCR9, MAP3K1, LSP and others	May 27, 2007
Crohn's Disease	IRGM	June 7, 2007
Diabetes, Type I	12q24 and others	June 7, 2007
Bipolar Disorder	16p12	June 7, 2007
Rheumatoid Arthritis	6p21, 1p13	June 7, 2007
Celiac Disease	IL-2, IL-21	June 10, 2007
Atrial Fibrillation	4q25	July 1, 2007

The Genomics Gold Rush

Disease	Gene or Loci	Date Reported
Diabetes, Type II	WFS1	July 1, 2007
Prostate Cancer	TCF2; 17p	July 1, 2007
Asthma (childhood)	ORMDL3	July 4, 2007
Colon, Prostate Cancer	8q24	July 8, 2007
Diabetes, Type I	KIAA0350	July 15, 2007
Gallstone Disease	ABCG8	July 15, 2007
Restless Leg Syndrome	MEIS1, BTBD9, MAP2K5	July 18, 2007
Coronary Artery Disease	6q25, 2q36	July 18, 2007
Age-Related Macular Degeneration	CF3	July 18, 2007
HIV Host Control	HLA-B*5701	July 19, 2007
Multiple Sclerosis	IL7Rα; IL2Rα	July 28, 2007
Amyotrophic Lateral Sclerosis	FLJ10986	August 1, 2007
Diabetes, Type I	IL2Rα	August 5, 2007
Glaucoma	LOXL1	August 9, 2007
Rheumatoid Arthritis	TRAF1-C5	August 31, 2007

The Genomics Gold Rush

Disease	Gene or Loci	Date Reported
Colorectal Cancer	SMAD7	October 14, 2007
Ankylosing Spondylitis	ARTS1, IL23R	October 21, 2007
Autoimmune Thyroid Disease	TSHR, FCRL3	October 21, 2007
Rheumatoid Arthritis	6q23	November 4, 2007
Psoriasis	β-Defensin CNV	December 2, 2007
Systemic Lupus Erythematosus	TNFSF4	December 2, 2007
Amyotrophic Lateral Sclerosis	DPP6	December 16, 2007
Colorectal Cancer	CRAC1 (HMPS)	December 16, 2007
Systemic Lupus Erythematosus	PXK, KIAA1542, BANK1, C8orf-BLK, ITGAM	January 20, 2008
Lipoprotein Disorders	MLX1PL and Multiple Others	January 13, 2008
Hypercholesterolemia	CELSR2	February 9. 2008
Prostate Cancer	2p15, Xp11.22 and Others	February 10, 2008
Gout	SLC2A9	March 9, 2008
Schizophrenia	ERBB4, SLC1A3 and Others	March 27, 2008

The Genomics Gold Rush

Disease	Gene or Loci	Date
Colorectal Cancer	10p14,8q23.3,18q21,11q23	March 30, 2008
Diabetes, Type 2	JA2F1 and others	March 30, 2008
Nicotine Add, Lung Ca, PAD	15q25	April 3, 2008
Hypertension	SLC12A3, SLC12A1,KCNJ1	April 6, 2008
Crohn's Disease and Ulcerative Colitis	ECM1and others PTPN2, HERC2, STAT3	April 27, 2008
Breast Cancer (ER +)	5p12	April 27, 2008
Osteoporosis	RANKL1,OPG, ESR	April 29, 2008
Obesity	MC4R	May 4, 2008
Neuroblastoma	6p22	May 7, 2008
Melanoma and Basal Cell Ca	20q11.22, ASIP, TYR	May 18, 2008
Gastric Cancer	PSCA	May 18, 2008
Macular Degeneration	ARMS2	May 30, 2008
Alzheimer's Disease	CALHM1	June 27, 2008
Crohn's Disease	JAK2, CDKAL1, ITLN1, more	June 29, 2008
Obesity	PCSK1	July 7, 2008
Knee Osteoarthritis	DVWA	July 14, 2008
Statin Myopathy	SLCO1B1	July 24, 2008

The Genomics Gold Rush

Disease	Gene or Loci	Date
Restless Leg Syndrome	PTPRD	July 27, 2008
Schizophrenia	1q21, 15q13	July 31, 2008
Systemic Lupus Erythematosus	TNAIP3	August 1, 2008
Sarcoidosis	ANXA11	August 10, 2008
Bipolar Disorder	ANK3, CACNA1C	August 17, 2008
Diabetes, Type II	KCNQ1	August 17, 2008
Crohn's Disease	IRGM	August 24, 2008
Prostate Cancer	HNF1B	August 31, 2008
CLL	2q13, 2q37, and others	August 31, 2008
Pediatric Inflammatory Bowel Dz	20q13, 21q22	August 31, 2008
Rheumatoid Arthritis	CD40, CD244, 10p15, 12q13, 22q13	September 14, 2008
Bladder Cancer	8q24	September 14, 2008
ESRD, Focal Glomerulosclerosis	MYH9	September 14, 2008
Narcolepsy	CPT1B, CHKB	September 28, 2008
Fatty Liver Disease (non-EtOH)	PNPLA3	September 28, 2008
Gout	SLC2A9, SLC17A3	October 1, 2008

The Genomics Gold Rush

Disease	Gene or Loci	Date
Male Pattern Baldness	20p11	October 12, 2008
Basal Cell Carcinoma	1p36, 1q42	October 12, 2008
Asthma	17q21	October 15, 2008
Lung Cancer	5p15, 6p21	November 2, 2008
Diabetes, Type 1	4q27, BACH2, PRKCQ	November 2 ,2008
Multiple Sclerosis	KIF1B	November 9, 2008
Intracranial Aneurysm	SOX17, 2p33	November 9, 2008
Colon Cancer	BMP4, CDH1, RHPN2, 20p12	November 16, 2008

Catalog of GWAS Studies
http://www.genome.gov/26525384

genome.gov
National Human Genome Research Institute
National Institutes of Health

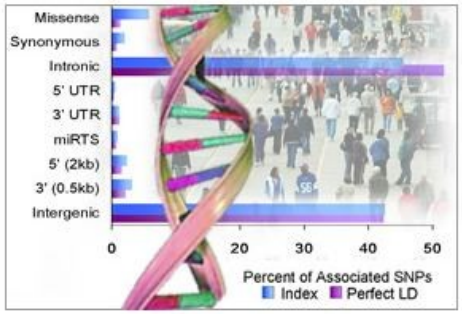
Google™ Search SEARCH

Research Funding | Research at NHGRI | Health | Education | Issues in Genetics | Newsroom | Careers & Training | About | For You

Home › About › Organizational Structure › About the Office of the Director › News Features from the Office of the Director › **Online GWAS Catalog Helps Guide Disease Research**

Online GWAS Catalog Helps Guide Disease Research

Share this page | Print



Researchers who want to sift through the biomedical literature to find genome-wide association results relevant to their research pursuits face an enormous challenge. But, thanks to the efforts of a dedicated team of National Human Genome Research Institute (NHGRI) scientists, they now have an online resource that can make the task a bit less daunting.

Genome-wide association studies, commonly called GWAS, efficiently scan markers across the DNA, or genomes, of large groups of people looking for variations between individuals with and without a health condition. Over the past few years, this approach has successfully plucked from our DNA code hundreds of the pesky one-letter genetic variations that contribute to the risk of common health conditions, such as obesity and Type 2 diabetes.

The recent deluge of GWAS results coming from these studies is exactly why NHGRI created a resource called A Catalog of Published Genome-Wide Association Studies. The catalog, which has been frequently cited in scientific publications, contains descriptive and association data on hundreds of published common genetic variations and their relationship with nearly 100 complex diseases and traits.

On a weekly basis, epidemiologists from NHGRI's Office of Population Genomics manually curate information from published GWAS and add them to the catalog. Researchers can search the catalog by journal name, first author, disease/trait, statistical significance of association and other categories. Data from the entire catalog can also be directly downloaded as an Excel spreadsheet.

See Also:

Potential etiologic and functional implications of genome-wide association loci for human diseases and traits PDF
PNAS, May 18, 2009

Keywords: what's this?

- GWAS
- catalog
- markers
- DNA
- genomes
- Office of Population Genomics
- OD News Features

Catalog of GWAS Studies
http://www.genome.gov/26525384

genome.gov
National Human Genome Research Institute
National Institutes of Health

Google™ Search SEARCH

Research Funding | Research at NHGRI | Health | Education | Issues in Genetics | Newsroom | Careers & Training | About | For You

Home › About › Organizational Structure › About the Office of the Director › Office of Population Genomics › **A Catalog of Published Genome-Wide Association Studies**

Office of Population Genomics

Share this page | Print

Overview | **A Catalog of Genome-Wide Association Studies** | Research Programs | Publications | Meetings & Workshops | Notices & Funding Opportunities | Staff Biographies | Contact

Keywords: what's this?
- GWAS
- GWAS studies
- OPG
- Teri Manolio
- RSS Web Feeds

A Catalog of Published Genome-Wide Association Studies

Potential etiologic and functional implications of genome-wide association loci for human diseases and traits

Click here to read our recent *Proceedings of the Academy of Sciences (PNAS)* article on catalog methods and analysis.

Go to the Catalog

Published Genome-Wide Associations
Credit: Darryl Leja and Teri Manolio

View as PDF

View PowerPoint Slide

The genome-wide association study (GWAS) publications listed here include only those attempting to assay at least 100,000 single nucleotide polymorphisms (SNPs) in the initial stage. Publications are organized from most to least recent date of publication, indexing from online publication if available. Studies focusing only on candidate genes are excluded from this catalog. Studies are identified through weekly PubMed literature searches, daily NIH-distributed compilations of news and media reports, and occasional comparisons with an existing database of GWAS literature (HuGE Navigator).

SNP-trait associations listed here are limited to those with p-values < 1.0 x 10-5 (see full methods for additiona details). Multipliers of powers of 10 in p-values are rounded to the nearest single digit; odds ratios and allele frequencies are rounded to two decimals. Standard errors are converted to 95 percent confidence intervals where applicable. Allele frequencies, p-values, and odds ratios derived from the largest sample size, typically a combined analysis (initial plus replication studies), are recorded below if reported; otherwise statistics from the initial study sample are recorded. For quantitative traits, information on % variance explained, SD increment, o unit difference is reported where available. Odds ratios < 1 in the original paper are converted to OR > 1 for the alternate allele. Where results from multiple genetic models are available, we prioritized effect sizes (OR's or beta-coefficients) as follows: 1) genotypic model, per-allele estimate; 2) genotypic model, heterozygote estimate, 3) allelic model, allelic estimate.

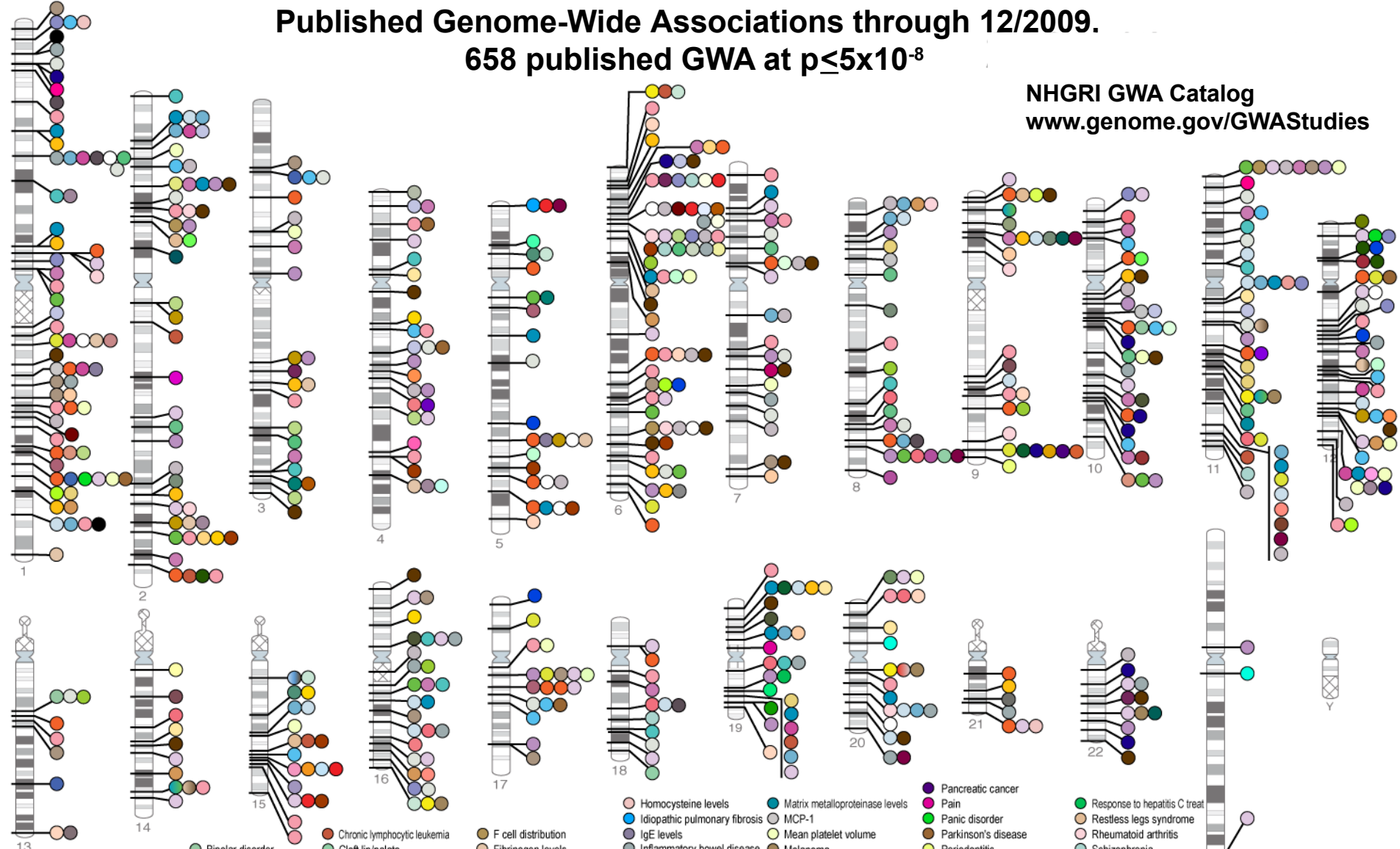
Catalog of GWAS Studies

As of 03/02/10, this table includes 496 publications and 2341 SNPs.

Date Added to Catalog (since 11/25/08)	First Author/Date/ Journal/Study	Disease/Trait	Initial Sample Size	Replication Sample Size	Region	Reported Gene(s)	Strongest SNP-Risk Allele	Risk Allele Frequency in Controls	P-value
03/02/10	Shi February 02, 2010 *Mol Psychiatry* Genome-wide association study of recurrent early-onset major depressive disorder	Major depressive disorder	1,020 European cases, 1,636 European controls	NR	18q22.1 21q21.2 1p13.3 5p13.2 2p23.2 3p14.2 1p13.3 7p15.3 13q21.33	*DSEL* *Intergenic* *GNAT2, GNAI3, AMPD2* *GDNF* *FAM179A, C2orf71* *FHIT, PTPRG* *ATXN7L2, SYPL2, CYB561D1* *SP4* *Intergenic*	rs17077540-G rs2828520-G rs6537837-T rs270545-G rs882632-T rs10514718-C rs12049330-G rs17144465-G rs9572423-G	0.11 0.31 0.17 0.69 0.29 0.94 0.14 0.04 0.88	2×10^{-7} 4×10^{-7} 1×10^{-6} 1×10^{-6} 2×10^{-6} 4×10^{-6} 6×10^{-6} 6×10^{-6} 9×10^{-6}
02/28/10	Petersen January 24, 2010 *Nat Genet* A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and	Pancreatic cancer	3,851 cases, 3,934 controls	NR	13q22.1 1q32.1 5p15.33	*KLF5, KLF12* *NR5A2* *CLPTM1L*	rs9543325-C rs3790844-T rs401681-T	0.37 0.76 0.45	3×10^{-11} 2×10^{-10} 7×10^{-7}

Published Genome-Wide Associations through 12/2009.
658 published GWA at $p \leq 5 \times 10^{-8}$

NHGRI GWA Catalog
www.genome.gov/GWAStudies



- Acute lymphoblastic leukemia
- Adiponectin levels
- Age-related macular degeneration
- AIDS progression
- Alcohol dependence
- Alzheimer disease
- Amyotrophic lateral sclerosis
- Arterial stiffness
- Asthma
- Atrial fibrillation
- Attention deficit hyperactivity disorder
- Autism
- Basal cell cancer
- Bipolar disorder
- Bilirubin
- Bladder cancer
- Blond or brown hair
- Blood pressure
- Blue or green eyes
- BMI, waist circumference
- Breast cancer
- Bone density
- C-reactive protein
- Cardiac structure/function
- Carnitine levels
- Carotenoid/tocopherol levels
- Celiac disease
- Chronic lymphocytic leukemia
- Cleft lip/palate
- Cognitive function
- Colorectal cancer
- Coronary disease
- Creutzfeldt-Jakob disease
- Crohn's disease
- Cutaneous nevi
- Dermatitis
- Drug-induced liver injury
- Eosinophil count
- Erythrocyte parameters
- Esophageal cancer
- Essential tremor
- Exfoliation glaucoma
- F cell distribution
- Fibrinogen levels
- Folate pathway vitamins
- Freckles and burning
- Gallstones
- Glioma
- Glycemic traits
- Hair color
- Hair morphology
- Heart rate
- Height
- Hepatitis
- Hirschsprung's disease
- HIV-1 control
- HDL cholesterol
- Homocysteine levels
- Idiopathic pulmonary fibrosis
- IgE levels
- Inflammatory bowel disease
- Intracranial aneurysm
- Iris color
- Iron status markers
- Ischemic stroke
- Juvenile idiopathic arthritis
- Kidney stones
- Leprosy
- LDL cholesterol
- Liver enzymes
- LP (a) levels
- Lung cancer
- Malaria
- Male pattern baldness
- Matrix metalloproteinase levels
- MCP-1
- Mean platelet volume
- Melanoma
- Menarche & menopause
- Multiple sclerosis
- Myeloproliferative neoplasms
- Narcolepsy
- Nasopharyngeal cancer
- Neuroblastoma
- Nicotine dependence
- Nonsyndromic cleft lip w/wo cleft palate
- Obesity
- Open personality
- Otosclerosis
- Other metabolic traits
- Ovarian cancer
- Pancreatic cancer
- Pain
- Panic disorder
- Parkinson's disease
- Periodontitis
- Peripheral arterial disease
- Phosphatidylcholine levels
- Primary biliary cirrhosis
- Prostate cancer
- Protein levels
- Psoriasis
- Pulmonary funct. COPD
- QT interval
- Quantitative traits
- Recombination rate
- Red vs.non-red hair
- Renal function
- Response to antipsychotic therapy
- Response to hepatitis C treat
- Restless legs syndrome
- Rheumatoid arthritis
- Schizophrenia
- Serum metabolites
- Skin pigmentation
- Soluble E-selectin
- Soluble ICAM-1
- Speech perception
- Sphingolipid levels
- Statin-induced myopathy
- Stroke
- Systemic lupus erythematosus
- Testicular germ cell tumor
- Thyroid cancer
- Total cholesterol
- Triglycerides
- Type 1 diabetes
- Type 2 diabetes
- Urate
- Venous thromboembolism
- Vitamin B12 levels
- Warfarin dose
- Weight
- White cell count
- YKL-40 levels

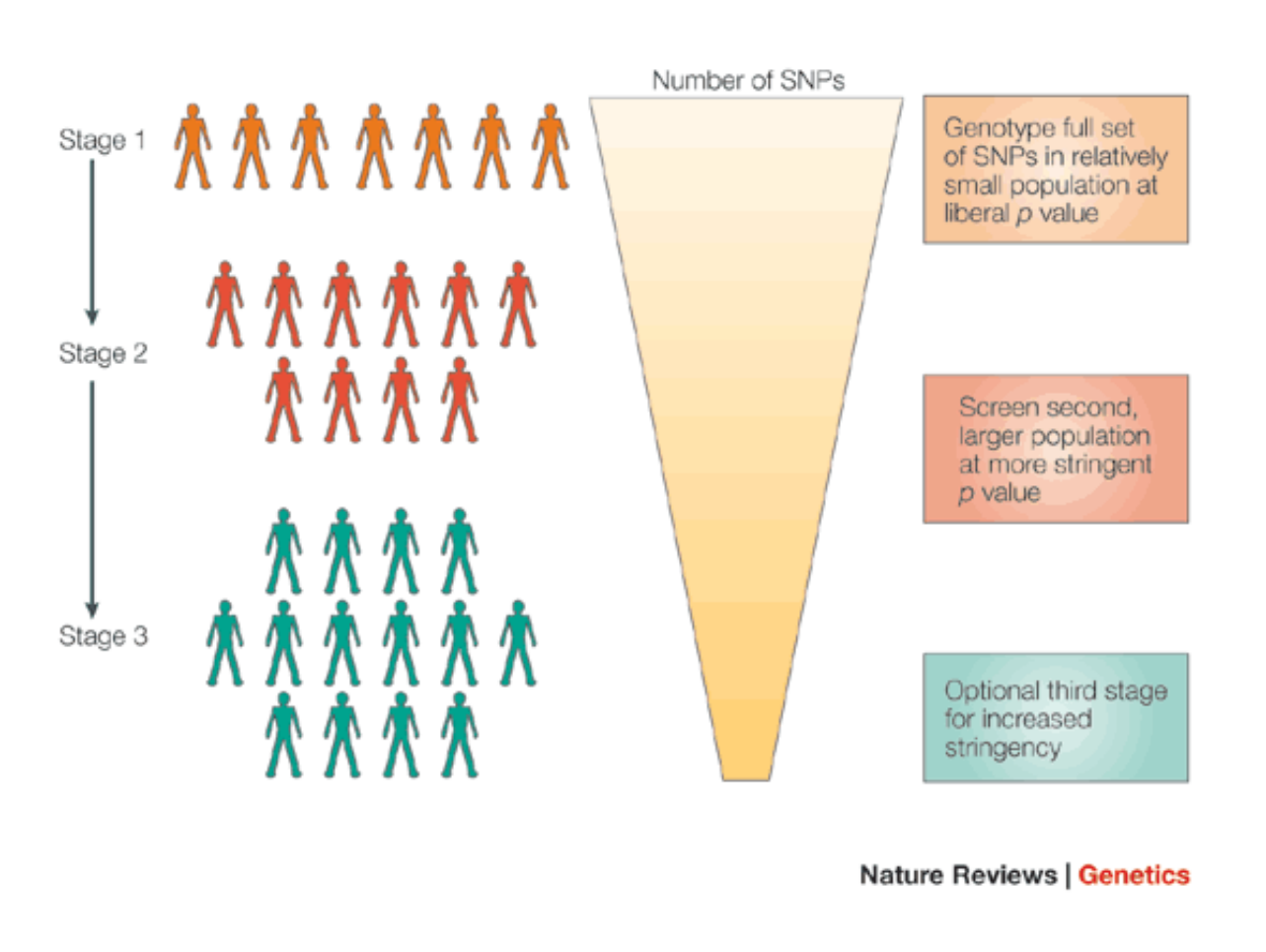
Study Designs Used in Genome-wide Association Studies

Table 1. Study Designs Used in Genome-wide Association Studies

	Case-Control	Cohort	Trio
Assumptions	Case and control participants are drawn from the same population Case participants are representative of all cases of the disease, or limitations on diagnostic specificity and representativeness are clearly specified Genomic and epidemiologic data are collected similarly in cases and controls Differences in allele frequencies relate to the outcome of interest rather than differences in background population between cases and controls	Participants under study are more representative of the population from which they are drawn Diseases and traits are ascertained similarly in individuals with and without the gene variant	Disease-related alleles are transmitted in excess of 50% to affected offspring from heterozygous parents
Advantages	Short time frame Large numbers of case and control participants can be assembled Optimal epidemiologic design for studying rare diseases	Cases are incident (developing during observation) and free of survival bias Direct measure of risk Fewer biases than case-control studies Continuum of health-related measures available in population samples not selected for presence of disease	Controls for population structure; immune to population stratification Allows checks for Mendelian inheritance patterns in genotyping quality control Logistically simpler for studies of children's conditions Does not require phenotyping of parents
Disadvantages	Prone to a number of biases including population stratification Cases are usually prevalent cases, may exclude fatal or short episodes, or mild or silent cases Overestimate relative risk for common diseases	Large sample size needed for genotyping if incidence is low Expensive and lengthy follow-up Existing consent may be insufficient for GWA genotyping or data sharing Requires variation in trait being studied Poorly suited for studying rare diseases	May be difficult to assemble both parents and offspring, especially in disorders with older ages of onset Highly sensitive to genotyping error

Pearson, T. A. et al. JAMA 2008;299:1335-1344

Replication A Must



Replication

Replication

Replication

Hirschhorn & Daly Nat. Genet. Rev. 6: 95, 2005
NCI-NHGRI Working Group on Replication Nature 447: 655, 2007

Examples of Multistage Designs in Genome-wide Association Studies

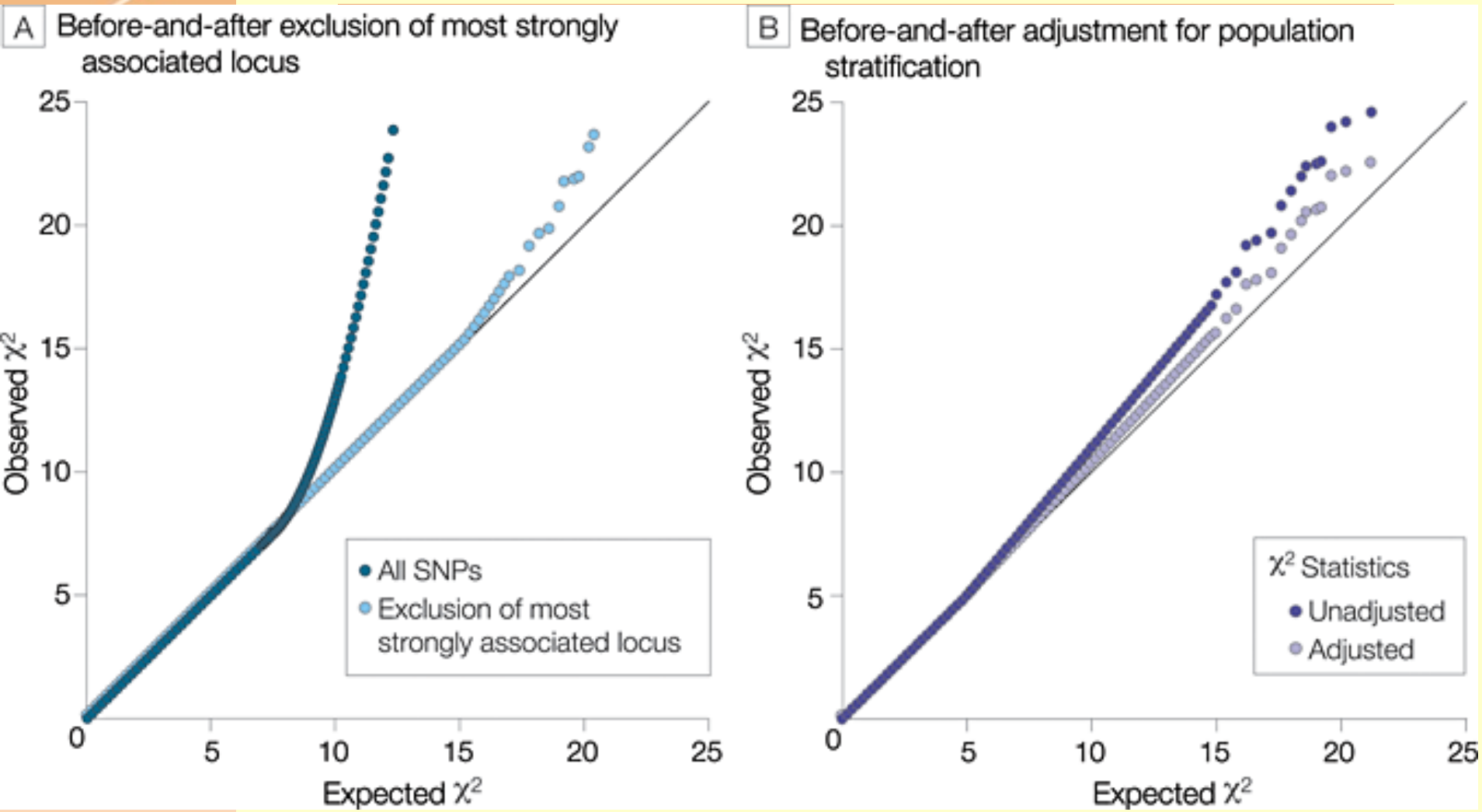
Table 2. Examples of Multistage Designs in Genome-wide Association Studies[a]

	3-Stage Study[b]		4-Stage Study[c]	
Stage	Case Participants/ Control Participants	SNPs Analyzed	Case Participants/ Control Participants	SNPs Analyzed
1	400/400	500 000	2000/2000	100 000
2	4000/4000	25 000	2000/2000	1000
3	20 000/20 000	25	2000/2000	20
4			2000/2000	5

Abbreviation: SNP, single-nucleotide polymorphism.
[a] Based on hypothetical data.
[b] Five SNPs associated with disease.
[c] Two SNPs associated with disease.

Pearson, T. A. et al. JAMA 2008;299:1335-1344

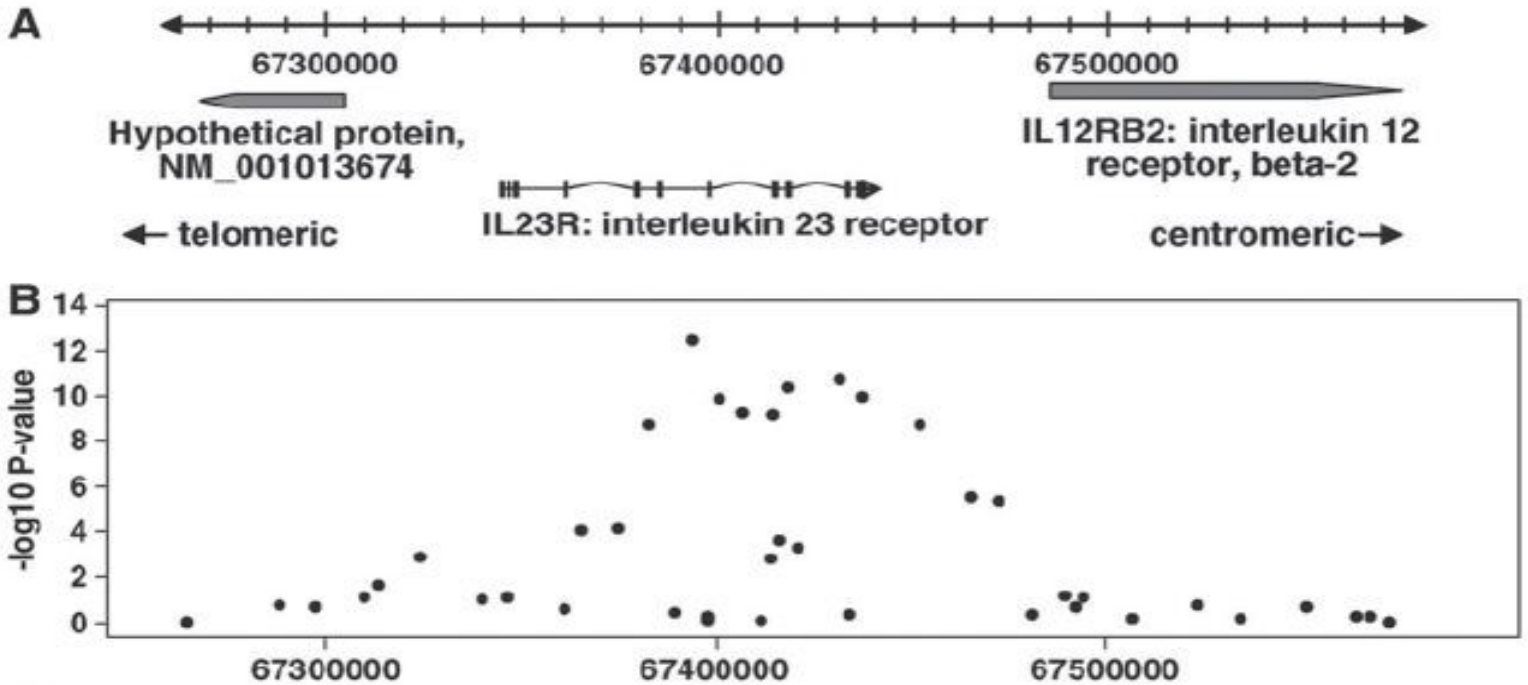
Hypothetical Quantile-Quantile Plots in Genome-wide Association Studies



Pearson, T. A. et al. JAMA 2008;299:1335-1344

Interleukin 23R & Inflammatory Bowel Disease

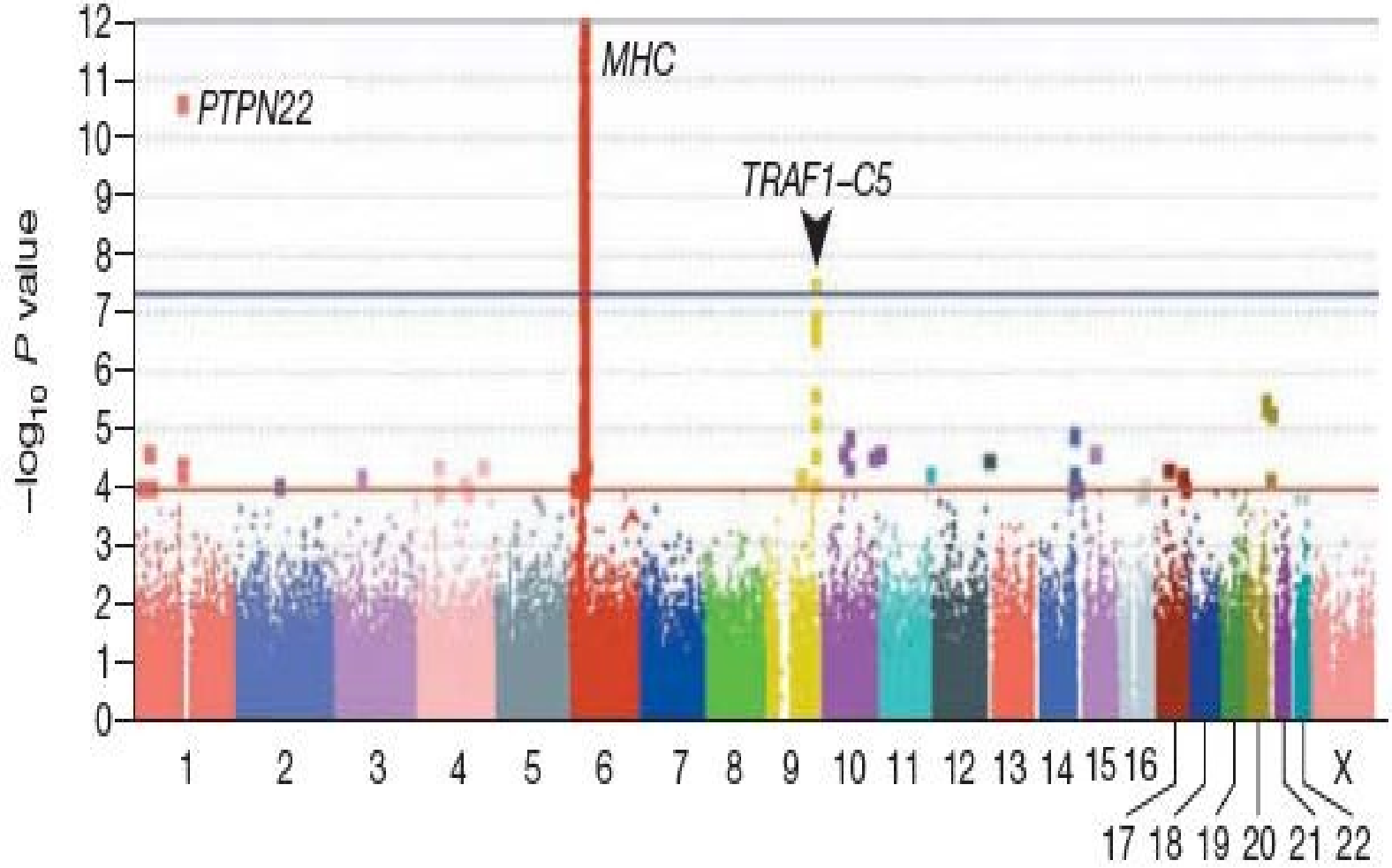
Figure 2. Associations in the *IL23R* Gene Region Identified by a Genome-wide Association Study of Inflammatory Bowel Disease



Pearson, T. A. et al. JAMA 2008;299:1335-1344

Genome-Wide Associations in Rheumatoid

Figure 3. Genome-wide Association Findings in Rheumatoid Arthritis



Pearson, T. A. et al. JAMA 2008;299:1335-1344

	Number and Frequency of rs6983267 Alleles in Colorectal Cancer					Number and Frequency of rs6983267 Genotypes in Colorectal Cancer						
	C	T	χ^2 (1*df*)	*P* Value	OR	CC	CT	TT	χ^2 (2*df*)	*P* Value	OR	OR
Cases	875 (56.5)	675 (43.5)	24.8	6.3×10^{-7}	1.35[b]	250 (32.3)	375 (48.4)	150 (19.4)	24.5	4.7×10^{-6}	1.33[c]	1.81[d]
Controls	1860 (48.9)	1940 (51.1)				460 (24.2)	940 (49.4)	500 (26.3)				

Abbreviation: OR, odds ratio.

[a] Data are hypothetical; adapted from Tomlinson et al.[66]

[b] Denotes allelic odds ratio.

[c] Denotes heterozygote odds ratio.

[d] Denotes homozygote odds ratio.

Pearson, T. A. et al. JAMA 2008;299:1335-1344

Ten Basic Questions to Ask About a Genome-wide Association Study Report

- 1. Are the cases defined clearly and reliably so that they can be compared with patients typically seen in clinical practice?
- 2. Are case and control participants demonstrated to be comparable to each other on important characteristics that might also be related to genetic variation and to the disease?
- 3. Was the study of sufficient size to detect modest odds ratios or relative risks (1.3-1.5)?
- 4. Was the genotyping platform of sufficient density to capture a large proportion of the variation in the population studied?
- 5. Were appropriate quality control measures applied to genotyping assays, including visual inspection of cluster plots and replication on an independent genotyping platform?
- 6. Did the study reliably detect associations with previously reported and replicated variants (known positives)?
- 7. Were stringent corrections applied for the many thousands of statistical tests performed in defining the *P value for significant associations?*
- *8. Were the results replicated in independent population samples?*
- *9. Were the replication samples comparable in geographic origin and phenotype definition, and if not, did the differences extend the applicability of the findings?*
- *10. Was evidence provided for a functional role for the gene polymorphism identified?*

Pearson, T. A. et al. JAMA 2008;299:1335-1344

The McDermott Center for Human Growth and Development

Center for Human Genetics

Helen H. Hobbs, M.D.
Howard Hughes Investigator
Director, McDermott Center
Chief, Division of Clinical Genetics, Internal Medicine
Professor of Internal Medicine and Molecular Genetics

Graduate Program:
Genetics and Development

Phone: 214-648-6724
Mailing Address:
5323 Harry Hines Blvd., Dallas, TX 75390-8591
E-mail: Helen.Hobbs@UTSouthwestern.edu
Fax: 214-648-7539

Research Interests:

- Genetic determinants of plasma lipid levels
- LDL metabolism
- Role of ABC transporters in lipid transport

Lab Personnel

Recent Publications:

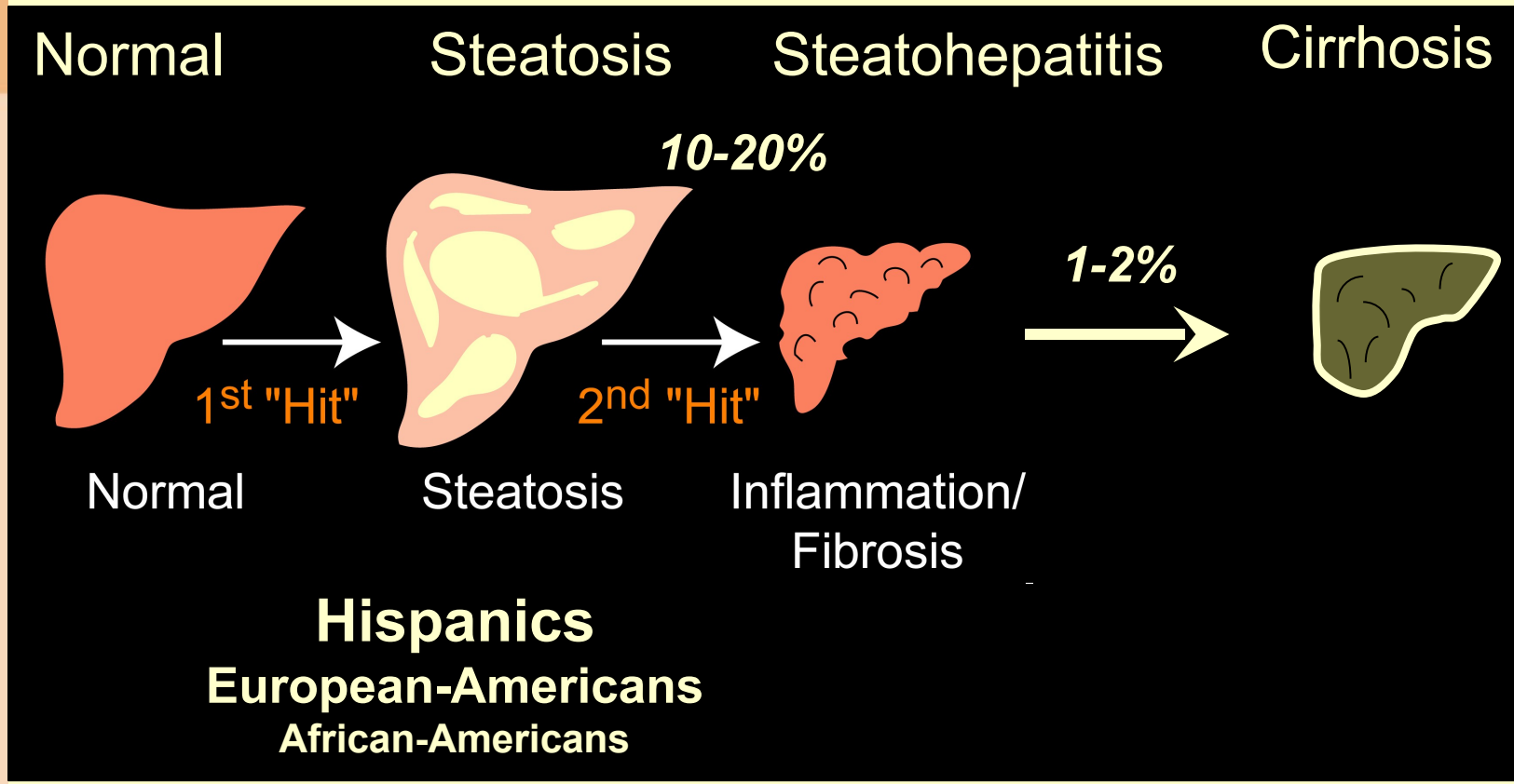
1. Romeo S., Pennacchio L.A., Fu Y., Boerwinkle E., Tybjaerg-Hansen A., Hobbs H.H., Cohen, JC (2007) Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL in Caucasians. Nat. Genet. 39:513-516.
2. McPherson R., Pertsemlidis A., Kavaslar N., Stewart A., Robert R., Cox D.R., Hinds D.A., Pennacch L.A., Tybjaerg-Hansen A., Folsom A.R., Boerwinkle E., Hobbs H.H., Cohen J.C. (2007) A common allele on chromosome 9 associated with coronary heart disease. Science 316:1488-1491.
3. Cohen J.C., Boerwinkle E., Mosley T.H., Hobbs H.H. (2006) Sequence variations in PCSK9, low LD and protection against coronary heart disease. N. Engl. J. Med. 354:1264-1272.
4. Cohen J.C., Kiss R.S., Pertsemlidis A., Marcel Y.L., McPherson R., Hobbs H.H. (2004) Multiple rare alleles contribute to low plasma levels of HDL cholesterol. Science 305:869-872.

For additional publications: Search PubMed

Education:

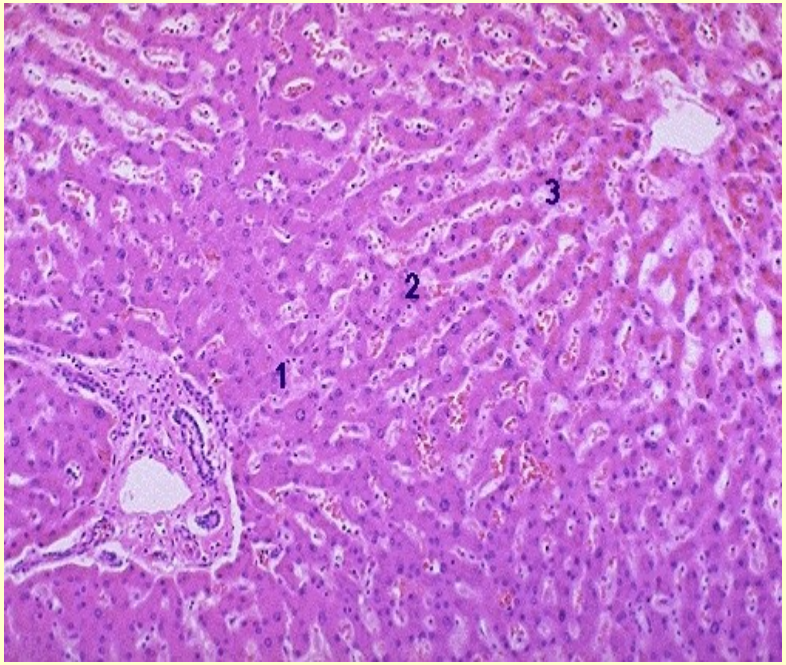
- Stanford University, Palo Alto, CA, B.A., Human Biology, 1974
- Case Western Reserve University School of Medicine, Cleveland , OH , M.D., Medicine 1979
- UT Southwestern Medical Center , Dallas, TX, Postdoctoral Fellow Endocrinology and Molecular Genetics, 1987

Do genetic differences between ethnic groups contribute to differences in fatty liver disease?

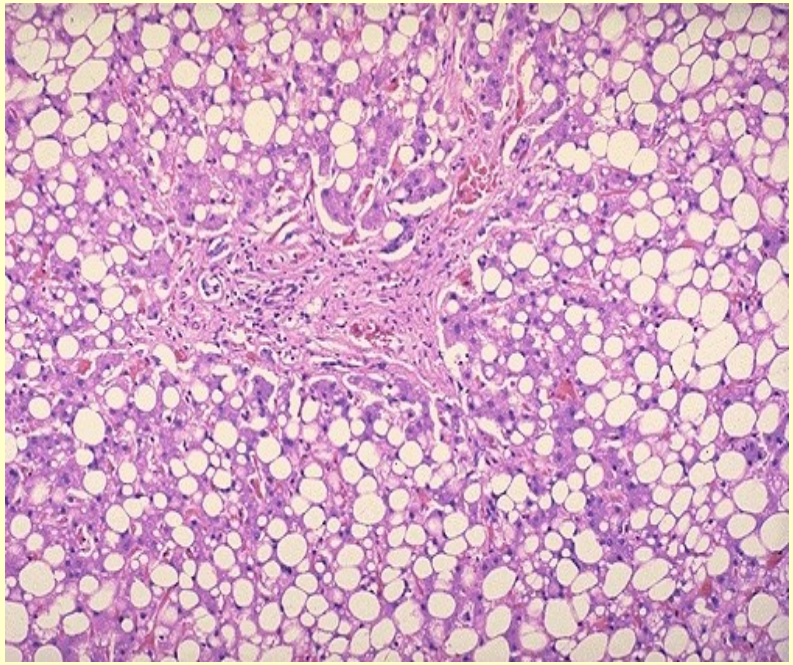


Hepatic Steatosis

Normal



Hepatic Steatosis



- Obesity
- Type 2 diabetes
- Ethanol
- Hepatitis C

Genome-Wide Association Study for Hepatic Triglyceride Content in the Dallas Heart Study

Restricted to nonsynonymous SNPs

Chip-based oligonucleotide hybridization (Perlegen)

Quality filter: n = 12,138 → 9,229

Association with hepatic fat, adjusted for ancestry

(2,270 ancestry informative SNPs)

1,032 African-Americans

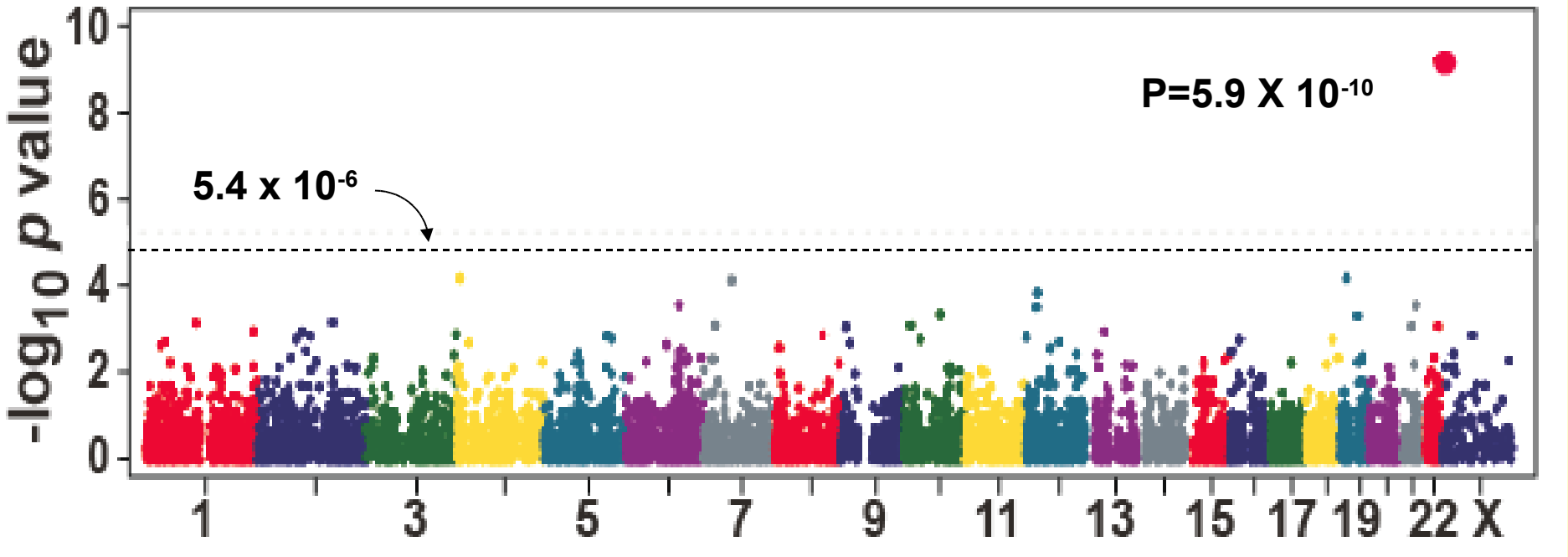
696 European-Americans

383 Hispanics

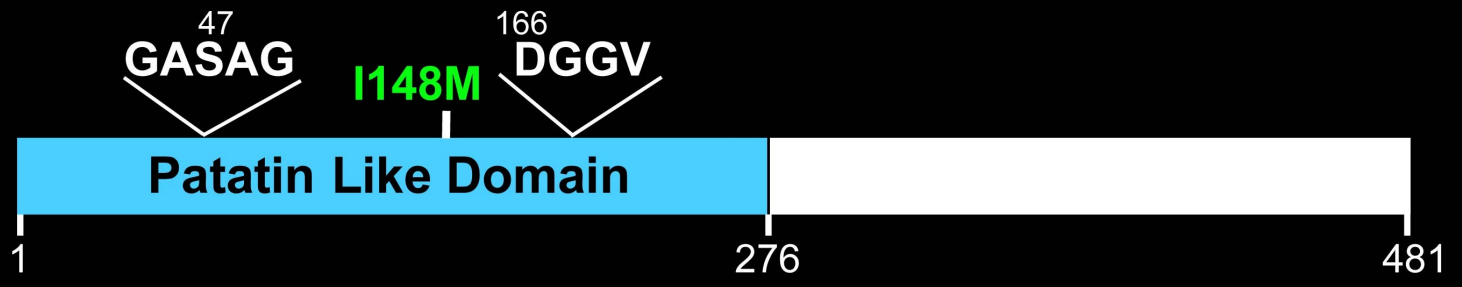
n = 2,111

Romeo, et al.(2008) Genetic Variation in PNPLA3 confers susceptib

Genome-wide Association Study in DHS
Non-synonymous SNPs (n = 9,229)



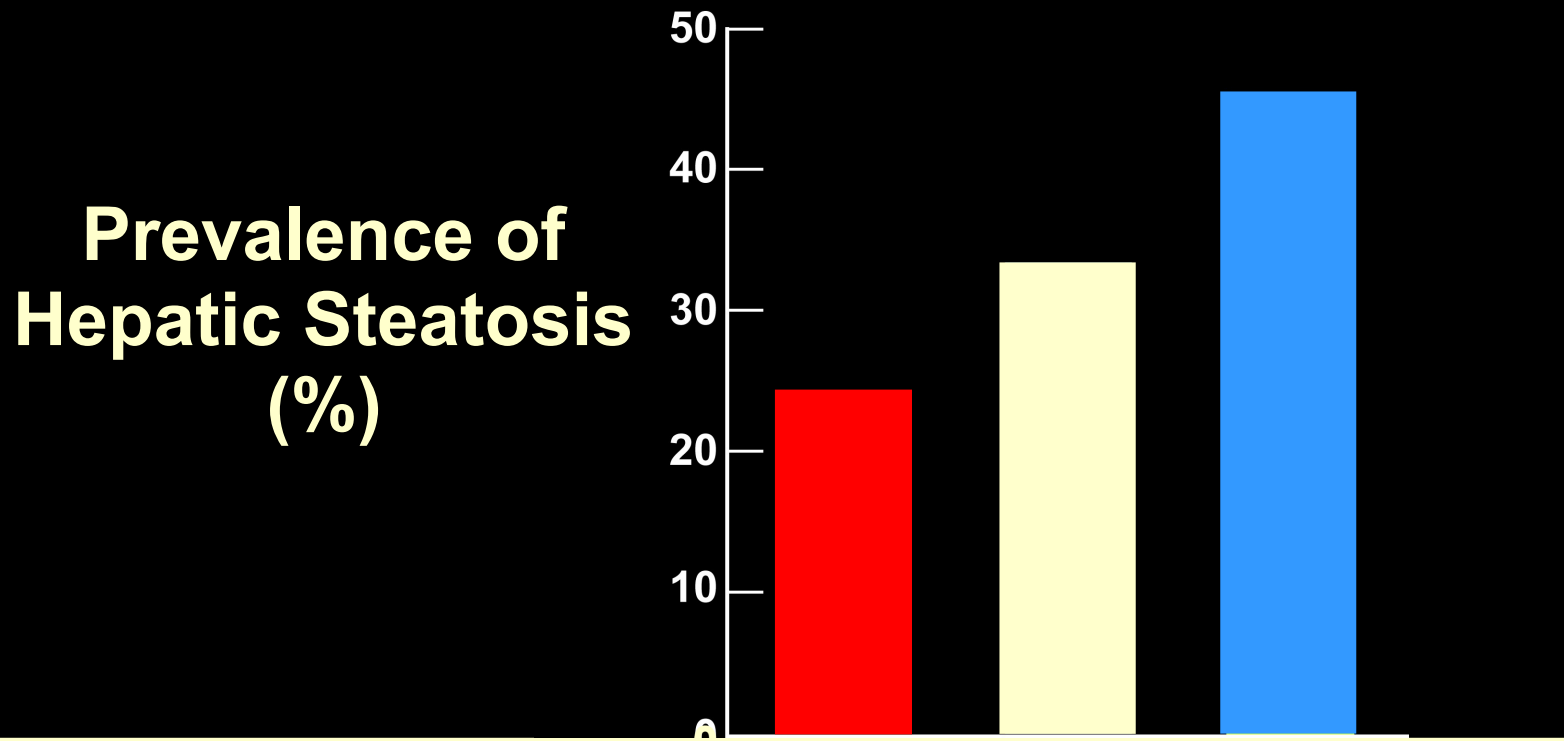
PNPLA3: A Member of the Patatin-like Phospholipase Family



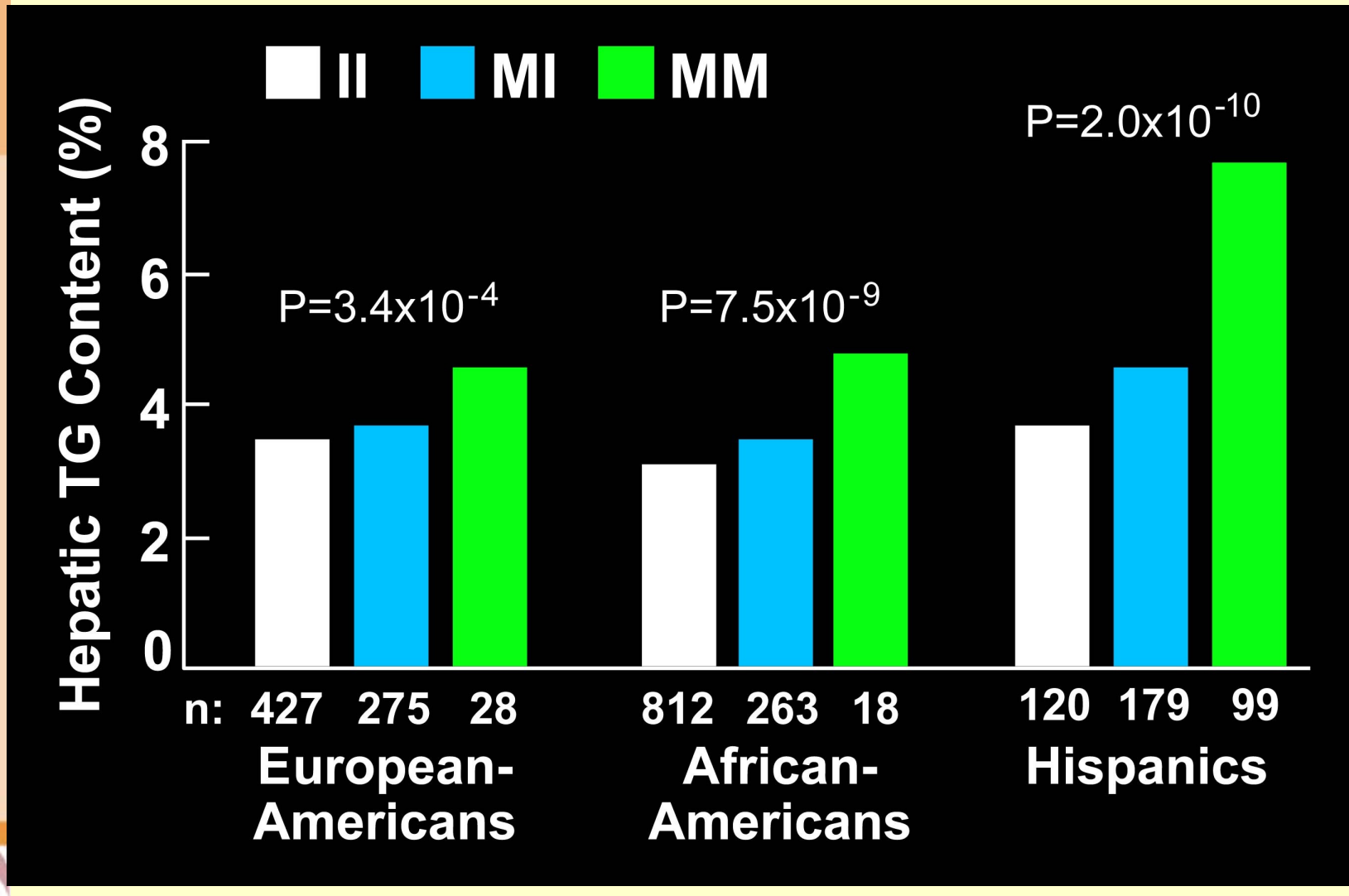
- Resembles patatin: major potato protein
- Nonspecific lipid acyl hydrolase activity (TG>PL)
- Expressed high level in fat & liver
- Increased with feeding (especially carbohydrates)

Ethnic Differences in the Frequency of PNPLA3-I148M

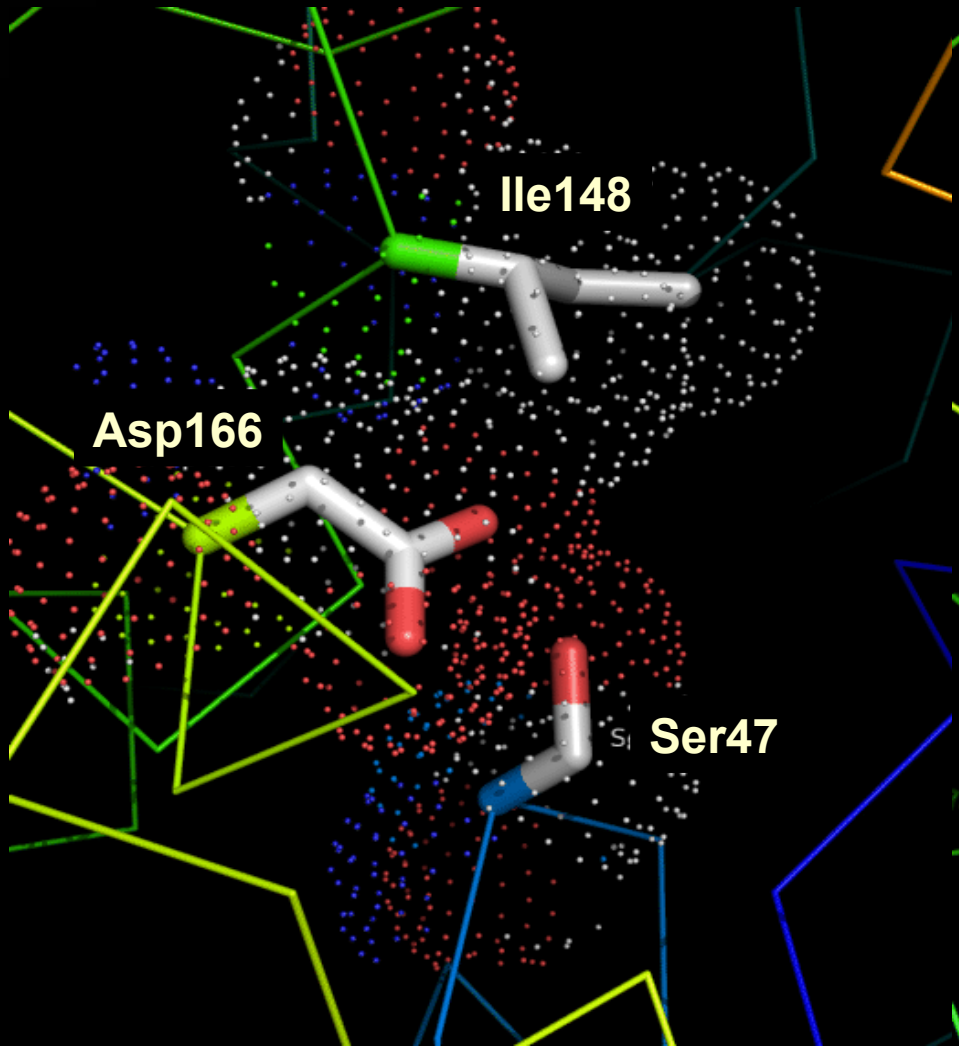
	African-Americans	European-Americans	Hispanics
Minor Allele Frequency	0.17	0.23	0.49

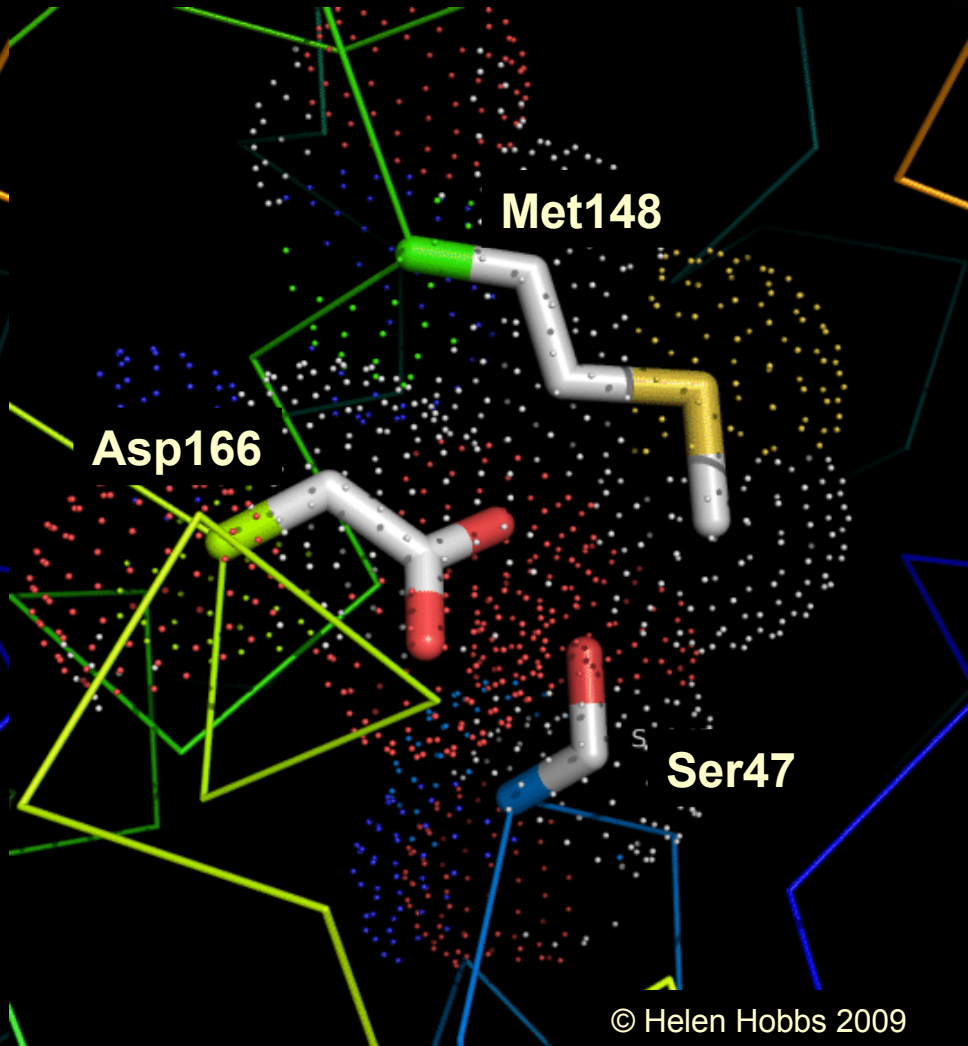


PNPLA3: I148M and Hepatic TG Content

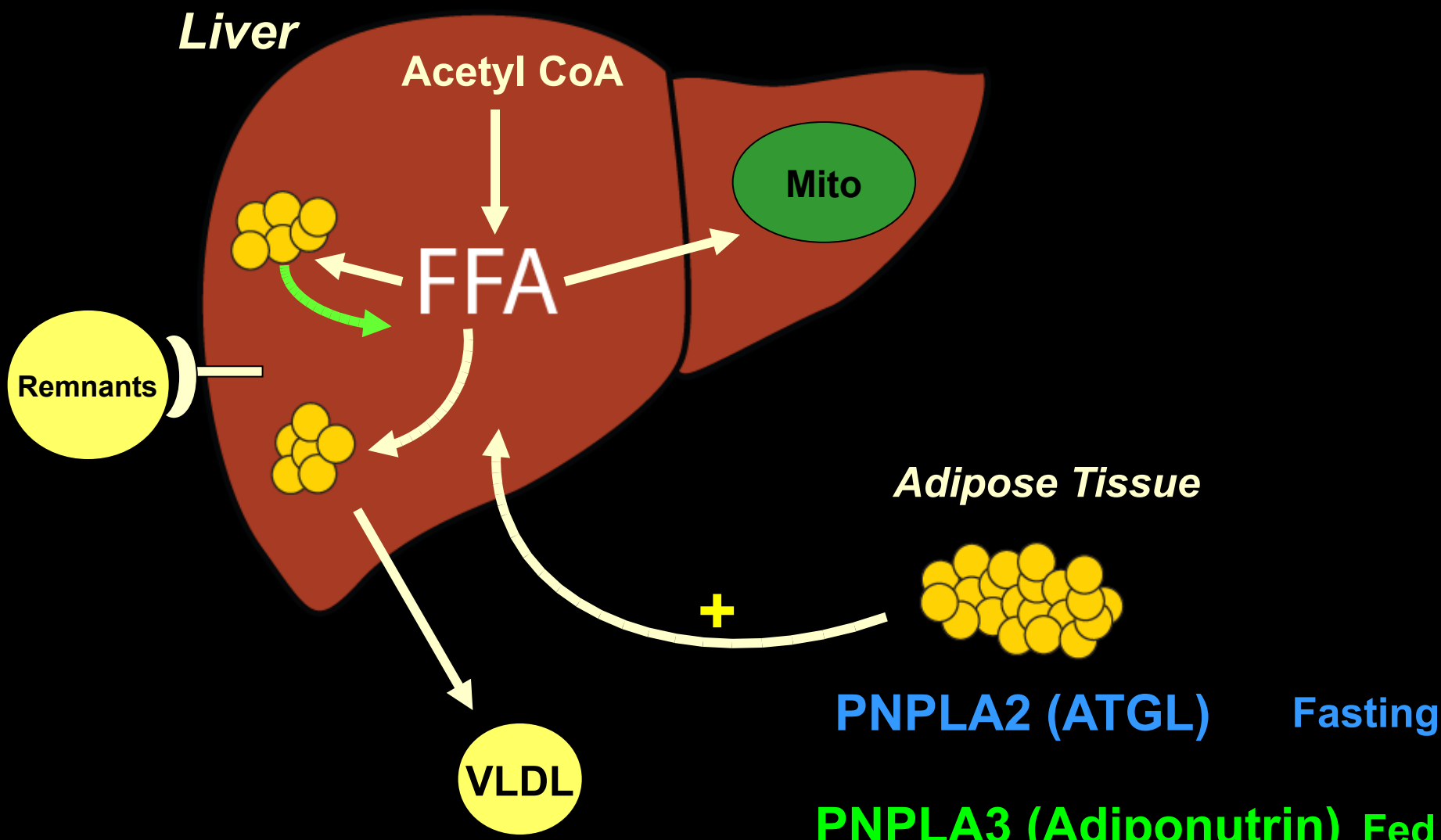


I148M & Catalytic Dyad of PNPLA3

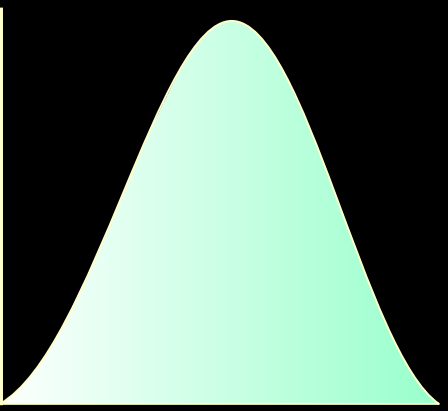


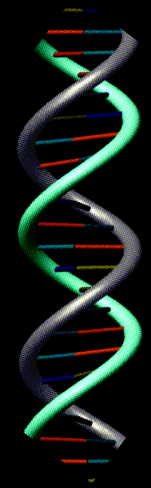


PNPLA3 & Hepatic Triglyceride Metabolism



TRAIT → GENE → PUBLIC HEALTH



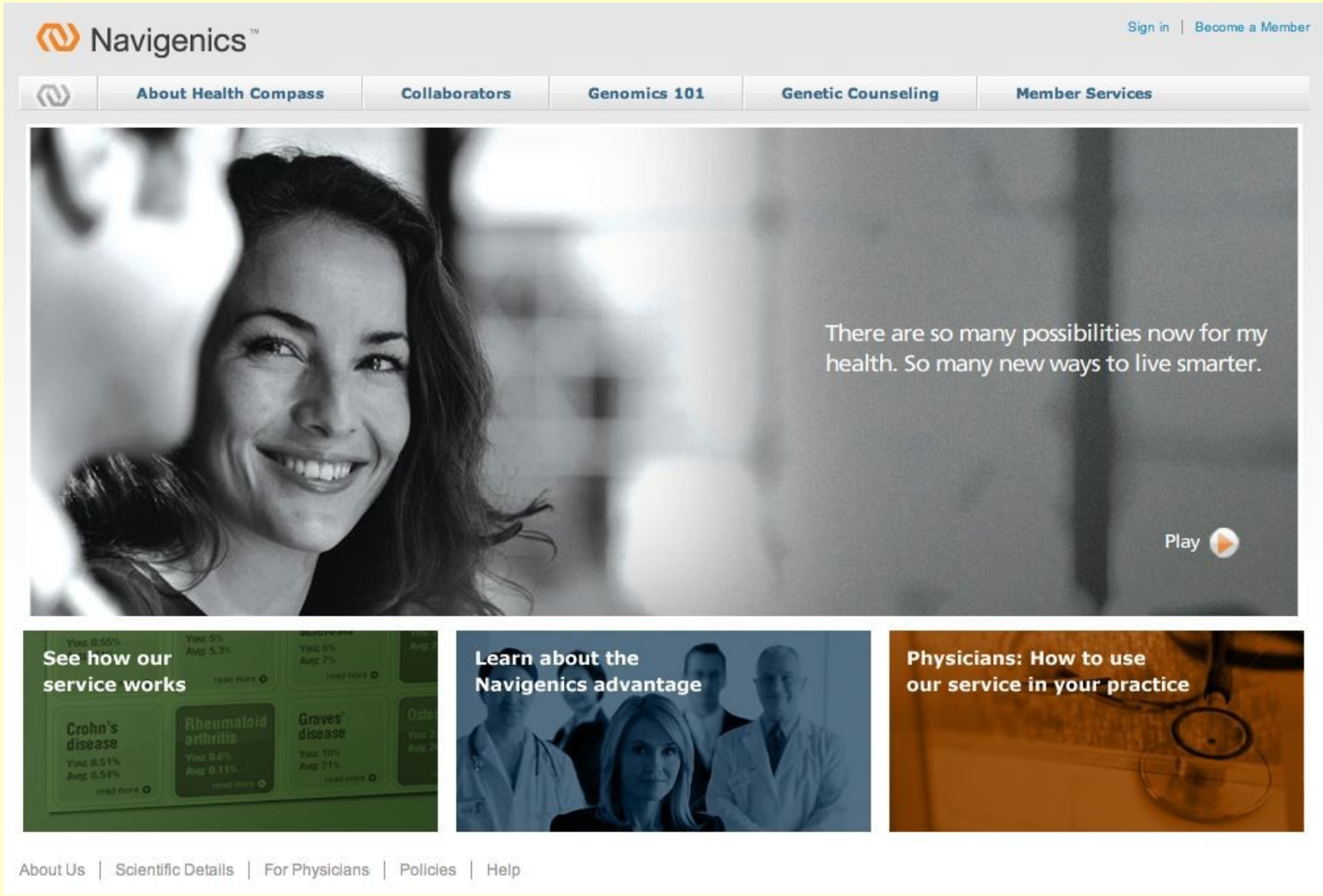


- *PNPLA3*: TG metabolism

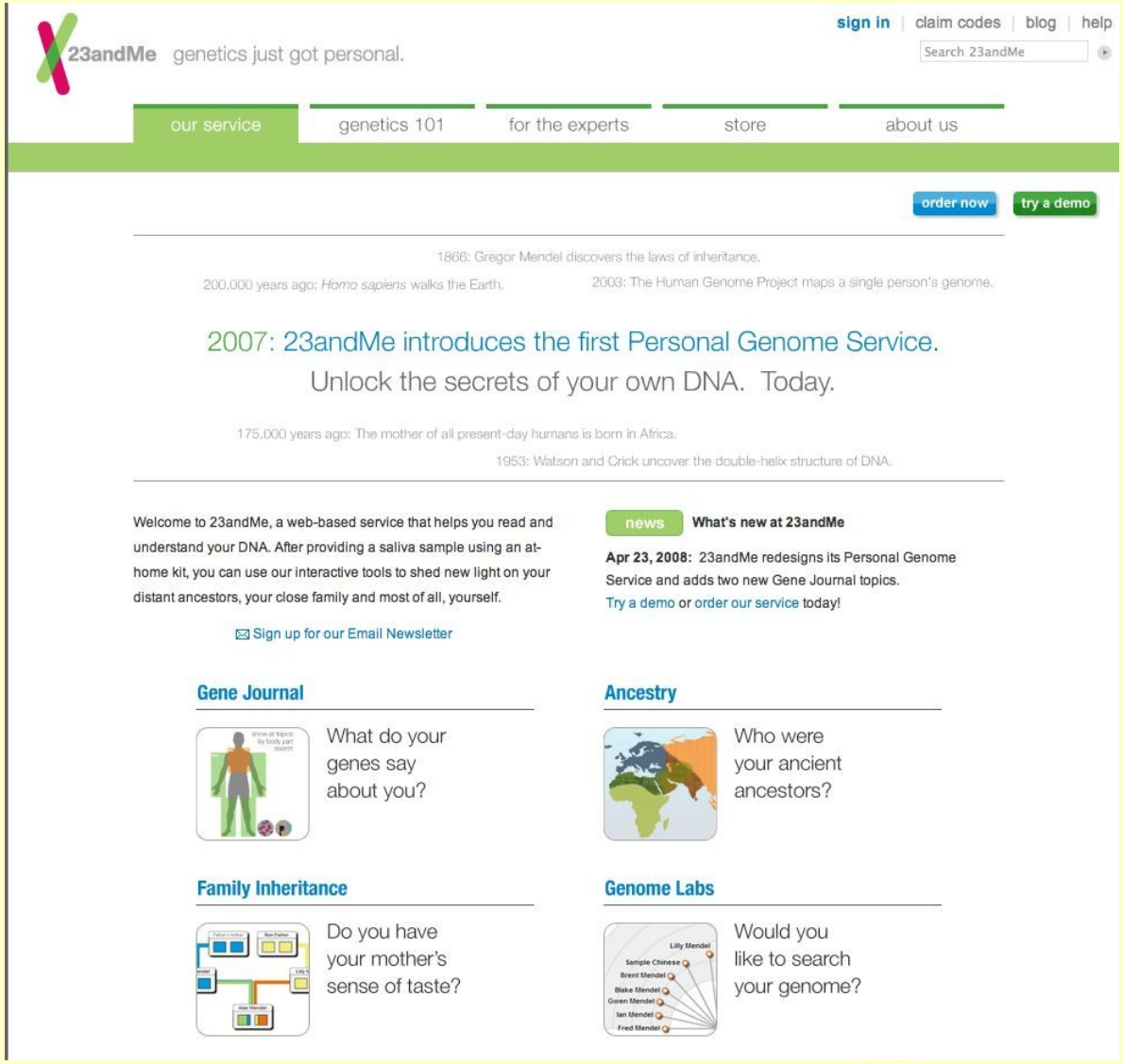
- Therapeutic target
- Prevention strategy
- Risk stratification

DNAdirect: Clinical Genetic Testing D

Navigenics



23andMe



23andMe Kit

23andMe Spittoon

23andMe Sample Tube

23andMe Tube in Envelope

23andMe Fedex Mailer

23andMe Login

23andMe

23andMe Search

Douglas Brutlag | Account | Help

My Home
Inbox

Health
Clinical Reports
Research Reports
Health Labs

Ancestry
Maternal Line
Paternal Line
Relative Finder
Ancestry Painting
Global Similarity
Ancestry Labs

Sharing & Community
Compare Genes
Family Inheritance
23andMe Community

23andWe
My Surveys (23)
Research Initiatives

welcome to you.

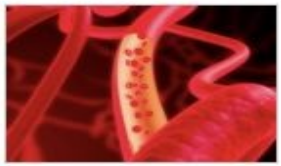
New at 23andMe

You've got relatives.

Introducing Relative Finder, 23andMe's premier tool to help you explore your ancestry.

Connect with living relatives and discover whole new branches of your family tree!

Visit Relative Finder



New articles added to **My Health and Traits**.

- Oral Contraceptives, Hormone Replacement Therapy and Risk of Venous Thromboembolism
- Drinking, Smoking, and Esophageal Cancer
- Response to Hepatitis C Treatment

» view all reports

Three new surveys added to **23andWe**.

- H1N1 Flu (Swine Flu)
- Longevity
- Sports Injuries

» view all surveys

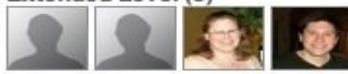
Support disease research through **Research Revolution**.

Join our most ambitious effort yet to let our customers direct and advance human knowledge of how genes and environment affect health. Any 23andMe customer can vote to support research.

» vote!

Currently Shared Genome

Extended Level (5)



» Edit sharing preferences

Help/Contact Us

- See our Frequently Asked Que
  For other questions or feedbac
  please email help@23andme.

Need more kits? Visit the s

Before you view your data...

Knowing your genetic information can have serious and unexpected consequences. Consider the following before yo view your genetic data regarding Breast/Ovarian Cancer:

- **The influence of environmental factors:** The risk for breast and ovarian cancer is only partially determined by genetics. Environmental factors, including but not limited to diet and lifestyle, also play significant roles.
- **This is not the entire genetic picture:** The mutations reported by 23andMe account for only a portion of the entire genetic contribution to breast and ovarian cancer. There are other known mutations, including many in BRCA1 and BRCA2, for which 23andMe does not provide data. If you are concerned about these, you should consult a medical professional about taking specific tests that offer a more complete assessment of these two genes. There are also unidentified genetic factors that affect breast cancer risk.
- **Your ancestry affects your chances of having these mutations:** Though extremely rare in the general population, these mutations are much more common in families with Ashkenazi Jewish ancestry.
- **The mutations described here cannot predict definitively whether you will develop breast or ovarian cancer:** Though having these mutations greatly increases the risk for both diseases, many people who have them will never get the disease. Conversely, lacking these mutations does not substantially reduce your breast or ovarian cancer risk.
- **These mutations are also relevant to men:** Although men are not at risk for ovarian cancer and are at very low risk for breast cancer, BRCA1 and BRCA2 mutations can increase a man's risk for prostate cancer and male breast cancer. Men who carry one of these mutations have a 50% chance of passing it on to their daughters, who would then be at increased risk for breast and ovarian cancer. The mothers and sisters of men who carry one of these mutations also have a 50% chance of being carriers.
- **The wishes of members in your account:** You are about to unlock results for everyone in your account, including the following individuals:

  Douglas Brutlag

  If any of these people do not want to know their genetic status with regard to Breast/Ovarian Cancer, it is your responsibility to ensure this information is not revealed to them or others. You may also request to transfer a profile to a separate account by emailing help@23andme.com.

If, after considering these points, you still wish to view your data, click here.

23andMe Clinical Reports

23andMe Clinical Reports

Douglas Brutlag | Account | Help

Search

My Home
Inbox

Health
- Clinical Reports
- Research Reports
- Health Labs

Ancestry
- Maternal Line
- Paternal Line
- Relative Finder
- Ancestry Painting
- Global Similarity
- Ancestry Labs

Sharing & Community
- Compare Genes
- Family Inheritance
- 23andMe Community

23andWe
- My Surveys (23)
- Research Initiatives

clinical reports

Show results for Douglas Brutlag

Print summary of elevated ri

« Return to Overview | Disease Risks | Carrier Status | Traits | Drug Response | Recently Updated

Elevated Risk

Name	Absolute Risk	Relative Risk	Last Update
No diseases in this category.			

Decreased Risk

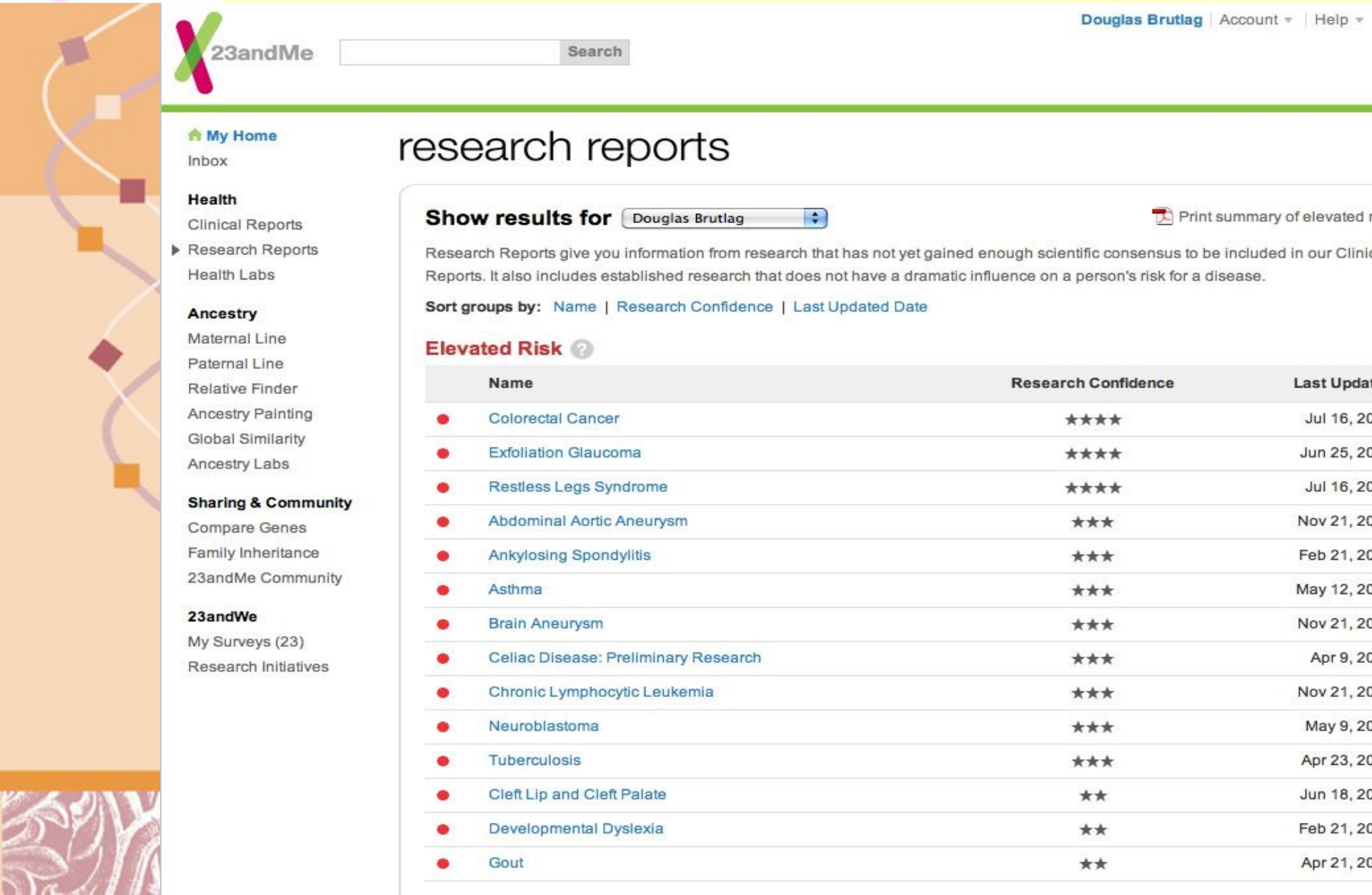
Name	Absolute Risk	Relative Risk	Last Update
Celiac Disease	0.03%	0.26	Jul 7, 20
Age-related Macular Degeneration	2.3%	0.33	May 21, 20
Crohn's Disease	0.3%	0.50	Jul 16, 20
Rheumatoid Arthritis	1.4%	0.59	Aug 6, 20
Type 2 Diabetes	17%	0.70	Feb 2, 20
Type 1 Diabetes	0.8%	0.78	Jul 30, 20

Typical Risk

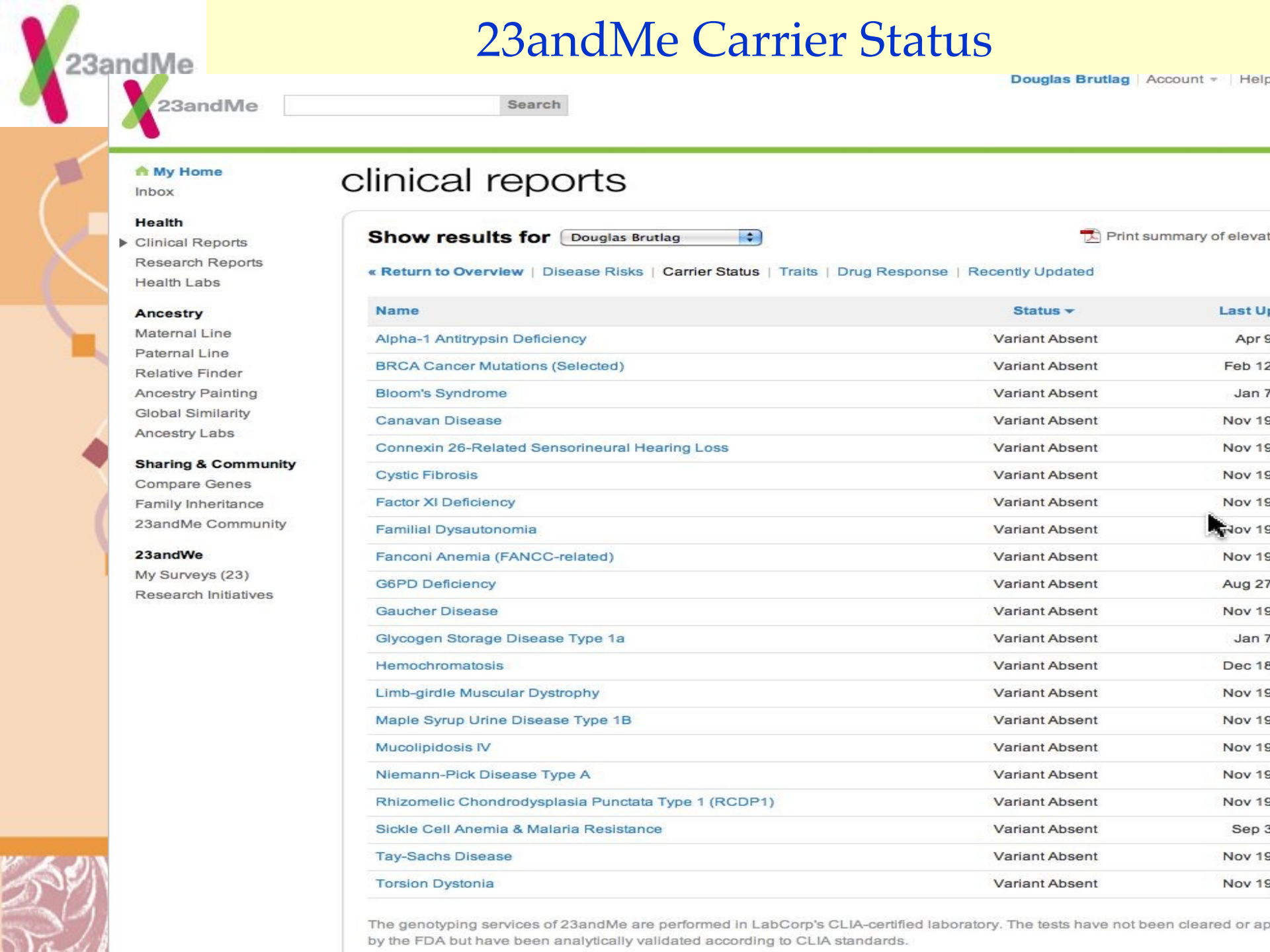
Name	Absolute Risk	Relative Risk	Last Update
Prostate Cancer ♂	18%	1.03	Oct 22, 20
Parkinson's Disease	1.6%	0.98	Sep 29, 20
Venous Thromboembolism	12%	0.96	Jul 30, 20
Psoriasis	9.9%	0.87	Jul 7, 20
Atrial Fibrillation	23%	0.85	Oct 29, 20
Breast Cancer ♀ update	Not Applicable		Feb 18, 20

The genotyping services of 23andMe are performed in LabCorp's CLIA-certified laboratory. The tests have not been cleared or approv by the FDA but have been analytically validated according to CLIA standards.

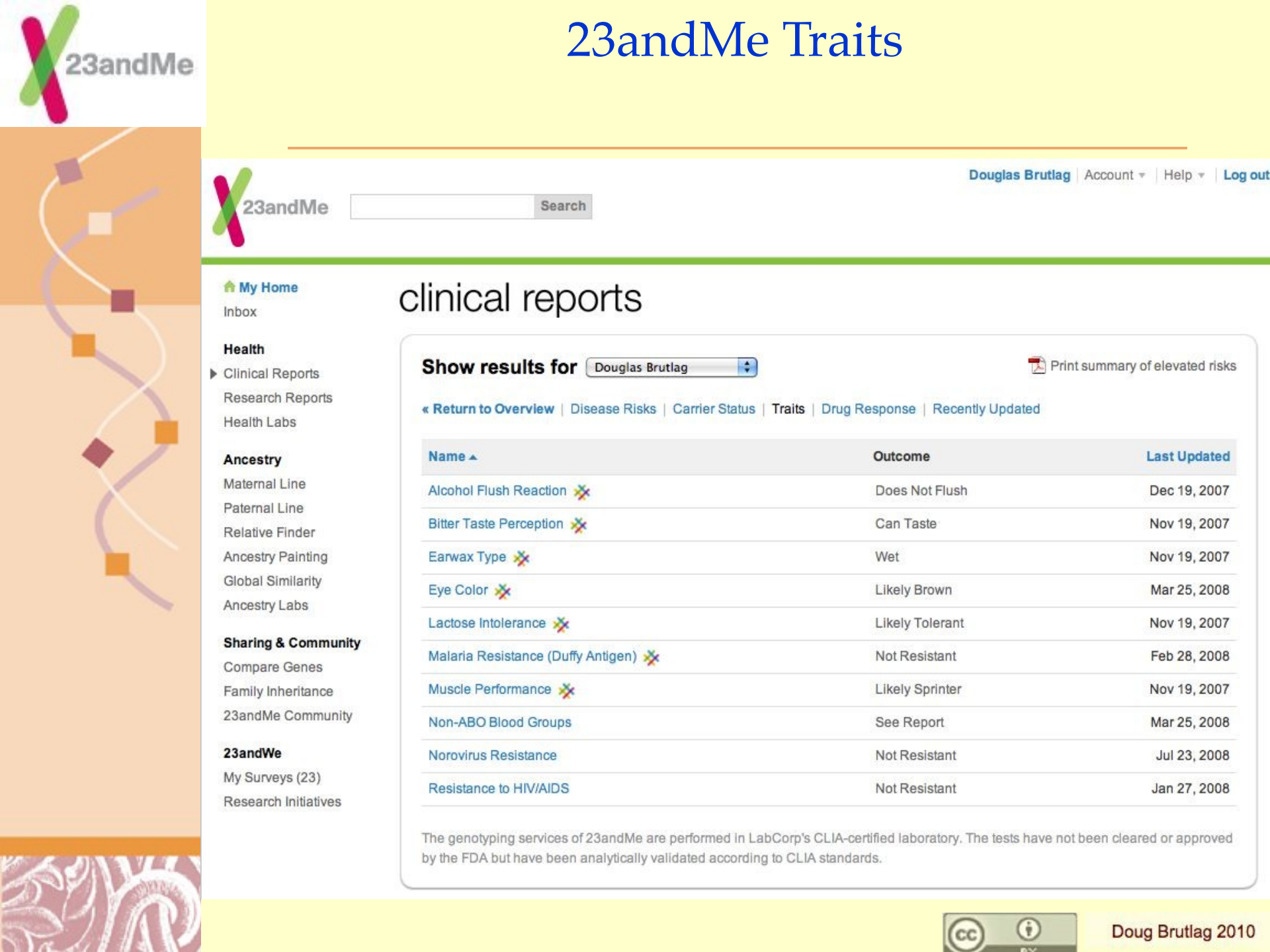
23andMe Research Reports



23andMe Carrier Status



23andMe Traits



23andMe Maternal Inheritance

23andMe genetics just got personal.

brutlag | genetics 101 | blog | help | sign out

me
- My Gene Journal (60)
- Browse Raw Data

family & friends
- Compare Genes
- Family Inheritance

my ancestors
- Maternal Line
- Paternal Line
- Ancestry Painting
- Global Similarity

account
- My Profiles
- Genome Sharing
- Settings
- Help/Contact Us

maternal line

Your mitochondrial DNA determines your maternal haplogroup. What is a haplogroup?

Map | History | Haplogroup Tree

Maternal Haplogroup: U5

Locations of haplogroup U5 circa 500 years ago, before the era of intercontinental travel.

Haplogroup U5 arose among early colonizers of Europe around 40,000 years ago; maternal descendants of those early colonizers persist in the region to this day. After the last Ice Age two subgroups of U5 expanded across Europe and into northern Africa and the Near East. Today, one subgroup, U5b1b, is shared by groups as diverse as the northern African desert-dwelling Berbers and the Scandinavian Arctic-dwelling Saami, also known as the Lapps.

Haplogroup: U5, a subgroup of U

Age: 40,000 years

Region: Europe, Near East, North Africa

Populations: Basques, Saami (Lapps) of northern Scandinavia

Highlight: Though primarily a European haplogroup, U5 was recently found in mitochondrial DNA extracted from the remains of a 6th-century AD Chinese chieftain.

Your Family and Friends

K1a1b1a	Simone Brutlag
U5b2	Douglas Brutlag
L3e	Nigerian Man
D5a2	Chinese Man
D4e2	Japanese Man

23andMe Paternal Inheritance

23andMe genetics just got personal.

me
My Gene Journal (60)
Browse Raw Data

family & friends
Compare Genes
Family Inheritance

my ancestors
Maternal Line
Paternal Line
Ancestry Painting
Global Similarity

account
My Profiles
Genome Sharing
Settings
Help/Contact Us

paternal line

Your Y chromosome DNA determines your paternal haplogroup. What is a haplogroup?

Map | History | Haplogroup Tree

Paternal Haplogroup: E3b1a

Locations of haplogroup E3b1a circa 500 years ago, before the era of intercontinental travel.

E3b is most common in northern Africa and southern Europe. It arose about 17,000 years ago in eastern Africa and spread into the Mediterranean region after the Ice Age. E3b1a, a subgroup of E3b, expanded out of the Near East 8,000 years ago into northern Africa and southern Europe. Today it is one of the most common haplogroups in those regions.

Haplogroup: E3b1a, a subgroup of E3b

Age: 14,000 years

Region: Northern Africa, Southern Europe

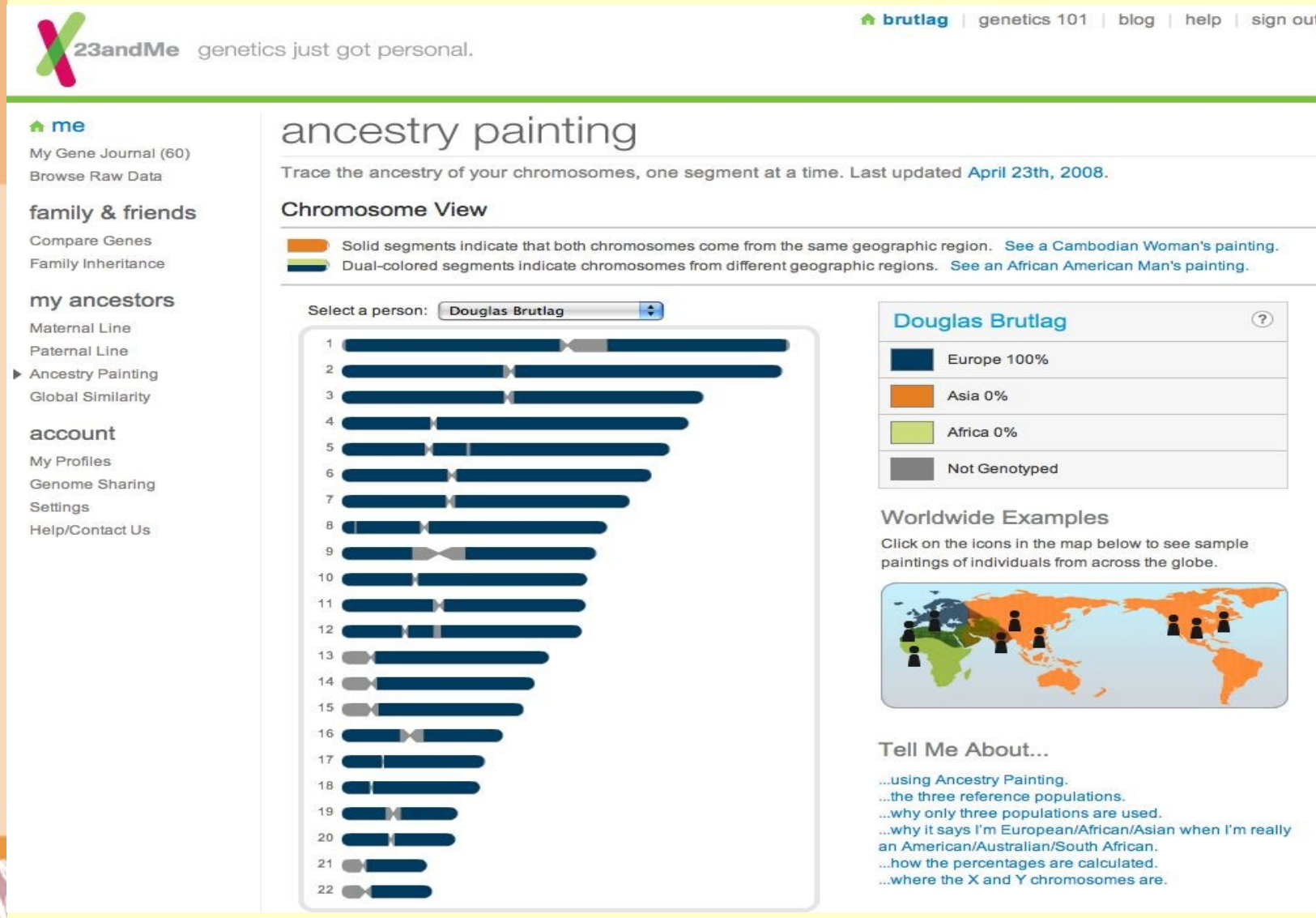
Populations: Berbers, Iberians, Balkans

Highlight: Two different migrations brought E3b1a into Europe.

Your Family and Friends

N	Chinese Man
E3a8a	Nigerian Man
E3b1a	Douglas Brutlag
D2	Japanese Man
N/A	Simone Brutlag

23andMe Ancestry



Genome-Wide Association Study References

How to Use an Article About Genetic Association: A: Background Concepts
John Attia; John P. A. Ioannidis; Ammarin Thakkinstian; et al. JAMA. 2009;301(1):74-81

How to Interpret a Genome-wide Association Study
Thomas A. Pearson; Teri A. Manolio JAMA. 2008;299(11):1335-1344

The Genomics Gold Rush
Eric J. Topol; Sarah S. Murray; Kelly A. Frazer JAMA. 2007;298(2):218-221

The Genome Gets Personal: Almost
W. Gregory Feero; Alan E. Guttmacher; Francis S. Collins JAMA. 2008;299(11):1351-1352

Mapping Genes for NIDDM: Design of the Finland–United States Investigation of NIDDM Genetics(FUSION) Study
Valle et al. DIABETES CARE, VOLUME 21, NUMBER 6, JUNE 1998

Romeo, et al.(2008) Genetic Variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nature Genetics 40, 1461-1465,

Francis O. Walker (2007) Huntington's Disease Review. Lancet 2007; 369: 218–28.

The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared cont

The HapMap and Genome-Wide Association Studies in Diagnosis and Therapy Manolio T. and Collins, F. Annual Review of Medicine (2009) 60

Finding the missing heritability of complex diseases. Manolio TA et al. Nature 2009 461: 747-753