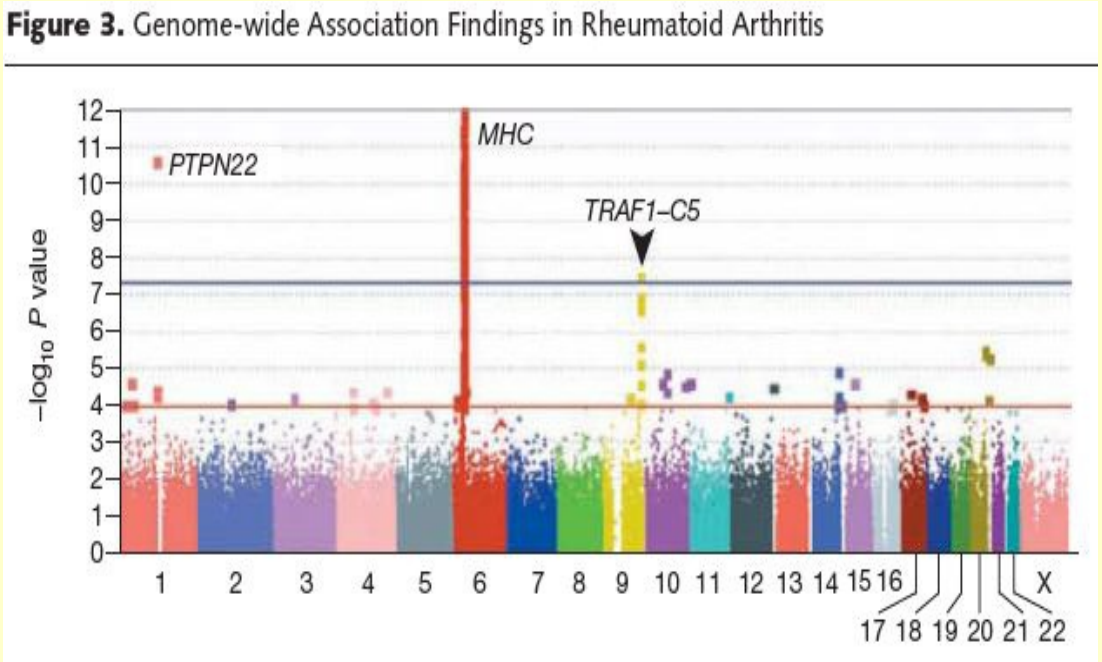
Genomics, Bioinformatics & Medicine

http://biochem158.stanford.edu/

Discovering Variations Associated with Disease

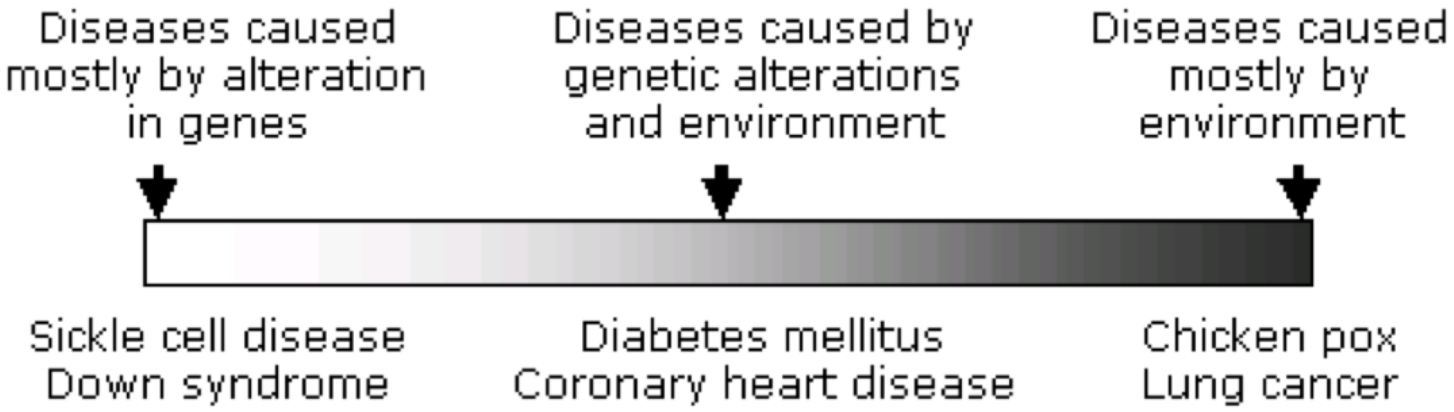
http://biochem158.stanford.edu/gwas.html



Doug Brutlag
Professor Emeritus of Biochemistry & Medicine
Stanford University School of Medicine

© Doug Brutlag 2015

Genetic Penetrance



Genetic diseases, at the left of the spectrum, are categorized as **single gene** or **chromosomal** disorders, depending on the specific genetic cause.

Diseases in the middle of the spectrum — including most common diseases — are **multifactorial**, and result from the interaction or additive effect of genetic and non-genetic factors.

© Doug Brutlag 2015

Genetic Penetrance of Inherited Diseases

- Many inherited diseases are Mendelian and highly penetrant
  - Sickle cell disease
  - Thalassemias
  - Huntington's disease
  - Color blindness
  - Cystic fibrosis
- Most common diseases are complex (multifactorial - caused by multiple genes or multiple pathways as well as multiple environmental factors) and of low penetrance
  - Familial
  - Predisposition to disease
  - Very large environmental and/or behavioral component
    - Type I diabetes and other autoimmune diseases (lupus, rheumatoid arthritis, hyperthyroidism, Crohn's disease, Celiac Sprue, irritable bowel disease etc.)
    - Type 2 diabetes
    - Coronary heart disease (atherosclerosis)
    - Asthma, COPD, pulmonary fibrosis
  - Many complex diseases can be avoided with diet, nutrition, exercise or behavioral modification
  - Many complex diseases can also be monitored by increased vigilance (another behavioral modification)

© Doug Brutlag 2015

Gene Variations Associated with Common Diseases

By comparing the frequencies of gene variations in patients with a disease (cases) and people without the disease (controls) one can often identify susceptibility and protective genes. The are called case-control studies.

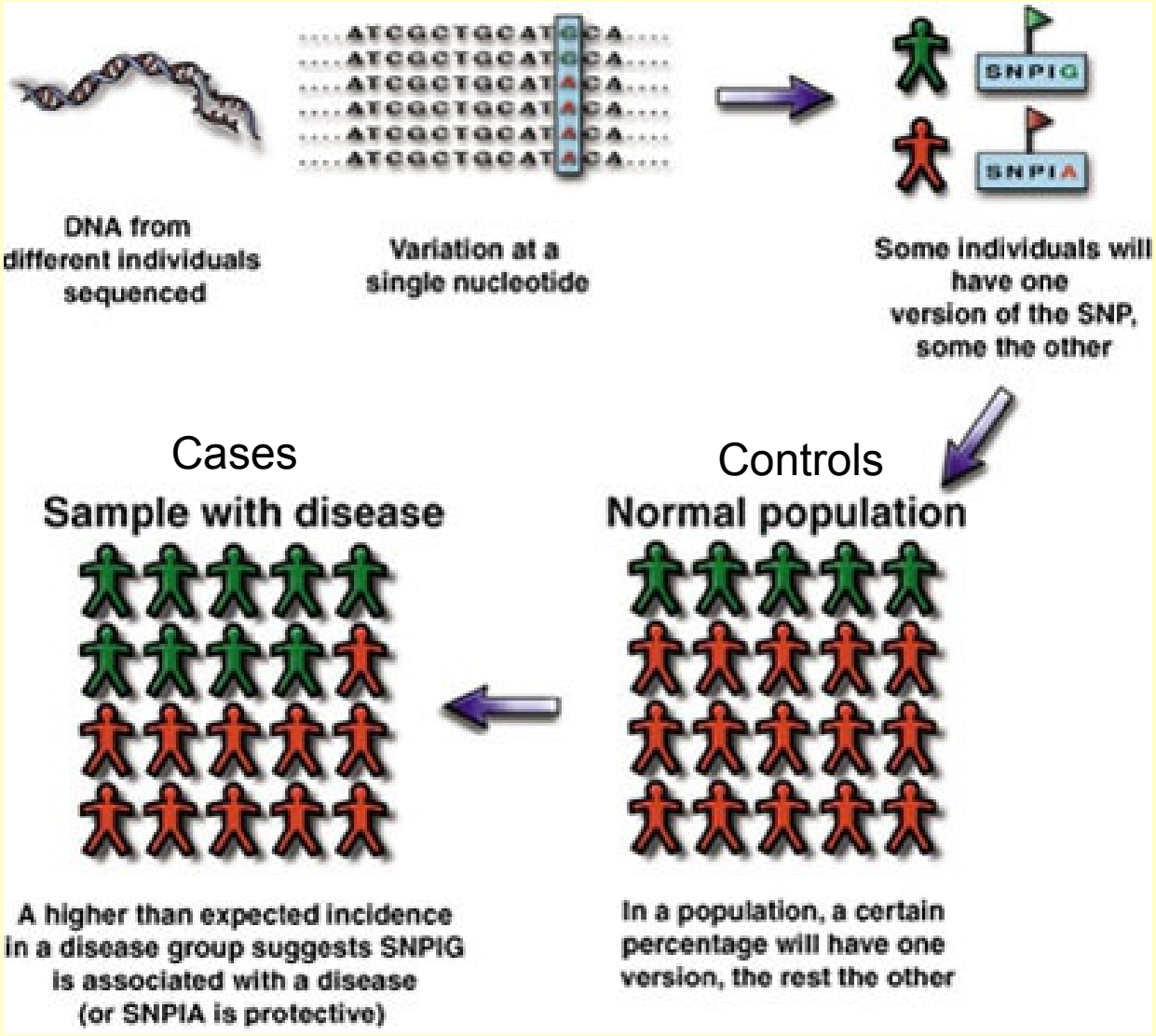
Case-Control studies primarily find correlations of genes with disease. Only rarely do case-control studies discover genes that cause the disease.

Phenotype	Gene	Variant
Peptic ulcer	ABO	O
IDDM*	HLA	DR3,4
Alzheimer dementia	APOE	E4
Deep venous thrombosis*	F5 (R506Q)	Leiden
*Falciparum malaria**	HBB	β^S
AIDS*	CCR5	$\Delta 32$
Colorectal cancer*	APC	3920A
NIDDM*	PPARγ	12A

© Gibson & Muse, A Primer of Genome Science

© Francis Collins, 2009

Using SNPs to Track Predisposition to Disease and other Genetic Traits

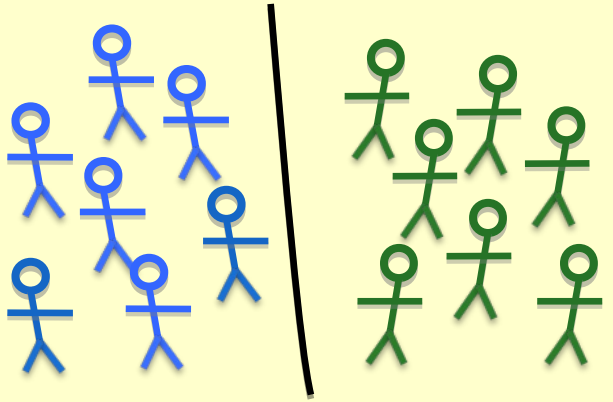


© Gibson & Muse, A Primer of Genome Science

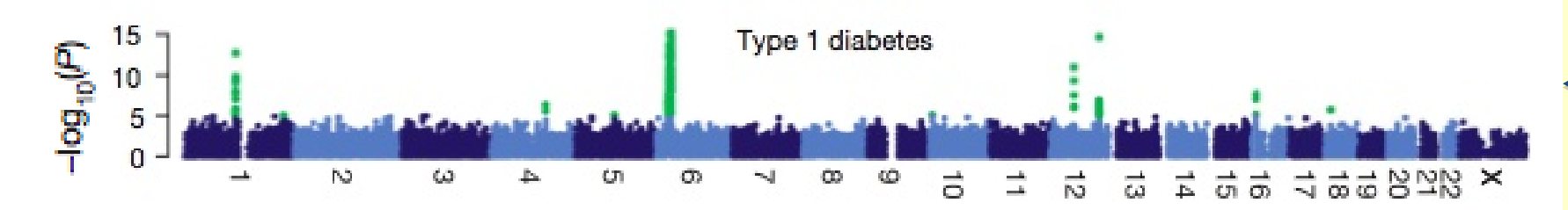
Genome-Wide Association Study: A Brief Primer

Control Population

Disease Population



SNP chip

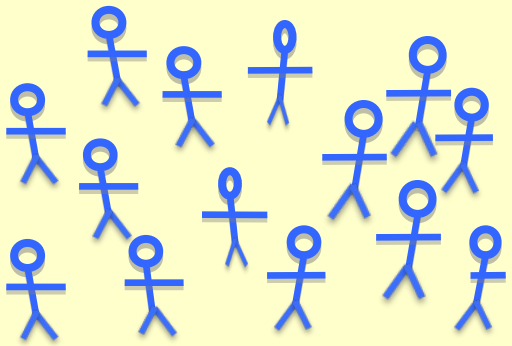


WTCCC, Nature 2007

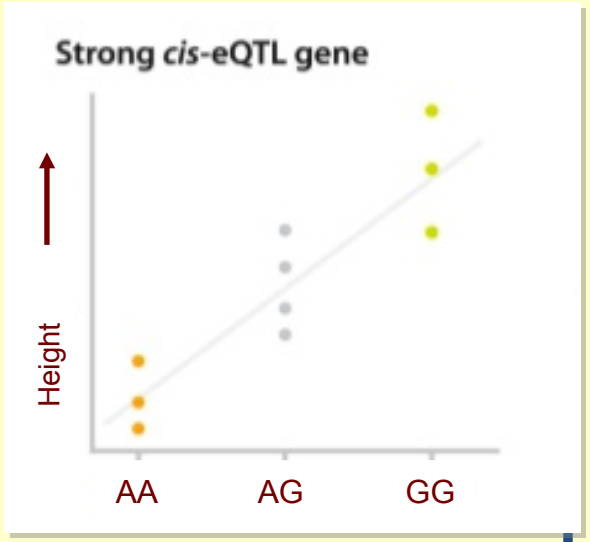
Courtesy of Daniel Newburger

Quantitative Trait Loci Associations

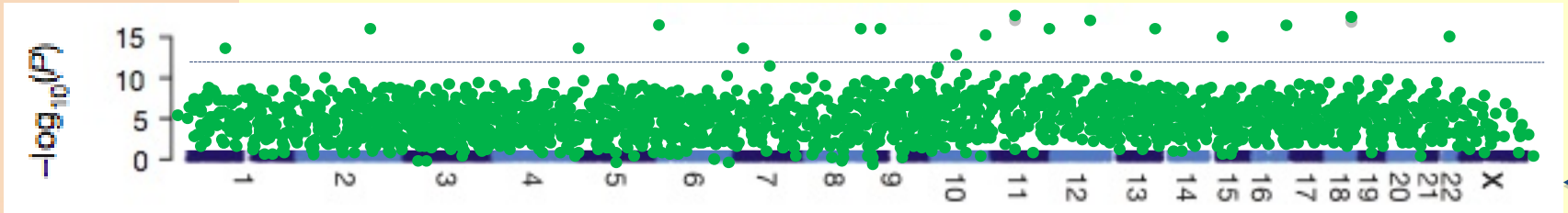
Sample Population



Measure Height



Quantitative Trait Loci (QTLs)



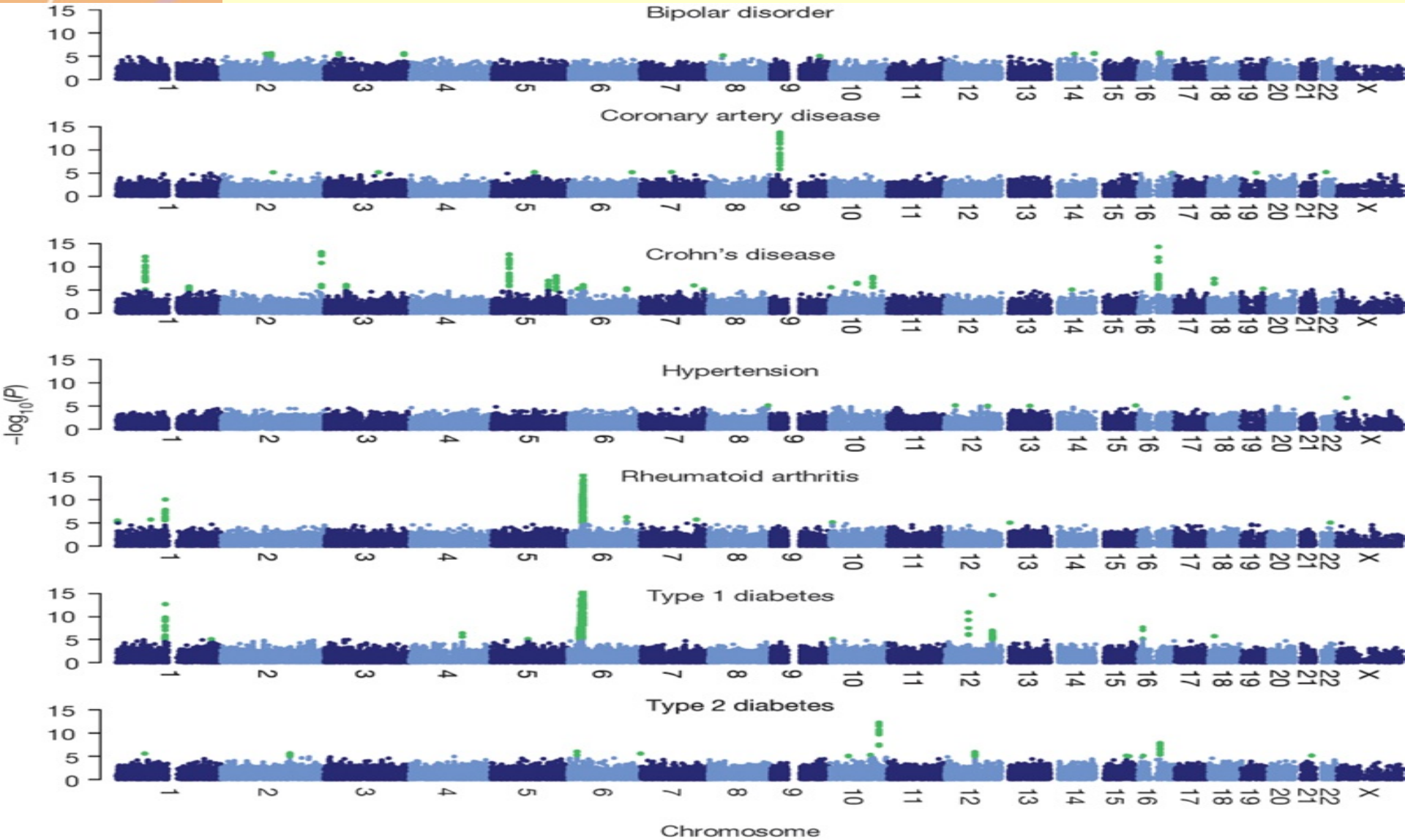
Modified from WTCCC, Nature 2007

Courtesy of Daniel Newburger

The Wellcome Trust Case Control Consortium

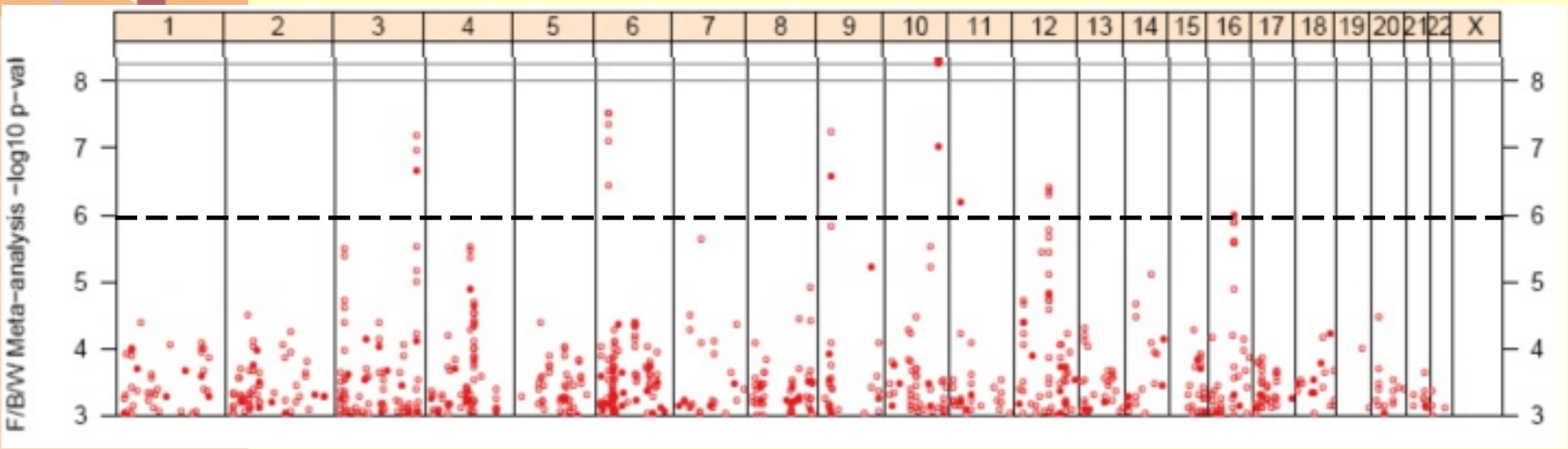
Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls

Nature 447, 661-678 (7 June 2007)



© Francis Collins, 2000

Genome Wide Association of type 2 Diabetes 4549 cases, 5579 controls & 317,503 SNPs

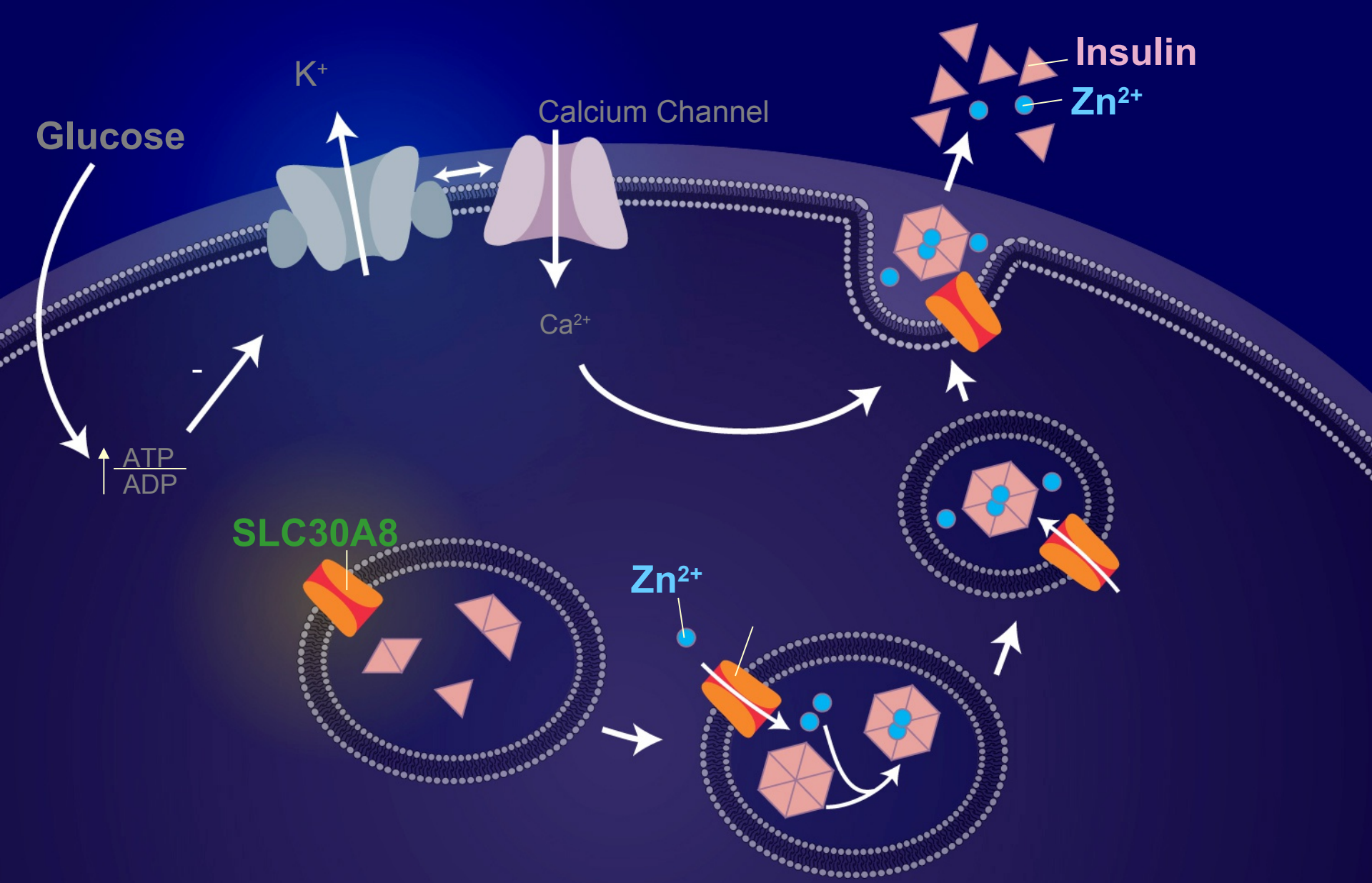


© Francis Collins, 2009

Top 10 Diabetes Genes from Genome-Wide Association Study

Gene	Statistics: Odds Ratio	Statistics: p-value
TCF7L2	1.37	1.0×10^{-48}
IGF2BP2	1.14	8.9×10^{-16}
CDKN2A/B	1.20	7.8×10^{-15}
FTO	1.17	1.3×10^{-12}
CDKAL1	1.12	4.1×10^{-11}
KCNJ11	1.14	6.7×10^{-11}
HHEX	1.13	5.7×10^{-10}
SLC30A8	1.12	5.3×10^{-8}
Chr 11	1.23	4.3×10^{-7}
PPARG	1.14	1.7×10^{-6}

© Francis Collins, 2009



SLC30A8 – A Beta Cell Zinc Transporter

© Francis Collins, 2009

The Great Wave of GWAS Studies

http://www.genome.gov/gwastudies/

Hokusai, K. *The Great Wave*

© Doug Brutlag 2015

Catalog of GWAS Studies

http://www.genome.gov/GWAStudies/

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 › Research Funding › Research Funding Divisions › Division of Genomic Medicine › GWAS Catalog

Share Print

Division of Genomic Medicine

A Catalog of Published Genome-Wide Association Studies

Division Staff ¦ Funding Opportunities ¦ Genomic Medicine Activities ¦ GWAS Catalog ¦ Meetings & Workshops ¦ Potential Sample Collections for Sequencing ¦ Programs ¦ Publications ¦ Trans-NIH Sequencing Inventory

Current uses of and future directions for the Genome-Wide Association Studies Catalog new

On Thursday, July 18th, 2013, the Division of Genomic Medicine held a webinar to highlight current uses and explore priorities and future directions for the GWAS catalog. See archived video and presentations.

Additional information has been added to the HTML catalog columns below. For a description of column headings for the HTML catalog, go to: Catalog Heading Descriptions PDF new

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

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

View the Interactive Diagram new | View the Full Catalog | Download the Catalog | Search the Catalog

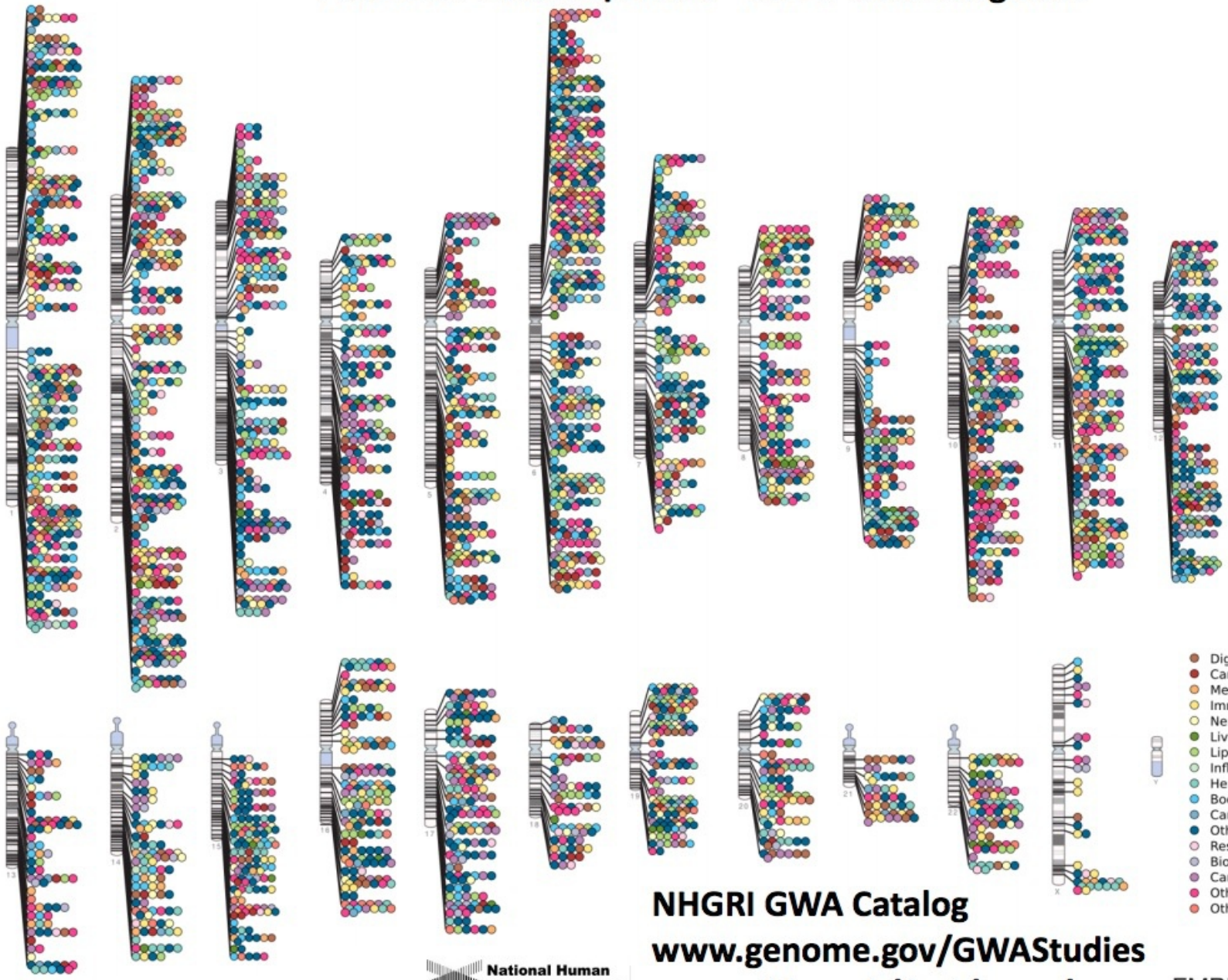
The genome-wide association study (GWAS) publications listed here include only those attempting to assay at least 100,000 single nucleotide polymorphisms (SNPs) in the initial stage. Publications are organized from most to least recent date of publication, indexing from online publication if available. 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 additional 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

© Doug Brutlag 2015

Published Genome-Wide Associations through 12/2013

Published GWA at p≤5X10^{-8} for 17 trait categories



NHGRI GWA Catalog
www.genome.gov/GWAStudies
www.ebi.ac.uk/fgpt/gwas/

National Human Genome Research Institute

EMBL-EBI


Abdominal aortic aneurysm
Acute lymphoblastic leukemia
Adhesion molecules
Adiponectin levels
Age-related macular degeneration
AIDS progression
Alcohol dependence
Alopecia areata
Alzheimer disease
Amyloid A levels
Amyotrophic lateral sclerosis
Angiotensin-converting enzyme activity
Ankylosing spondylitis
Arterial stiffness
Asparagus anosmia
Asthma
Atherosclerosis in HIV
Atrial fibrillation
Attention deficit hyperactivity disorder
Autism
Basal cell cancer
Behcet's disease
Bipolar disorder
Biliary atresia
Bilirubin
Bitter taste response
Birth weight
Bladder cancer
Bleomycin sensitivity
Blond or brown hair
Blood pressure
Blue or green eyes
BMI, waist circumference
Bone density
Breast cancer
C-reactive protein
Calcium levels
Cardiac structure/function
Cardiovascular risk factors
Carnitine levels
Carotenoid/tocopherol levels
Celiac disease
Celiac disease and rheumatoid arthritis
Cerebral atrophy measures
Chronic lymphocytic leukemia
Chronic myeloid leukemia
Cleft lip/palate
Coffee consumption
Cognitive function
Conduct disorder
Colorectal cancer
Corneal thickness
Coronary disease
Creutzfeldt-Jakob disease
Crohn's disease
Crohn's disease and celiac disease
Cutaneous nevi
Cystic fibrosis severity
Dermatitis
DHEA-s levels
Diabetic retinopathy
Dilated cardiomyopathy
Drug-induced liver injury
Drug-induced liver injury (amoxicillin-clavulanate)
Endometrial cancer
Endometriosis
Eosinophil count
Eosinophilic esophagitis
Erectile dysfunction and prostate cancer treatment
Erythrocyte parameters
Esophageal cancer
Essential tremor
Exfoliation glaucoma
Eye color traits
F cell distribution
Fibrinogen levels
Folate pathway vitamins
Follicular lymphoma
Fuch's corneal dystrophy
Freckles and burning
Gallstones
Gastric cancer
Glioma
Glycemic traits
Hair color
Hair morphology
Handedness in dyslexia
HDL cholesterol
Heart failure
Heart rate
Height
Hemostasis parameters
Hepatic steatosis
Hepatitis
Hepatocellular carcinoma
Hirschsprung's disease
HIV-1 control
Hodgkin's lymphoma
Homocysteine levels
Hypospadias
Idiopathic pulmonary fibrosis
IFN-related cytopeni
IgA levels
IgE levels
Inflammatory bowel disease
Insulin-like growth factors
Intracranial aneurysm
Iris color
Iron status markers
Ischemic stroke
Juvenile idiopathic arthritis
Keloid
Kidney stones
LDL cholesterol
Leprosy
Leptin receptor levels
Liver enzymes
Longevity
LP (a) levels
LpPLA(2) activity and mass
Lung cancer
Magnesium levels
Major mood disorders
Malaria
Male pattern baldness
Mammographic density
Matrix metalloproteinase levels
MCP-1
Melanoma
Menarche & menopause
Meningococcal disease
Metabolic syndrome
Migraine
Moyamoya disease
Multiple sclerosis
Myeloproliferative neoplasms
Myopia (pathological)
N-glycan levels
Narcolepsy
Nasopharyngeal cancer
Natriuretic peptide levels
Neuroblastoma
Nicotine dependence
Obesity
Open angle glaucoma
Open personality
Optic disc parameters
Osteoarthritis
Osteoporosis
Otosclerosis
Other metabolic traits
Ovarian cancer
Pancreatic cancer
Pain
Paget's disease
Panic disorder
Parkinson's disease
Periodontitis
Peripheral arterial disease
Personality dimensions
Phosphatidylcholine levels
Phosphorus levels
Photic sneeze
Phytosterol levels
Platelet count
Polycystic ovary syndrome
Primary biliary cirrhosis
Primary sclerosing cholangitis
PR interval
Progranulin levels
Progressive supranuclear palsy
Prostate cancer
Protein levels
PSA levels
Psoriasis
Psoriatic arthritis
Pulmonary funct. COPD
QRS interval
QT interval
Quantitative traits
Recombination rate
Red vs.non-red hair
Refractive error
Renal cell carcinoma
Renal function
Response to antidepressants
Response to antipsychotic therapy
Response to carbamazepine
Response to clopidogrel therapy
Response to hepatitis C treat
Response to interferon beta therapy
Response to metaformin
Response to statin therapy
Restless legs syndrome
Retinal vascular caliber
Rheumatoid arthritis
Ribavirin-induced anemia
Schizophrenia
Serum metabolites
Skin pigmentation
Smoking behavior
Speech perception
Sphingolipid levels
Statin-induced myopathy
Stroke
Sudden cardiac arrest
Suicide attempts
Systemic lupus erythematosus
Systemic sclerosis
T-tau levels
Tau AB1-42 levels
Telomere length
Testicular germ cell tumor
Thyroid cancer
Thyroid volume
Tooth development
Total cholesterol
Triglycerides
Tuberculosis
Type 1 diabetes
Type 2 diabetes
Ulcerative colitis
Urate
Urinary albumin excretion
Urinary metabolites
Uterine fibroids
Venous thromboembolism
Ventricular conduction
Vertical cup-disc ratio
Vitamin B12 levels
Vitamin D insuffiency
Vitiligo
Warfarin dose
Weight
White cell count
White matter hyperintensity
YKL-40 levels

GWAS Catalog
http://www.genome.gov/gwastudies/

Search By:

Journal: Select Journal

First Author:
(last name)

Disease/Trait:
(string search)

Tip: Expand your search by using the OR operator (returns results with either term), or narrow your search using the AND operator (returns results with both terms).

or

5-HTT brain serotonin transporter levels
Abdominal aortic aneurysm
Acenocoumarol maintenance dosage
Acne (severe teenage)
Acne (severe)
Activated partial thromboplastin time
Acute graft versus host disease
Acute lung injury
Acute lymphoblastic leukemia (B-cell precursor)
Acute lymphoblastic leukemia (childhood)
Acute myeloid leukemia
Acute urticaria and angioedema (non-steroidal anti-inflammatory drug-induced)

Tip: Hold Ctrl-key to select multiple entries.

Chromosomal Region:
(e.g., "13q21.31")

Gene:
(e.g., "LRP5")

SNP:
(e.g., "rs20755555")

The SNP data in the catalog has been mapped to dbSNP Build 142 .

GRCh38/hg37.p13

OR greater than:

p-Value threshold:
Enter the exponent. For example, enter "5" for $p<10^{-5}$

Genome-Wide Association Studies

http://gwas.nih.gov/

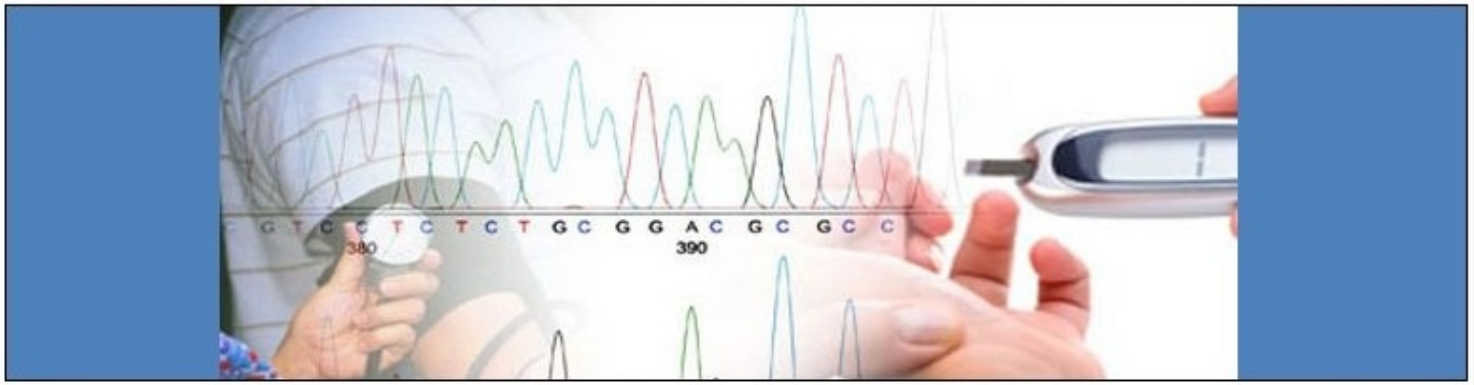
U.S.Department of Health & Human Services www.hhs.gov

www.nih.gov

NIH Genomic Data Sharing (GDS)

Google™ Custom Search Search

- Home
- Policy
- Policy Oversight
- Researchers
- Institutions & IRBs
- Data Repositories
- FAQs
- Related Resources
- Subscribe to the GDS Listserv



Introduction

Genomic research advances our understanding of factors that influence health and disease. In January 2008, NIH established expectations for sharing data obtained through NIH-funded genome-wide association studies (GWAS) with the implementation of the GWAS Policy. GWAS research compares DNA markers across the genome (an individual's complete genetic material) in people with a disease or particular trait to people without the disease or trait.

Information and resources related to the GWAS Policy can be found on this website. Any questions about the Policy can be e-mailed to GWAS@mail.nih.gov.

© Doug Brutlag 2015

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., Pennacchio 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 LDL, 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.

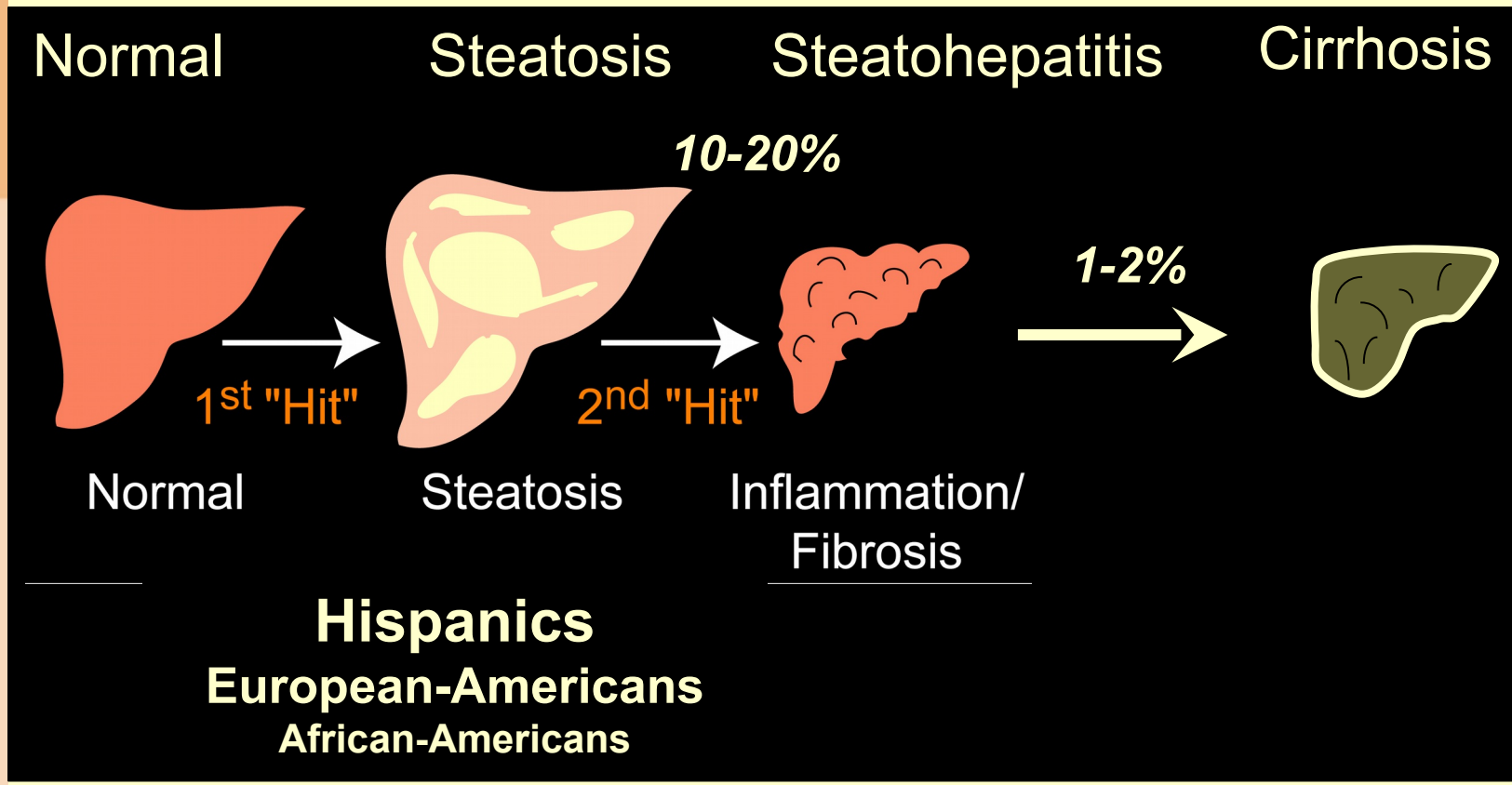
For additional publications: Search PubMed

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

© Helen Hobbs 2009

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



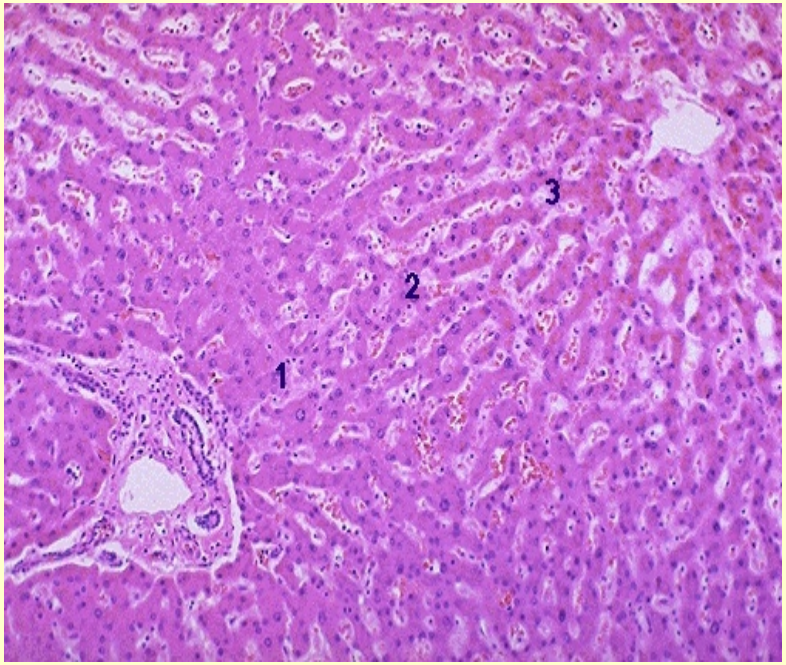
First Hit
- Obesity
- Type 2 diabetes
- Ethanol
- Hepatitis C

Second Hit
- Oxidative Stress
- Lipid Peroxidation
- Anti-virals
- Cytokines

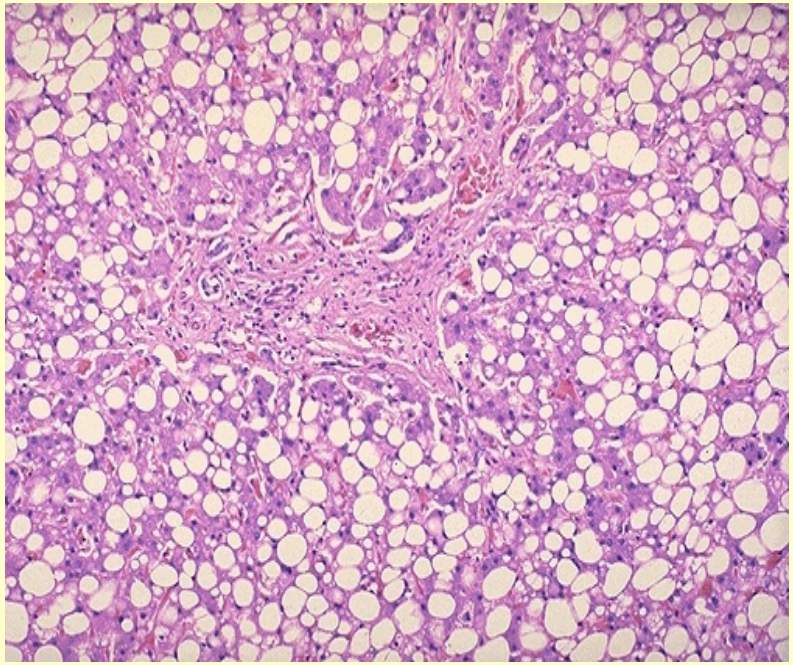
© Helen Hobbs 2009

Hepatic Steatosis

Normal



Hepatic Steatosis



- Obesity
- Type 2 diabetes
- Ethanol
- Hepatitis C

© Helen Hobbs, Nature Genetics V40, pp 1461, 2008

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

- Restricted to non-synonymous SNPs
- Chip-based oligonucleotide hybridization (Perlegen)
- Quality filter: N = 12,138 è 9,229
- Association with hepatic fat, adjusted for ancestry

(2,270 ancestry informative SNPs)

1,032 African-Americans
696 European-Americans
383 Hispanics

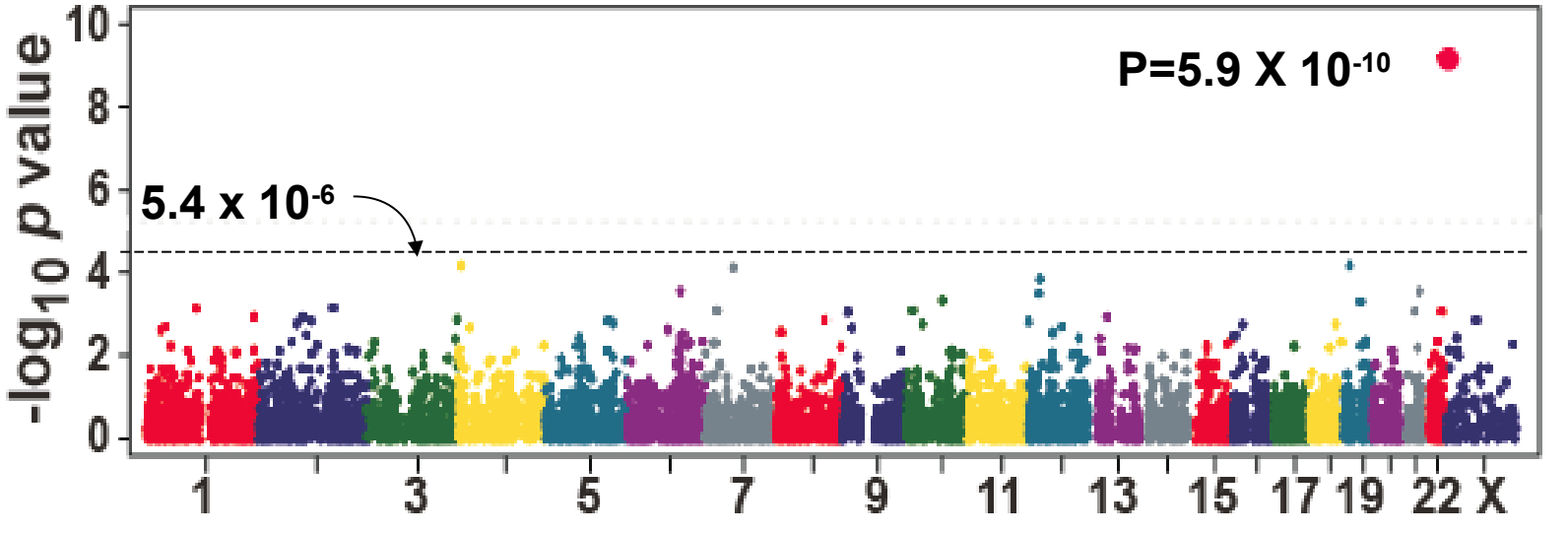
N = 2,111

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

© Helen Hobbs, Nature Genetics V40, pp 1461-1465, 2008

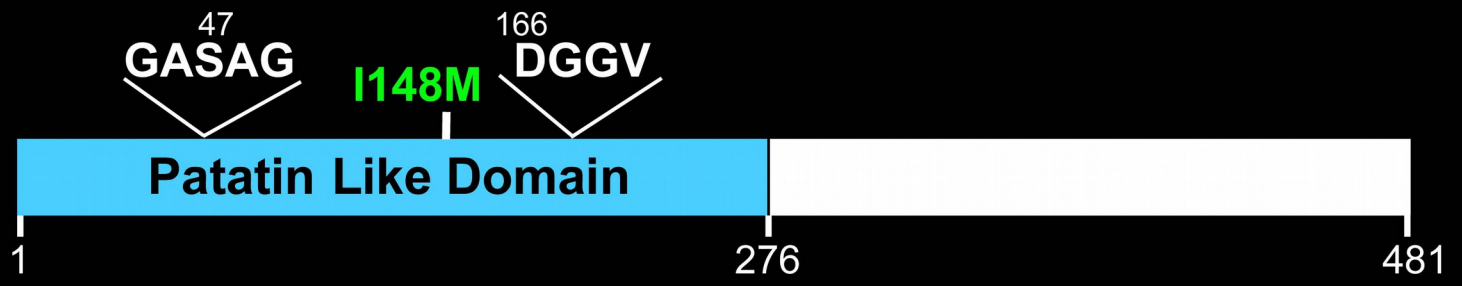
Genome-wide Association Study of Fatty Liver in Dallas Heart Study Cohort

(2,111 patients and 9,299 Non-synonymous SNPs)



© Helen Hobbs, Nature Genetics V40, pp 1461, 2008

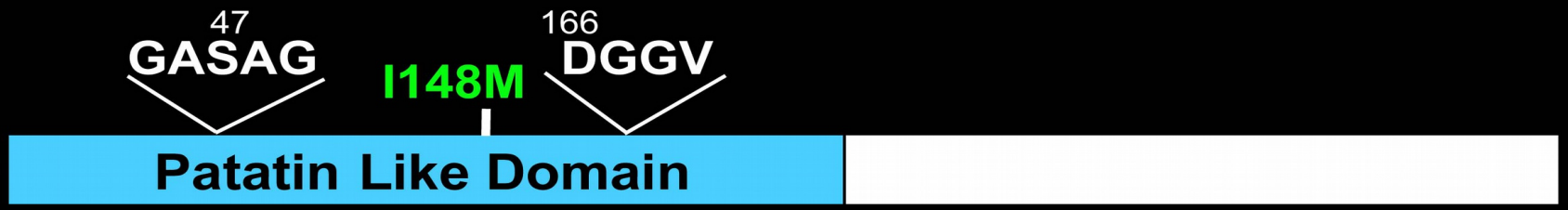
PNPLA3: A Member of the Patatin-like Phospholipase Family

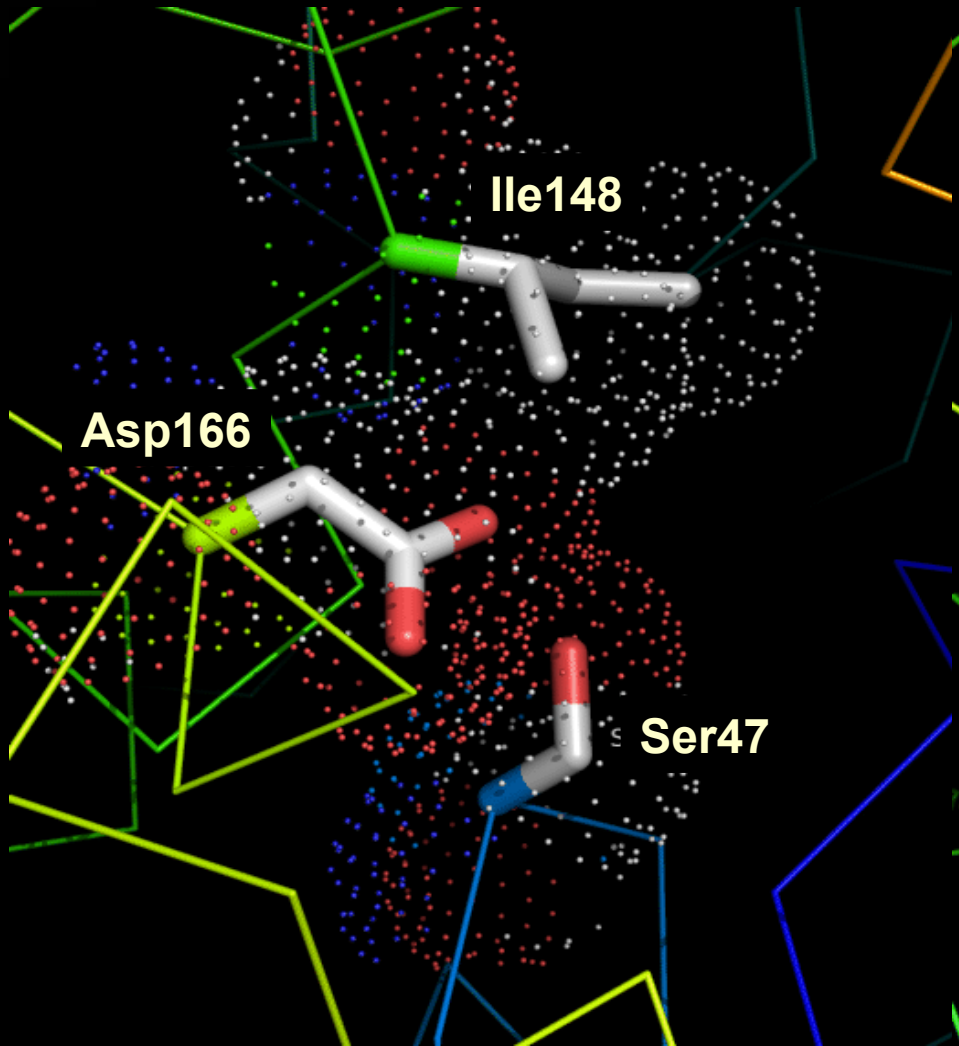


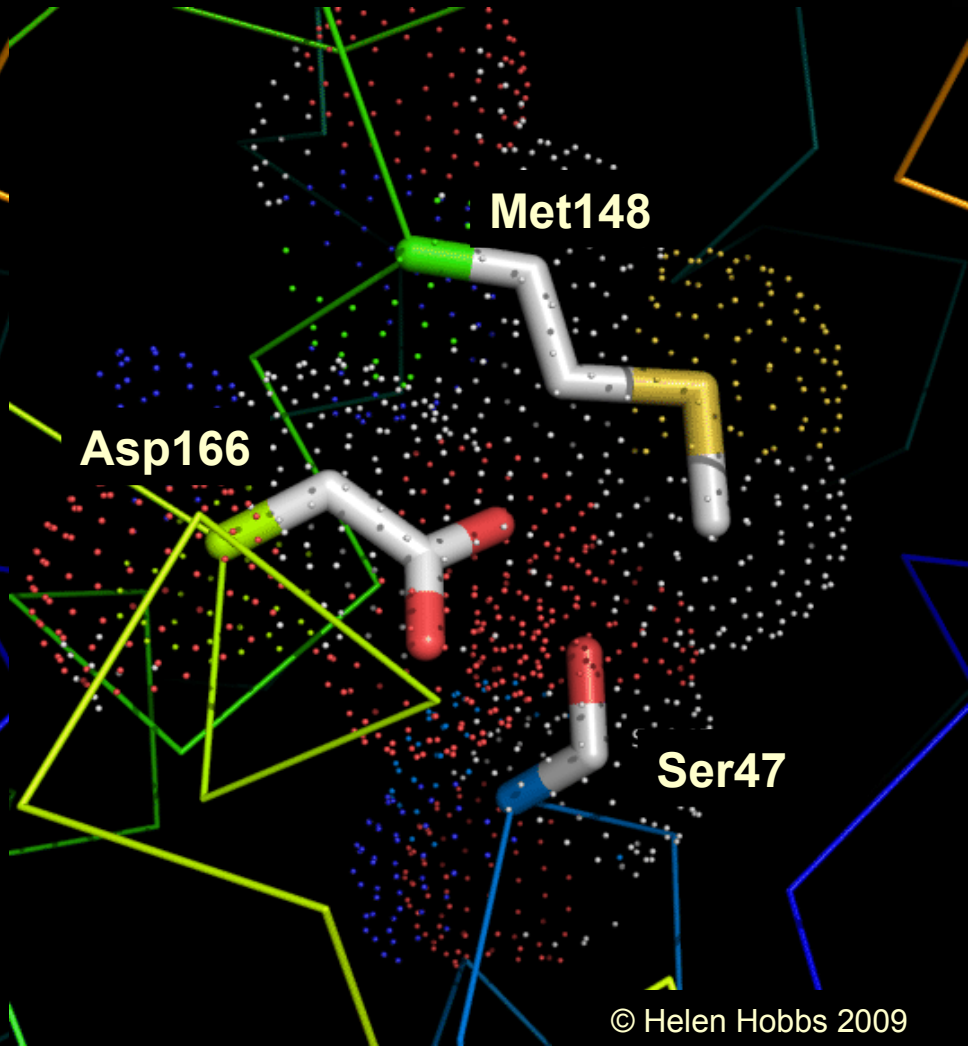
- Resembles patatin: major potato protein
- Nonspecific lipase activity (breaks down fat)
- Expressed high level in fat & liver
- Increased with feeding

© Helen Hobbs, Nature Genetics V40, pp 1461, 2008

I148M & Catalytic Site of PNPLA3

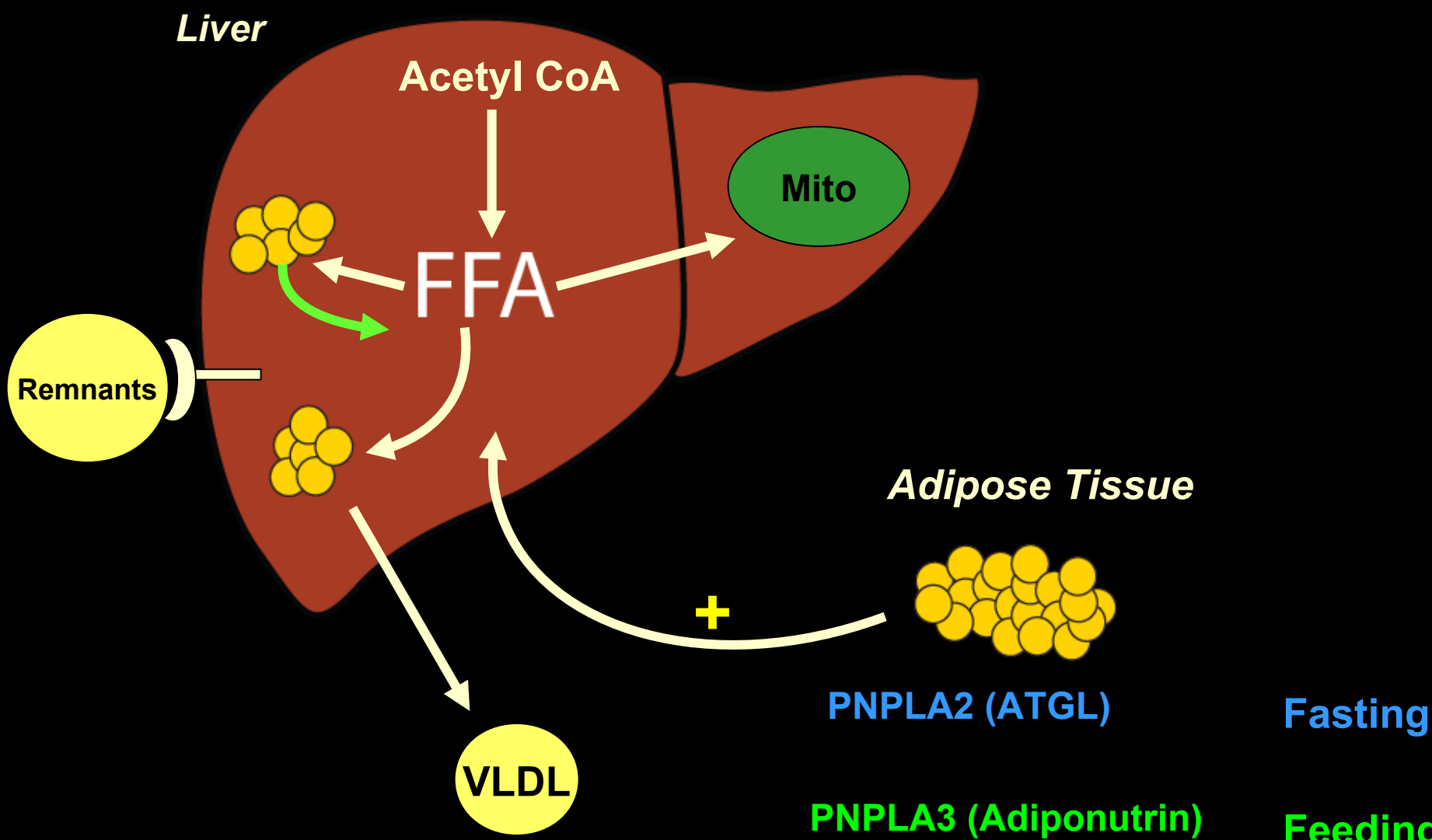




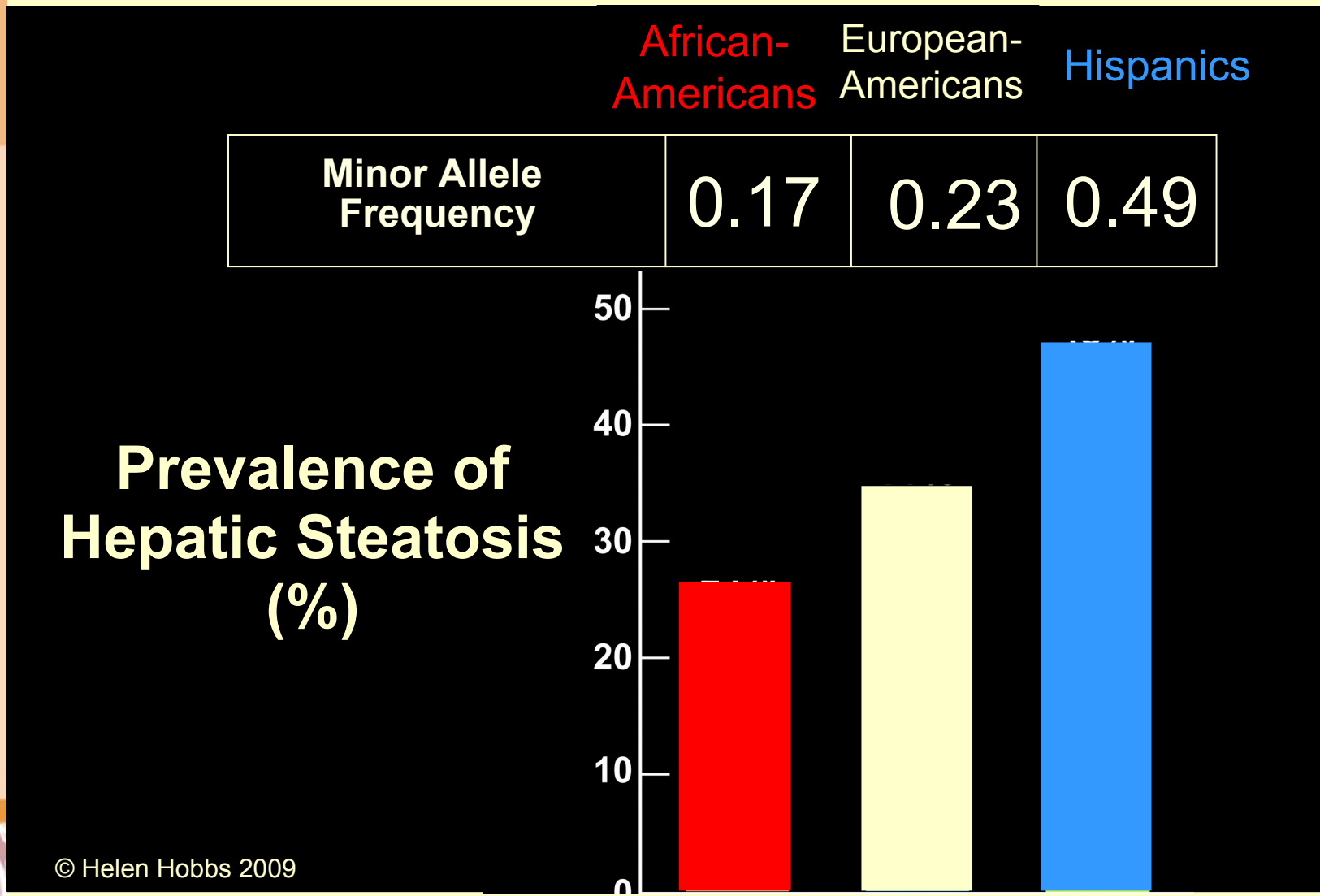


© Helen Hobbs 2009

PNPLA3 & Hepatic Triglyceride Metabolism

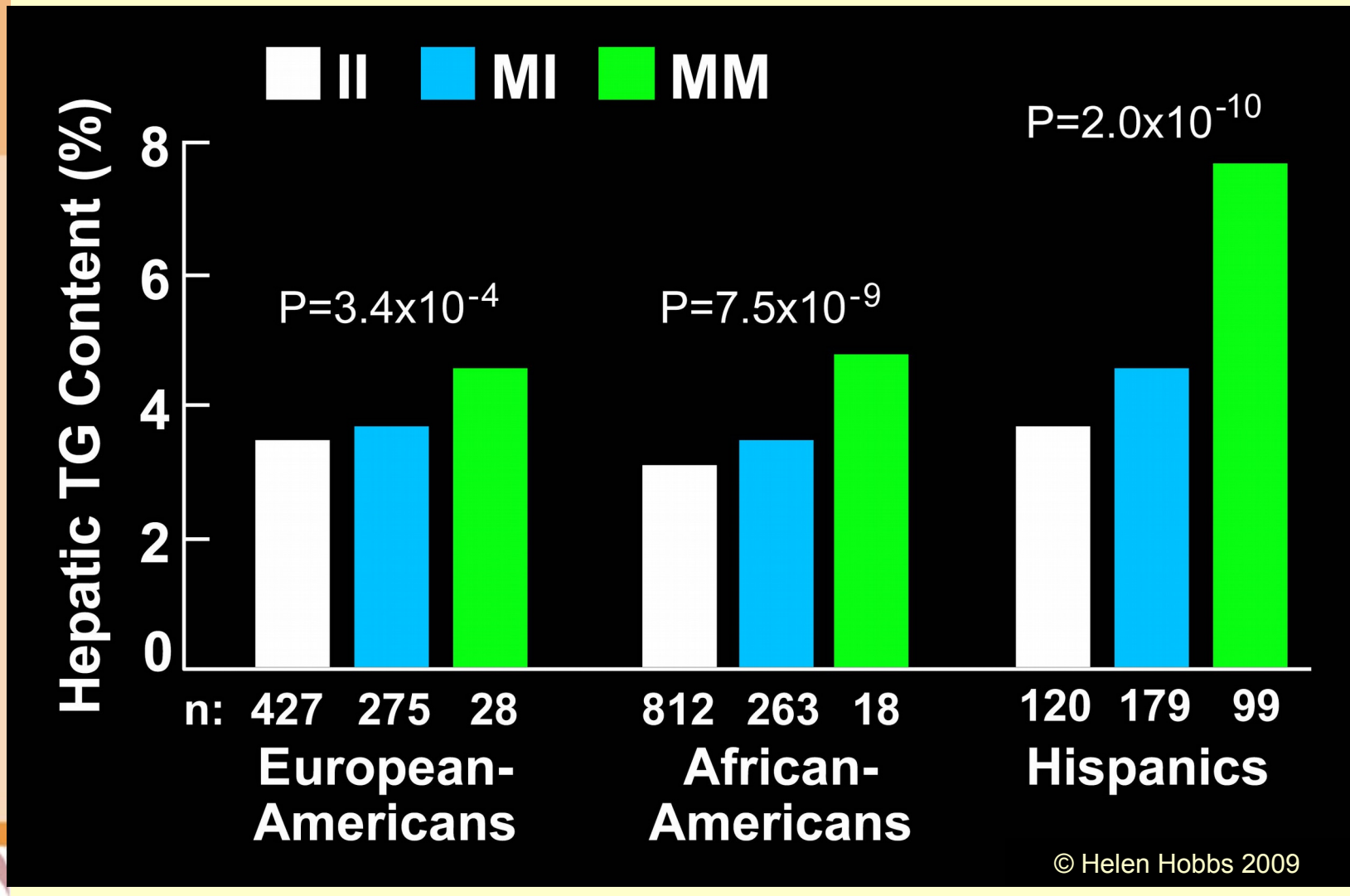


Genetic Contribution to Ethnic Differences in Hepatic Steatosis



© Helen Hobbs, Nature Genetics V40, pp 1461, 2008

PNPLA3: I148M Genotype and Hepatic Triglyceride Content



© Helen Hobbs, Nature Genetics V40, pp 1461, 2008

Genome-Wide Association Project

http://biochem158.stanford.edu/gwas-project.html

Read Thomas A. Pearson; Teri A. Manolio (2008) How to Interpret a Genome-wide Association Study. JAMA, March 19, 2008; 299: 1335 - 1344.

Please search either PubMed or Google Scholar or the GWAS Catalog for a multifactorial disease disease of interest to you AND (GWAS or "Genome wide association study"). To help you with the PubMed search, "Genome-Wide Association Study" is a defined MeSH term so your search can look like this: '"Genome-Wide Association Study"[MaJR] AND Disease-name-or-Disease-MeSH-term

For Google Scholar you will have to do two searches, one with the phrase "Genome-Wide Association Study" AND disease-name and another search for "GWAS AND disease-name".

Read the papers which have performed genome-wide associatio studies on your disease of interest.

Please write a 4-5 page summary of the genome-wide association studies on your disease of interest. Please include the following information in your summary and the implication of each observation:

1) The URL or UID of the papers you read.

2) The genes or SNPs that are most highly correlated with the disease.

3) The odds ratio and heritability of each SNP correlation.

4) Have the association studies been repeated in different laboratories, or different populations or subpopulations?

5) Have causal mutations been detected or suggested from any of the data?

6) Also please report if knowledge of those SNPs or genes sheds any light on the molecular basis for the disease.

© Doug Brutlag 2015

Experimental 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

Examples of Multistage Designs in Genome-wide Association Studies

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

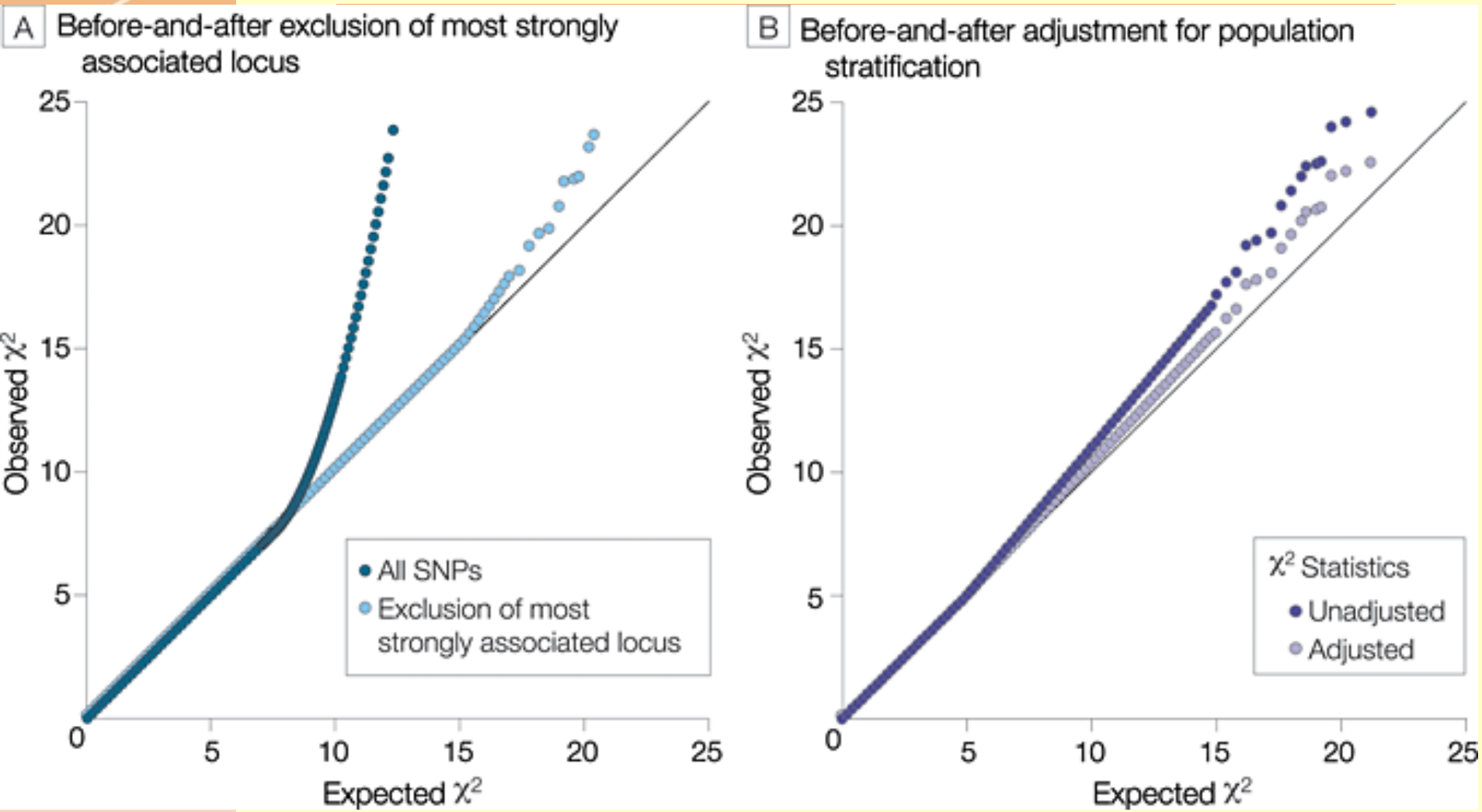
	3-Stage Study[b]		4-Stage Study[c]	
Stage	Case Participants/ Control Participants	SNPs Analyzed	Case Participants/ Control Participants	SNPs Analyzed
1	400/400	500 000	2000/2000	100 000
2	4000/4000	25 000	2000/2000	1000
3	20 000/20 000	25	2000/2000	20
4			2000/2000	5

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

Copyright restrictions may apply.

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

Quantile-Quantile Plots in Genome-wide Association Studies

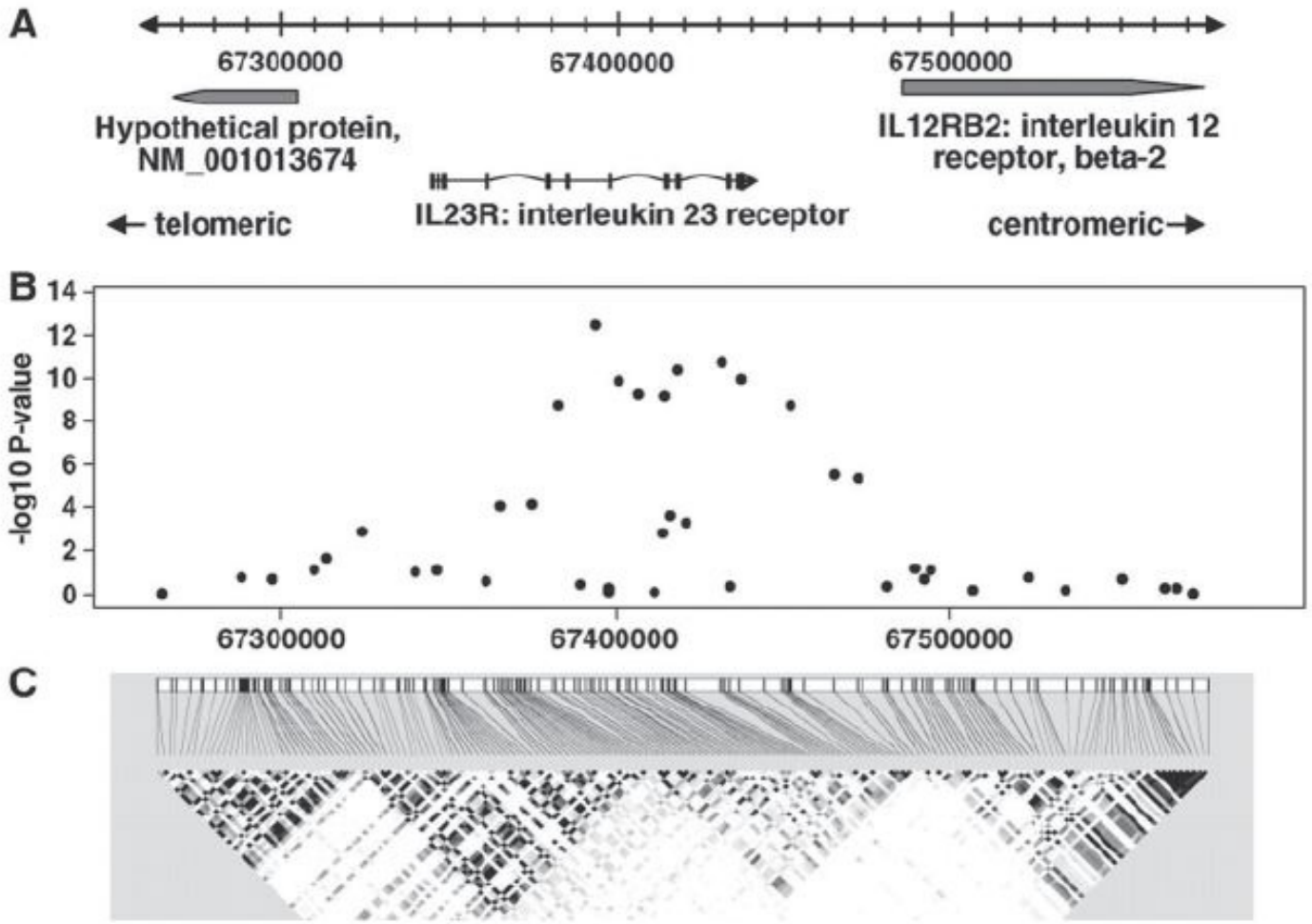


Copyright restrictions may apply.

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

Interleukin 23R & Inflammatory Bowel Disease

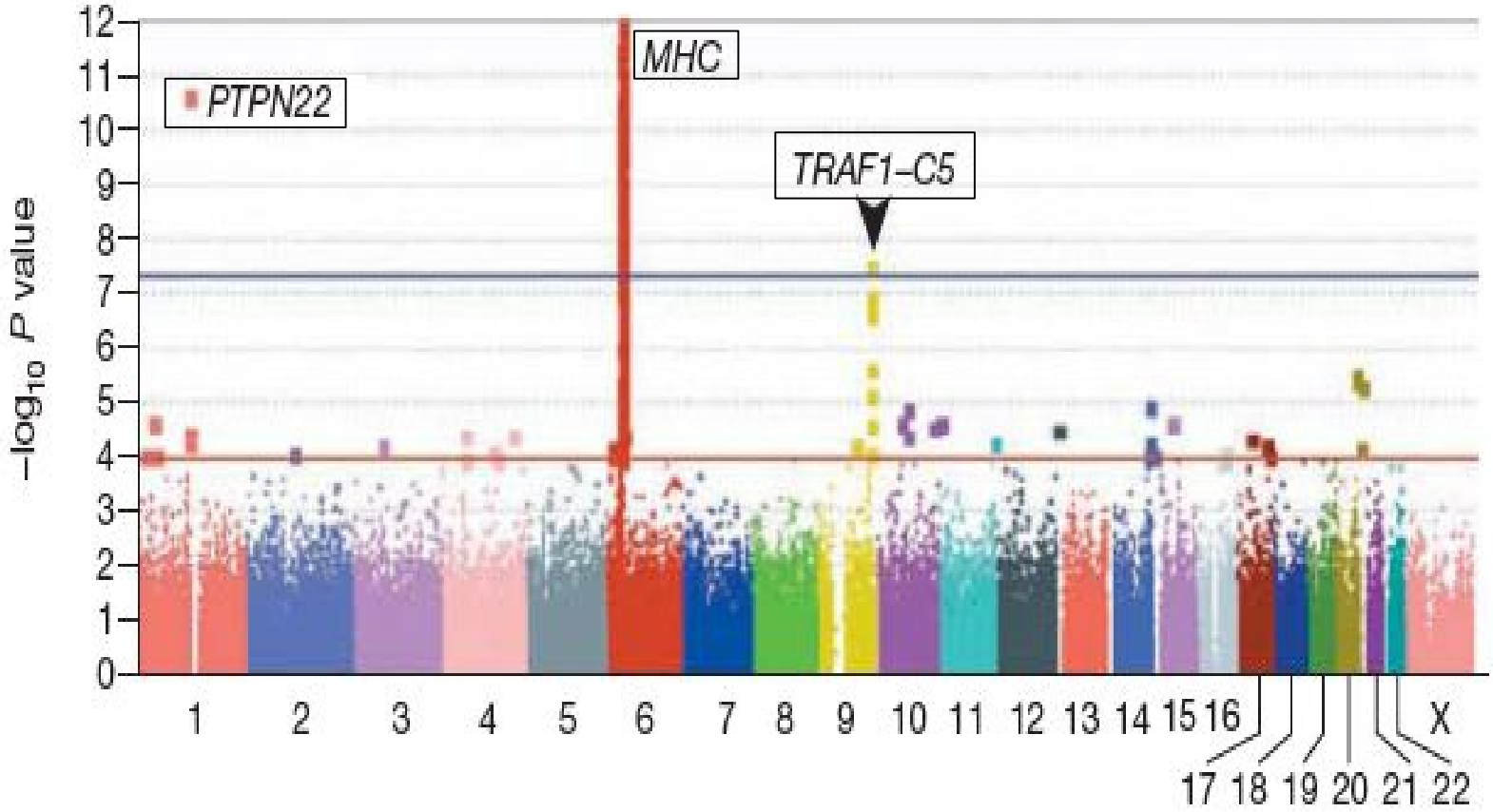
Figure 2. Associations in the *IL23R* Gene Region Identified by a Genome-wide Association Study of Inflammatory Bowel Disease



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

Genome Wide Associations in Rheumatoid Arthritis

Figure 3. Genome-wide Association Findings in Rheumatoid Arthritis



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

Association of Alleles & Genotypes

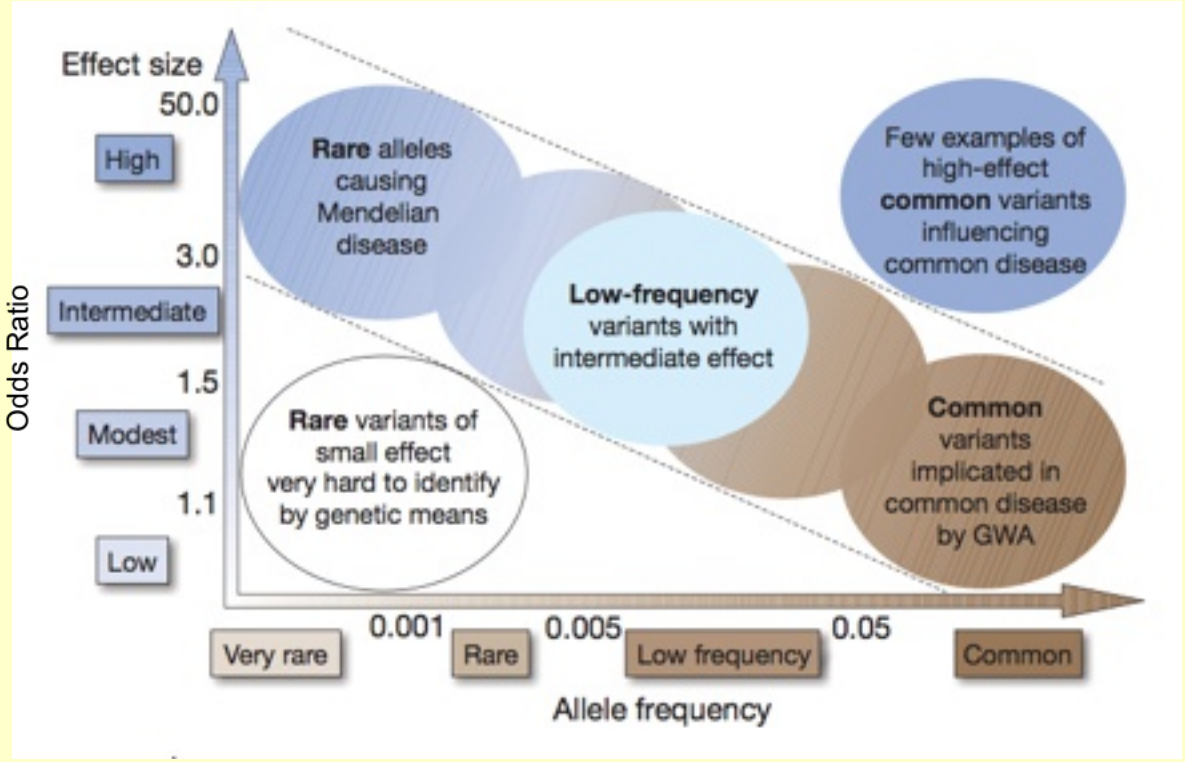
Table 3. Association of Alleles and Genotypes of rs6983267 on Chromosome 8q24 With Colorectal Cancer[a]

	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.[56]
[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

Low Heritability of Common SNPs



- Rare High Penetrance Variants Carry High Risk
- Common SNPs Carry Low Risk
- Multiple Variants May Increase Risk Synergistically
- Common SNPs Associated with Genes Containing High Risk Alleles
- Common SNPs Associations can Suggest Regions to Sequence in Cohorts or Trios

Manolio et al. Nature 461, 747-753 (2009)

Genetic Loci Associated with Hypertriglyceridemia

http://www.ncbi.nlm.nih.gov/pubmed/20657596

Table 2 Genetic loci associated with HTG

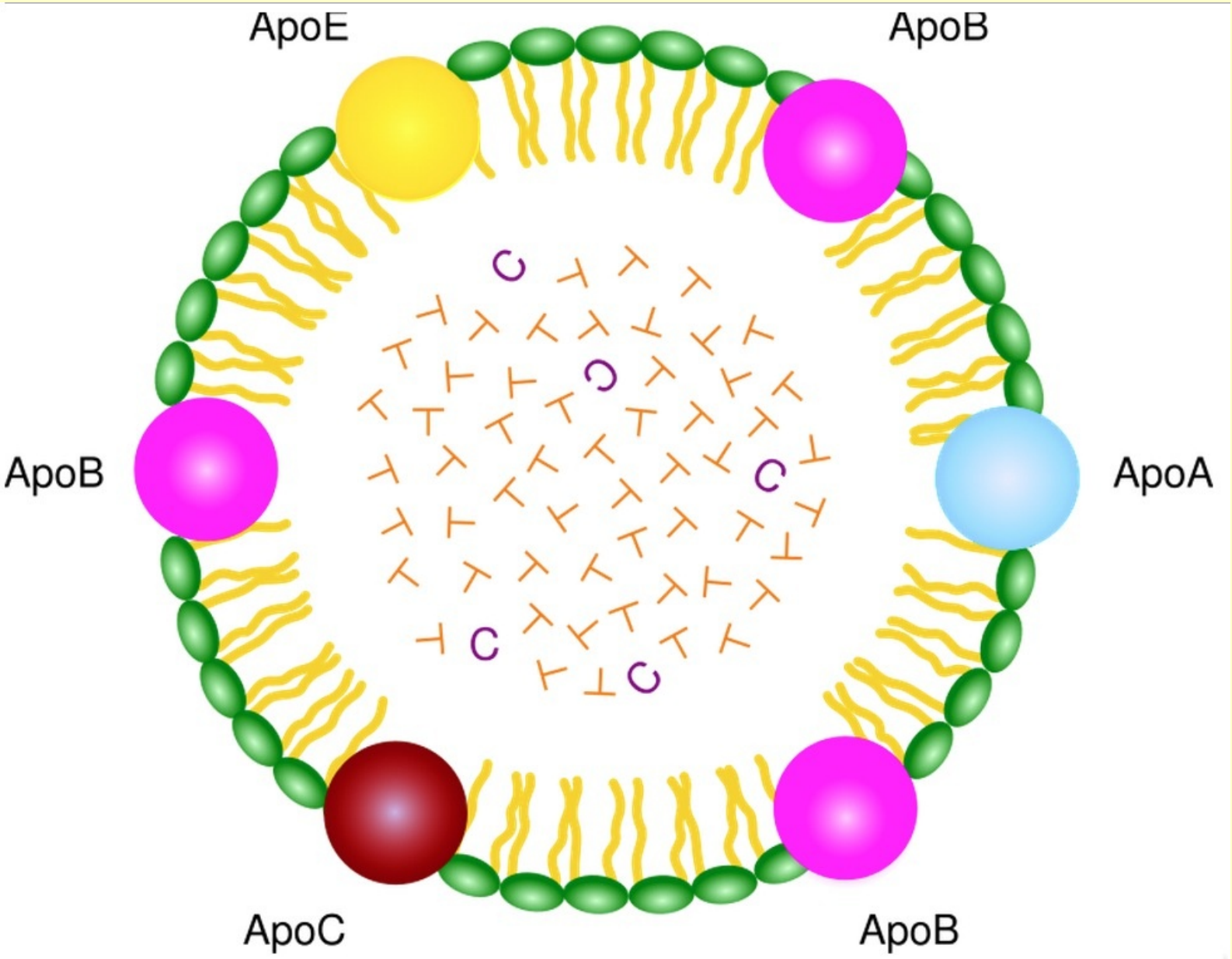
Locus	SNP	Chr.	Position	Minor allele	HTG MAF	Control MAF	OR (95% CI)	*P*
APOA5	rs964184	11	116.2	G	0.33	0.14	3.28 (2.61–4.14)	5.4×10^{-24}
GCKR	rs1260326	2	2.8	T	0.52	0.41	1.75 (1.45–2.12)	6.5×10^{-9}
LPL	rs7016880	8	19.9	C	0.03	0.10	0.32 (0.21–0.49)	2.0×10^{-7}
APOB	rs4635554	2	21.2	G	0.39	0.31	1.67 (1.38–2.02)	2.0×10^{-7}
MLXIPL	rs714052	7	72.5	G	0.07	0.13	0.44 (0.31–0.62)	0.000003
TRIB1	rs2954029	8	126.6	T	0.37	0.46	0.71 (0.59–0.86)	0.0004
ANGPTL3	rs10889353	1	62.9	C	0.27	0.32	0.73 (0.59–0.89)	0.002
NCAN	rs17216525	19	19.5	T	0.07	0.09	0.71 (0.50–1.00)	0.05
FADS	rs174547	11	61.3	C	0.40	0.33	1.20 (0.99–1.44)	0.07
XKR6	rs7819412	8	11.1	G	0.46	0.50	0.87 (0.72–1.05)	0.14
PLTP	rs7679	20	44.0	C	0.20	0.19	1.17 (0.94–1.47)	0.16

Nat Genet. 2010 Aug;42(8):684-7. Epub 2010 Jul 25.

Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia.

© Doug Brutlag 2015

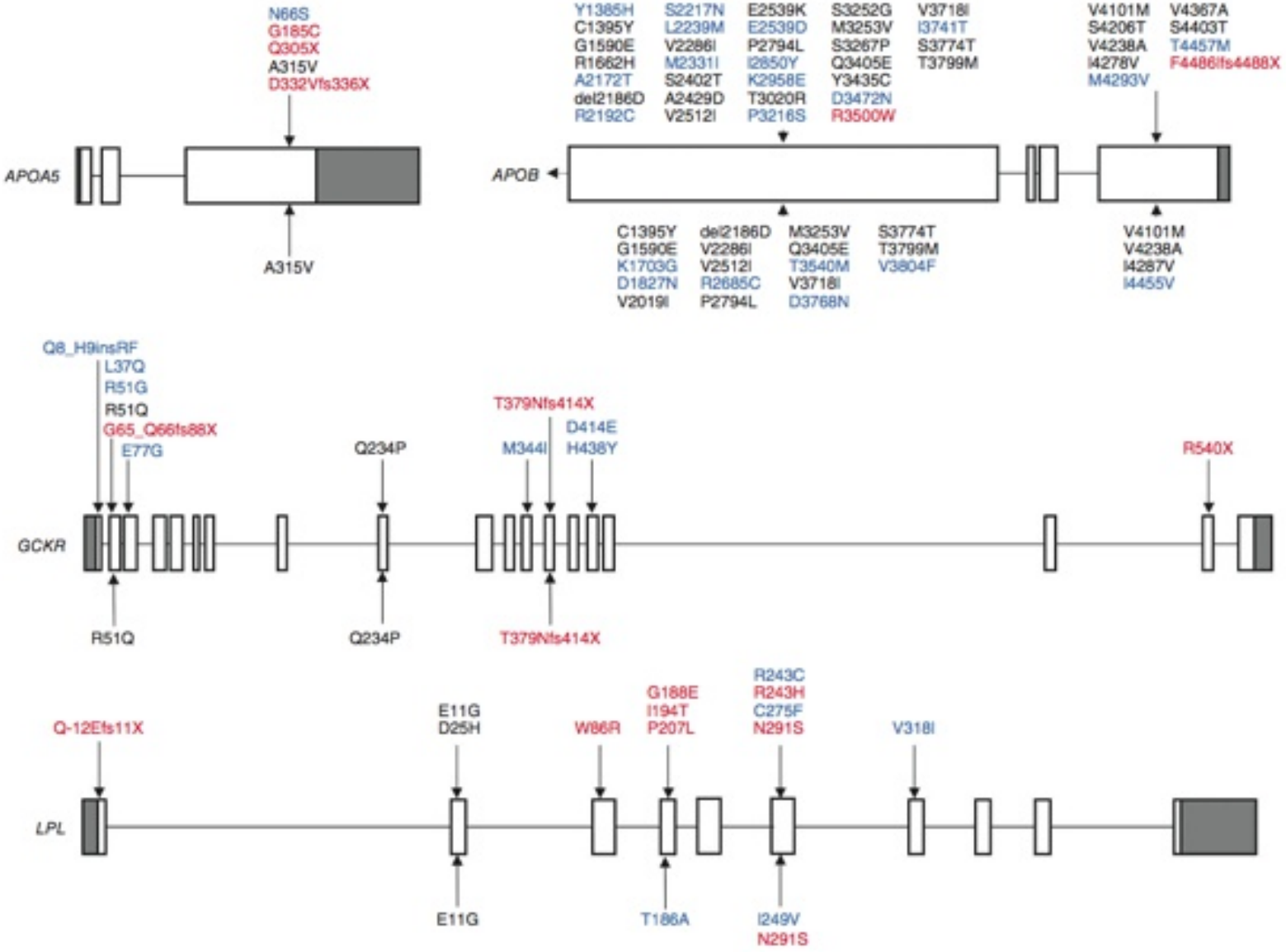
Chylomicrons



http://en.wikipedia.org/wiki/Chylomicron

© Doug Brutlag 2015

Novel Rare Variants in GWAS Genes for Hypertriglyceridemia
http://www.ncbi.nlm.nih.gov/pubmed/20657596



© Doug Brutlag 2015

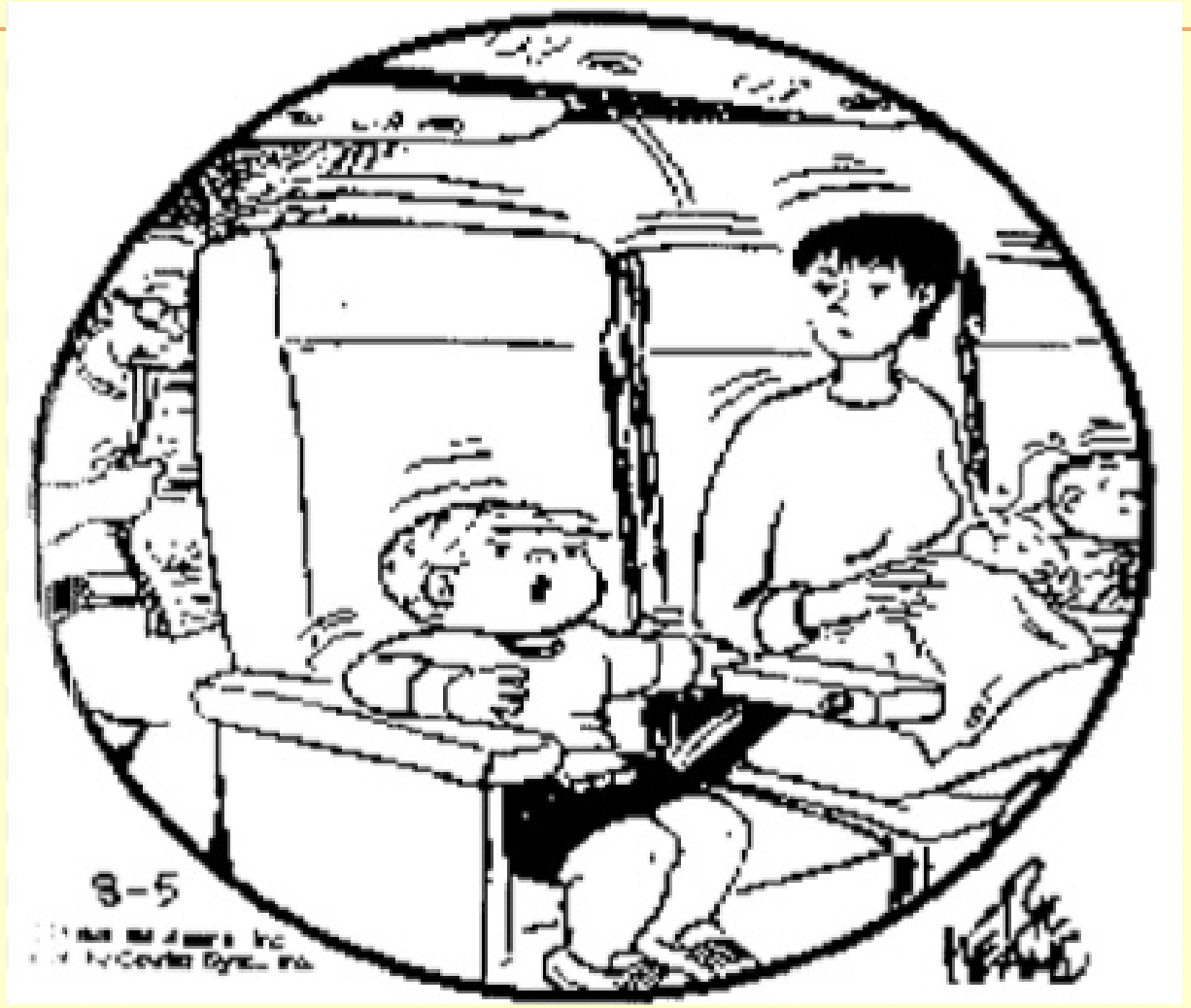
Summary

- Genome-wide association studies make no assumptions about disease mechanism or cause
- Genome-wide association studies usually discover only gene regions correlated with disease, NOT genes that cause the disease.
- Genome-wide associations indicate
  - Genes and regions to reanalyze by complete sequencing for causal genes or variations
  - Subpopulations that may be enriched for causal variations
  - Genes and gene products for functional and structural studies
  - Genes to examine for regulatory studies
- Genome-wide association studies coupled with proper biological and structural studies can lead to:
  - Unexpected causes for disease that could not have been predicted
  - Unexpected mechanisms for disease (missense mutations, regulatory changes, alternative splicing, copy number variation etc.)
  - Multiple pathways and multiple genes involved in disease
  - Novel diagnostics and prognosis
  - Novel treatments

© Doug Brutlag 2015

Association versus Causality

http://en.wikipedia.org/wiki/Correlation_does_not_imply_causation



I wish they didn't turn on that seatbelt sign so much!
Every time they do, it gets bumpy.

Courtesy David Feldman

Summary

- Genome-wide association studies make no assumptions about disease mechanism or cause
- Genome-wide association studies usually discover only gene regions correlated with disease, NOT genes that cause the disease.
- Genome-wide associations indicate
  - Genes and regions to reanalyze by complete sequencing for causal genes or variations
  - Subpopulations that may be enriched for causal variations
  - Genes and gene products for functional and structural studies
  - Genes to examine for regulatory studies
- Genome-wide association studies coupled with proper biological and structural studies can lead to:
  - Unexpected causes for disease that could not have been predicted
  - Unexpected mechanisms for disease (missense mutations, regulatory changes, alternative splicing, copy number variation etc.)
  - Multiple pathways and multiple genes involved in disease
  - Novel diagnostics and prognosis
  - Novel treatments

© Doug Brutlag 2015

GWAS References

How to Use an Article About Genetic Association: A: Background Concepts John Attia et al. (2009)
JAMA 301, 74-81

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

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

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)

Manolio, T.A. et. al.,(2009) Finding the missing heritability of complex diseases. Nature 461, 747-753.

Dickson, S. P.. et al. (2010) Rare Variants Create Synthetic Genome-Wide Associations. PLOS Biology 8, 1-12.

Johansen et al. (2010) Excess of rare variants in genes identified by genome-wide association study of hypertriglyc

© Doug Brutlag 2015