Linking Genes to Diseases: Leveraging the Human Genome

Doug Brutlag, Professor Emeritus of Biochemistry & Medicine (by courtesy) Stanford University School of Medicine
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 (smoking, drinking, drugs & other addictions)
  – Many complex diseases can also be monitored by increased vigilance (another behavioral modification)
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.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Gene</th>
<th>Variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peptic ulcer</td>
<td>ABO</td>
<td>O</td>
</tr>
<tr>
<td>IDDM*</td>
<td>HLA</td>
<td>DR3,4</td>
</tr>
<tr>
<td>Alzheimer dementia</td>
<td>APOE</td>
<td>E4</td>
</tr>
<tr>
<td>Deep venous thrombosis*</td>
<td>F5 (R506Q)</td>
<td>Leiden</td>
</tr>
<tr>
<td><em>Falciparum malaria</em></td>
<td>HBB</td>
<td>βs</td>
</tr>
<tr>
<td>AIDS*</td>
<td>CCR5</td>
<td>Δ32</td>
</tr>
<tr>
<td>Colorectal cancer*</td>
<td>APC</td>
<td>3920A</td>
</tr>
<tr>
<td>NIDDM*</td>
<td>PPARγ</td>
<td>12A</td>
</tr>
</tbody>
</table>
Using SNPs to Track Predisposition to Disease and other Genetic Traits

DNA from different individuals sequenced

Variation at a single nucleotide

Some individuals will have one version of the SNP, some the other

Sample with disease

Normal population

A higher than expected incidence in a disease group suggests SNPIG is associated with a disease (or SNPIA is protective)

In a population, a certain percentage will have one version, the rest the other

© Gibson & Muse, A Primer of Genome Science
Genome-Wide Association Study: A Brief Primer

Control Population               Disease Population

WTCCC, Nature 2007
Manhattan at Night
A Quantitative Gene-Expression Association

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)
Genome Wide Association of type 2 Diabetes
4549 cases, 5579 controls & 317,503 SNPs
# Top 10 Diabetes Genes from Genome-Wide Association Study

<table>
<thead>
<tr>
<th>Gene</th>
<th>Odds Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCF7L2</td>
<td>1.37</td>
<td>1.0 x 10^{-48}</td>
</tr>
<tr>
<td>IGF2BP2</td>
<td>1.14</td>
<td>8.9 x 10^{-16}</td>
</tr>
<tr>
<td>CDKN2A/B</td>
<td>1.20</td>
<td>7.8 x 10^{-15}</td>
</tr>
<tr>
<td>FTO</td>
<td>1.17</td>
<td>1.3 x 10^{-12}</td>
</tr>
<tr>
<td>CDKAL1</td>
<td>1.12</td>
<td>4.1 x 10^{-11}</td>
</tr>
<tr>
<td>KCNJ11</td>
<td>1.14</td>
<td>6.7 x 10^{-11}</td>
</tr>
<tr>
<td>HHEX</td>
<td>1.13</td>
<td>5.7 x 10^{-10}</td>
</tr>
<tr>
<td>SLC30A8</td>
<td>1.12</td>
<td>5.3 x 10^{-8}</td>
</tr>
<tr>
<td>Chr 11</td>
<td>1.23</td>
<td>4.3 x 10^{-7}</td>
</tr>
<tr>
<td>PPARG</td>
<td>1.14</td>
<td>1.7 x 10^{-6}</td>
</tr>
</tbody>
</table>
GLUCOSE → ATP → ADP

K^+

Calcium Channel

Ca^{2+} → Zn^{2+}

Insulin

KCNJ11

SLC30A8 – A Beta Cell Zinc Transporter

© Francis Collins, 2009
The Great Wave of GWAS Studies

http://gwas.nih.gov/
Catalog of GWAS Studies
https://www.ebi.ac.uk/gwas/
Published Genome-Wide Associations through 6/2010
http://www.genome.gov/GWAStudies

Published GWA Reports, 2005 – 2013

Total Number of Publications

Calendar Quarter

© Doug Brutlag 2016
Published Genome-Wide Associations as of February 10, 2016
http://www.ebi.ac.uk/gwas/diagram#
I wish they didn’t turn on that seatbelt sign so much! Every time they do, it gets bumpy.
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:


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?

- Hispanics
- European-Americans
- African-Americans

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

Second Hit
- Oxidative Stress
- Lipid Peroxidation
- Anti-virals
- Cytokines
Genome-wide Association Study of Fatty Liver in Dallas Heart Study Cohort (2,111 patients and 2,299 controls)

- \( P = 5.9 \times 10^{-10} \)

- \( 5.4 \times 10^{-6} \)
PNPLA3 & Hepatic Triglyceride Metabolism

Liver

Acetyl CoA → FFA

FFA

Remnants

Liver → VLDL

Adipose Tissue

PNPLA2 (ATGL)

PNPLA3 (Adiponutrin)

Mito

Fasting

Feeding
I148M & Catalytic Site of PNPLA3

Patatin Like Domain

Ser47  Asp166  Ile148  Met148

© Helen Hobbs 2009
Genetic Contribution to Ethnic Differences in Hepatic Steatosis

<table>
<thead>
<tr>
<th>Minor Allele Frequency</th>
<th>African-Americans</th>
<th>European-Americans</th>
<th>Hispanics</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.17</td>
<td>0.23</td>
<td>0.49</td>
<td></td>
</tr>
</tbody>
</table>

Prevalence of Hepatic Steatosis (%)

© Helen Hobbs 2009
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
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 or Subpopulations

Genome Wide Association Study
Third Homework Suggestion

• Step 1: Read:
  • How to Use an Article About Genetic Association: A: Background Concepts
    John Attia et al. (2009) JAMA 301, 74-81

• Please search the GWAS Catalog for a disease of interest to you.

• After you find a GWA Study on a disease of interest to you, please read the paper describing the genome-wide association study and report to me 1) the reference for the paper and 2) genes or SNPs that are most highly correlated with the disease. 3) the odds ratio and heritability of each gene and 4) Also please tell me if knowledge of those SNPs or genes sheds any light on the basis for the disease.

• A more advanced treatment of reading GWAS papers is:
  • How to Interpret a Genome-wide Association Study
    Thomas A. Pearson; Teri A. Manolio (2008) JAMA 299, 1335-1344
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


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

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)


Genome Wide Associations in Rheumatoid Arthritis

Figure 3. Genome-wide Association Findings in Rheumatoid Arthritis