Genetics and Inflammatory Bowel Disease

Louis Ghanem M.D, Ph.D.
Attending Physician, The Children’s Hospital of Philadelphia
Division of Gastroenterology, Hepatology and Nutrition
Etiology of IBD

Genetic Predisposition

Mucosal Immune System (Adaptive/Innate)

Environmental Triggers (Luminal Bacteria, Infection)

IBD
What Is the Human Genome?

**Genome:**

The complete supply of DNA

- Includes all the genes and the spaces in between
DNA

- What is DNA?
  - Deoxyribose Nucleic Acid

- What is it made of?
  - Adenine, Guanine, Thymine, Cytosine

- What does it do?
  - Acts as our body's instruction manual

Diagram:
- Bases
- Sugar
- Phosphate
- Backbone
- Base pair
- Adenine (A)
- Cytosine (C)
- Thymine (T)
- Guanine (G)
SNPs: Frequently Occurring Genetic Variants

Most of the population

At least 1 percent of the population

G to C

SNP site

Common sequence

Variant sequence

Functional protein

Functional protein

National Cancer Institute
Summary

- Human genome consists of DNA
- DNA is packaged into chromosomes
- 25,000 genes in the genome (3%)
- DNA (Genes) is made up of 4 bases – A, C, T, G
- Most genes makes proteins
- SNPs are sites of common genetic variation
Genetic epidemiology of IBD

- 1 in 300 to 500 Caucasians are affected by IBD in North America & Europe

<table>
<thead>
<tr>
<th>Study</th>
<th>Crohn’s disease</th>
<th>Ulcerative colitis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Monozygotic</td>
<td>Dizygotic</td>
</tr>
<tr>
<td>Sweden</td>
<td>9/18 (50%)</td>
<td>1/26 (4%)</td>
</tr>
<tr>
<td>Britain</td>
<td>5/25 (20%)</td>
<td>3/46 (7%)</td>
</tr>
<tr>
<td>Denmark</td>
<td>5/10 (50%)</td>
<td>0/27 (0%)</td>
</tr>
</tbody>
</table>

Identical

Fraternal
The rules of genetics:

- Well-defined physical patterns
- No environmental influence
- One gene to one trait
- If you have the gene you have the trait (100% penetrance)
- Two traits act independently
Crohn’s Disease and ulcerative colitis: Do not follow the rules!
IBD
Genetics of Complex Disease

Complex Disease:

• Several genes involved in the expression of the clinical trait

• Gene-gene interactions and gene-environment interactions

• Incomplete penetrance – age-dependent
SNPs: Frequently Occurring Genetic Variants

Most of the population

At least 1 percent of the population

G to C

Common sequence

Variant sequence

SNP site

Functional protein

Functional protein

SNP

National Cancer Institute
Genetic Discoveries in IBD

1859
1st "impact" description of UC

1932
Crohn's "terminal ileitis" JAMA article

1950's
Family clusters

1972
1st HLA associations

1990's
IBD1
DRB1*1502
DRB1*0103
NOD2

2000
IBD1

2001

2002-Present

Genome Wide Association Studies
## CD Genetics 2010

Genome-wide meta-analysis increases to 71 the number of confirmed Crohn’s disease susceptibility loci

<table>
<thead>
<tr>
<th>Prior</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOD2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5q31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MHC</td>
<td>TNFSF15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8q24</td>
<td>TAGAP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1q24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1q32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6q21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7p12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8q24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10p11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13q14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19p13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21q21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITLN1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTPN22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL12B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDKAL1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCR6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NNX2-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JAK2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTPN2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C11orf30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZNF365</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IRGM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LRRK2/MUC19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORMDL3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAT3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAP3K7IP1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19q13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCKR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNMT3A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNFSF11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3CR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTMR3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNFRSF6B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRDX5/ESRRA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UBE2D1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL2RA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

McGovern, D, Cedar Sinai MC
ImmunoChip Consortium
Specific to 12 immunologically related human diseases:

- Type 1 Diabetes
- Ankylosing spondylitis
- Celiac disease
- Multiple sclerosis
- Psoriasis
- SLE
- Autoimmune thyroid disease
- Crohn’s disease
- IgA deficiency
- Primary biliary cirrhosis
- Rheumatoid arthritis
- Ulcerative colitis

• 38,565 cases & 37,747 controls

• Meta-analysis:
  - GWAS
  - New cases genotyped on Immunochip
    • 14,763 CD cases
    • 10,920 UC cases
    • 15,977 Controls

• 71 new loci→163 genome-wide significant loci

Jostins et al, Nature 2011
McGovern, D, Cedar Sinai MC
‘Shared’ genes immune-mediated diseases

Lees et al Gut 2011
Relative Importance of Biologic Mechanisms in IBD

Targets for treatment
Genetic Contribution to Disease

Variant Frequency

High

Low

Number of People

Very Few

Many

Next Generation DNA Sequencing

Single Gene Disorders

GWAS

Common Variants Complex Disorders
Penetrance

- Genotype and phenotype present
- Genotype and phenotype absent
- Genotype present, phenotype (disease) absent
Unanswered Research Questions in IBD Genetics

• Explained Genetic Variance
  - 13.6% in CD
  - 7.5% in UC
• What are **ALL** the genes that predispose to, protect against, or determine subtypes and course of IBD?
  – Most of the 163 risk regions are associated with multiple genes
• What are the **functional mechanisms** of IBD genes?
• How do IBD genes interact with each other and with environmental factors to predispose to or modify the course of IBD?
Unanswered Research Questions in IBD Genetics

• Can IBD risk genes predict:
  – Likelihood of developing IBD?
  – Disease subtype (CD vs UC)?
  – Clinical course?
  – Response to Therapy?
Advances in Genetic Discovery

• Next Generation Sequencing:
  – Exome Sequencing – Gene regions only
    • Coding region of genes that effect protein function
    • 1% of the entire genome
    • 2012 - $1200-2400; 4-6 weeks for analysis
  – Whole Genome Sequencing – Everything!
    • up to 3 billion bases per run
    • 1 human genome 30x per day
    • 2010 - $20000 - $50000
    • 2012 - ~$5000; analysis – variable, not standardized
  – Critical: Must balance data acquisition with bioethical considerations!

• Identify less common, higher-risk variants and potentially novel mutations
  – Family history of IBD

• CHOP - pilot studies using exome sequencing underway
Advances in Genetic Discovery

• Functional Consequences of Genetic Variation
  – A HUGE UNDERTAKING!
  – Animal models of Disease
  – Genomics, Epigenetics, RNA sequencing, Metabolomics, Proteomics…

• GWAS in non-white populations

• Further meta-analysis of existing studies

• Define the interplay between a host’s genetics, the environment and gut bacteria

• Personalized medicine…
What is personalized medicine?

• Customized health care **Informed** by the individual’s unique genomic, clinical & environmental information
  – Individual’s Susceptibility to Disease
  – Management of Disease Progression
  – Response to & toxicity from therapies

• IBD is well suited to this approach
  – Genetically Heterogeneous (Complex) Disease
  – Understanding of the Genetic Architecture of IBD
  – Ability to alter the Natural History of Disease
Charting a course for genomic medicine from base pairs to bedside

Eric D. Green¹, Mark S. Guyer¹ & National Human Genome Research Institute*
Charting a course for genomic medicine from base pairs to bedside

Eric D. Green, Mark S. Guyer & National Human Genome Research Institute

BOX 2
Imperatives for genomic medicine

Opportunities for genomic medicine will come from simultaneously acquiring foundational knowledge of genome function, insights into disease biology and powerful genomic tools. The following imperatives will capitalize on these opportunities in the coming decade.

Making genomics-based diagnostics routine. Genomic technology development so far has been driven by the research market. In the next decade, technology advances could enable a clinician to acquire a complete genomic diagnostic panel (including genomic, epigenomic, transcriptomic and microbiomic analyses) as routinely as a blood chemistry panel.

Defining the genetic components of disease. All diseases involve a genetic component. Genome sequencing could be used to determine the genetic variation underlying the full spectrum of diseases, from rare Mendelian to common complex disorders, through the study of upwards of a million patients; efforts should begin now to organize the necessary sample collections.

Comprehensive characterization of cancer genomes. A comprehensive genomic view of all cancers\textsuperscript{1-7} will reveal molecular taxonomies and altered pathways for each cancer subtype. Such information should lead to more robust diagnostic and therapeutic strategies and a roadmap for developing new treatments\textsuperscript{24,25}.

Practical systems for clinical genomic informatics. Thousands of genomic variants associated with disease risk and treatment response are known, and many more will be discovered. New models for capturing and displaying these variants and their phenotypic consequences should be developed and incorporated into practical systems that make information available to patients and their healthcare providers, so that they can interpret and reinteract the data as knowledge evolves.

The role of the human microbiome in health and disease. Many diseases are influenced by the microbial communities that inhabit our bodies (the microbiome)\textsuperscript{101}. Recent initiatives\textsuperscript{102,103} (http://www.human-microbiome.org) are using new sequencing technologies to catalogue the resident microflora at distinct body sites, and studying correlations between specific diseases and the composition of the microbiome.\textsuperscript{104}. More extensive studies are needed to build on these first revelations and to investigate approaches for manipulating the microbiome as a new therapeutic approach.
Charting a course for genomic medicine from base pairs to bedside

Eric D. Green¹, Mark S. Guyer¹ & National Human Genome Research Institute

BOX 2
Imperatives for genomic medicine

Opportunities for genomic medicine will come from simultaneously acquiring foundational knowledge of genome function, insights into disease biology and powerful genomic tools. The following imperatives will capitalize on these opportunities in the coming decade.

Making genomics-based diagnostics routine. Genomic technology development so far has been driven by the research market. In the next decade, technology advances could enable a clinician to acquire a complete genomic diagnostic panel (including genomic, epigenomic, transcriptomic and microbiomic analyses) as routinely as a blood chemistry panel.

Defining the genetic components of disease
The role of the human microbiome in health and disease......
Making genomics-based diagnostics/prognostics routine......
Practical systems for clinical genomics informatics……
One Major Conclusion of Genetic Studies to Date: 

*Host-microbe interactions are critical determinants of the predisposition to IBD*
Questions?