Finding new therapeutic targets through genetics & sequencing

Judy H. Cho, M.D.
Yale University
6.29.2012
Overview

- Examples from inflammatory bowel disease
  - IL-23 pathway
  - NOD2, mycobacterial diseases, & innate immune cells
  - TNF pathway
- Systematically leveraging high throughput sequencing to prioritize new targets
  - Phenotype driven
  - Genotype (Encode data) driven
The IL-23 pathway in immune-mediated diseases
Multiple signals in IL23R gene region: uncommon protective Arg381Gln allele
IL-23 signaling (Th17 cells)

Cytokine

- IL23A(p19) on chr12q13
- IL12B(p40) on chr5q33

Receptor

- IL23R on chr1p31
- IL12RB1 on chr19p13

5/7 members of the primary IL-23 pathway associated in IBD

Th17 cells:
- Patrol mucosal surfaces
- Fungal and bacterial defense

* IBD associated
Arg381Gln protective allele in IL23R is a loss-of-function allele

- Anti-p40 treatment (blocks IL12/23):
  - Approved in psoriasis
  - IBD phase III studies ongoing
  - Issues: How to block IL23 pathway?
  - Blocking IL-23 alone vs. IL12/23??
  - Blockade at what level? Receptor? JAK?
NOD2, mycobacterial disease & innate immune cells
The Immunochip effort in IBD: a large scale international collaboration

- 38,565 cases, 37,747 controls
- 71 new loci → 163 loci with genome-wide significant association (~1500 genes)
163 loci → improved network analysis—key role of directionality

- 140 CD & IBD GWAS SNPs
- Genes in 59 cis eSNP loci
  - From liver, omental, subq cis eSNPs from multiple tissues
- CD genes
  - Within GWAS loci
  - Human only
- Microarray datasets
  - Across numerous tissues
- Enrichment
- Modules to be screened
  - Enrichment in green module
  - IBD subnetwork
- Bayesian network
  - Causal regulators
  - Color-coded subnetwork
- TB co-expression modules
  - GO pathways
- Color-coded CD-TB overlap
- Color-coded GO annotation

- cis → trans validation

- Top module: omental adipose (macrophage enriched) from obese patients

Eric Schadt
Ken Hui
163 loci \(\rightarrow\) improved network based analyses based on gene co-expression

- Co-expression modules: tracking similar gene expression based on large microarray datasets
- The co-expression module with the greatest enrichment of IBD-associated genes: 523 gene module in omental adipose tissue (macrophage-enriched gene expression) — value of direct ex-vivo tissue analysis

Gene in IBD-associated locus

NOD2
NOD2-centric view of the submodule: 7 IBD-associated genes near NOD2

- LGALS9
- Autophagy
- Induced with Mtb infection
- Modulates mycobacteriosis
- M. tuberculosis susceptibility
- SLC11A1 (aka NRAMP1)
- Vitamin D receptor
- HCK: key for differentiation of M2 macrophages (anti-inflammatory $\rightarrow$ IL10)

Highly correlated RNA expression between NOD2, IL10 & HCK (hematopoietic cell kinase)
The TNF pathway
IBD is a TNF-mediated disorder

- TNF-overexpressing mice develop ileitis and arthritis
- Anti-TNF is a highly effective treatment for IBD
- GWAS: multiple TNF-mediated signals
  - NF-κB (NFKB1, REL, RELA, TNFAIP3)
- TNF: crucial in pathogen eradication—reactivation of tuberculosis a side effect of anti-TNF therapy
Molecular integration of TNF and 3’UTRs: crucial role of kinetics/functional responses

The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules

Shengli Hao & David Baltimore

TNF
A20 (TNFAIP3)

CCL2—max association in 3’UTR

Kinetics of gene expression: multiple ub/dub associations: NDFIP1, CPEB4, CUL2, UBE2L3, as well as TNFAIP3 (15 loci involved (p < 0.001)

Few AREs (<2), very stable mRNA

Some AREs (2-4), moderately stable mRNA

Many AREs (4-10), unstable mRNA
Systematically leveraging high throughput sequencing to prioritize new targets: phenotype to genotype (1)

- LOF, protective alleles as ideal therapeutic targets
  - PCSK9 & CAD
  - IL23R & psoriasis/IBD/ankylosing spondylitis
  - CCR5 & HIV
  - IFIH1 & T1DM?

- Value of sequencing
  - Targeted re-sequencing of GWAS signals: enormous structure-function data useful for improved targeting
Systematically leveraging high throughput sequencing to prioritize new targets: phenotype to genotype (2)

- Early onset, severe cases: medical resequencing
  - LOF IL10 pathway genes $\rightarrow$ bone marrow transplantation
- New biology: Nick Volker—young boy with early onset IBD $\rightarrow$ XIAP mutation (essential for NOD2-signaling)
- -omics data & systems biology
  - RNASeq: improved quantification should improve predictive models
- Systematic interrogation of disease-associated transcription factors: ChIPSeq
- Cross-phenotype analyses: immune-mediated diseases & infectious diseases
Striking overlap of loci between diseases: the genetics of infectious diseases

6/7 leprosy loci also IBD loci
NOD2
IL23R
TNFSF15
RIPK2
LRRK2
C13ORF31

82

53

Immune-mediated diseases

IBD loci

6/8 MSMD genes within IBD loci
IL12B
STAT1
IRF8
TYK2
STAT3
IFNGR2

MSMD, Mendelian susceptibility to mycobacterial disease
Gains of glycosylation comprise an unexpectedly large group of pathogenic mutations

Guillaume Vogt¹, Ariane Chappier¹, Kun Yang¹,², Nadia Chuzhanova³,⁴, Jacqueline Feinberg¹, Claire Fieschi¹,⁵, Stéphanie Boisson-Dupuis¹, Alexandre Alcais¹, Orchidée Filipe-Santos¹, Jacinta Bustamante¹, Ludovic de Beaucoudrey¹, Ibrahim Al-Mohsen⁶, Sami Al-Hajjar⁶, Abdulaziz Al-Ghonaium⁶, Parisa Adimi⁷, Mehdi Mirsaedi⁷, Soheila Khalilzadeh⁷, Sergio Rosenzweig⁸,¹⁷, Oscar de la Calle Martin⁹, Thomas R Bauer¹⁰, Jennifer M Puck¹¹, Hans D Ochs¹², Dieter Furthner¹³, Carolin Engelhorn¹⁴, Bernd Belohradsky¹⁴, Davood Mansouri⁷, Steven M Holland⁸, Robert D Schreiber¹⁵, Laurent Abel¹, David N Cooper⁴, Claire Soudais¹ & Jean-Laurent Casanova¹,²,¹⁶

Mutations involving gains of glycosylation have been considered rare, and the pathogenic role of the new carbohydrate chains has never been formally established. We identified three children with mendelian susceptibility to mycobacterial disease who were homozygous with respect to a missense mutation in IFNGR2 creating a new N-glycosylation site in the IFNγR2 chain. The resulting additional carbohydrate moiety was both necessary and sufficient to abolish the cellular response to IFNγ. We then searched the Human Gene Mutation Database for potential gain-of-N-glycosylation missense mutations; of 10,047 mutations in 577 genes encoding proteins trafficked through the secretory pathway, we identified 142 candidate mutations (~1.4%) in 77 genes (~13.3%). Six mutant proteins bore new N-linked carbohydrate moieties. Thus, an unexpectedly high proportion of mutations that cause human genetic disease might lead to the creation of new N-glycosylation sites. Their pathogenic effects may be a direct consequence of the addition of N-linked carbohydrate.
Genotype to phenotype: the Encode approach

- Covalent modifications: missense mutations &
  - Glycosylation
  - Phosphorylation
  - Ubiquitination/sumoylation

- Regulation of expression
  - Conserved sequences
  - AU-rich elements: RNA-binding protein sites in 3’UTR
  - TF-binding sites, miRNA-binding sites, splice sites

- Analysis and information dissemination: validity & magnitude of effects
  - Bioinformatic probability vs. experimental validation
  - Frequency, population specificity
  - Distinguishing negative selection from drift
Acknowledgements

- NIDDK IBD Genetics Consortium
- Steven Brant, Richard Duerr, Dermot McGovern, John Rioux, Mark Silverberg, Mark Daly
- DCC: Phil Schumm, Yashoda Sharma, Clarence Zhang, Kaida Ning

- International IBD Genetics Consortium