The Complexity of Simple Genetics
The ciliopathies: a journey into variable penetrance and expressivity

Allelism at a single locus is insufficient to explain phenotypic variability

Bardet-Biedl Syndrome
Moderate

Meckel-Gruber Syndrome
Severe

Nephronophthisis
Mild

Senior Loken Syndrome
Moderate

Joubert Syndrome
Moderate

17 years of BBS gene discovery...
25+ genes implicated as recessive drivers
Lesson 1: IFT139

Functional Validation Reveals Association Signal

♦ 28 individual novel variants unique to patients:
  2 frame-shifting
  2 nonsense
  4 splice junction
  20 missense

♦ 10 individual novel variants unique to controls:
  (all missense)

♦ 10 individual variants found in cases and controls:
  (all missense)
Most *IFT139* missense variants are pathogenic

Functional testing of missense alleles in an in vivo model: Rescue efficiency of morpholino phenotypes

27/30 *IFT139* alleles detected in patients are deleterious in vivo

43 patients harbor pathogenic variants (5.7%)

Davis et al., Nat Genet. 43, 189-196 (2011).
Mutations in *IFT139* cause JATD and NPHP
Total pathogenic variant frequency is enriched ~5-fold in N. European cases vs. controls

Davis et al., Nat Genet. 43, 189-196
Lesson 2: it’s not all in the sequence
Systematic Testing for Genomic Variants

~800 ciliary genes

100 BBS patients
CNVs contribute to driver loci

**Driver CNVs**

**Mendelizers of SNVs**
~18% of BBS cases have CNVs

Significant enrichment in Cases vs Controls (p<0.0001)

Figure 1

Lindstrand et al 2016. AJHG 4, 318-336
Overall mutational burden in BBS: n=17 cases

Lindstrand et al 2016. AJHG 4, 318-336
Lesson 3: Overall mutational burden matters (CNVs + SNVs)
Examined ONLY “solved” cases
Mutational burden in BBS
Contributory alleles in cases vs controls

A. Discovery and Replication cohorts using in-cohort allele frequencies
- NEU cases in Discovery cohort (N=84)
- Entire Discovery cohort (N=102)
- Replication cohort (N=89)

B. Discovery cohort vs NEU control cohort

C. Replication cohort vs Exome control cohort

D. Collapsed case and control cohorts
- Combined control cohorts
- BBS cases meta-analysis
- Discovery cohort
- Non-BBS recessive cohort
Interactions are not distributed linearly...

Fold change in cases vs controls

- BBS9-BBS12: p=0.0043
- BBS5-BBS10: p=0.0477
- BBS4-BBS12: p<0.0001
- BBS4-MKKS: p<0.0001
- BBS1-BBS4: p=0.0226
Initial glimpses of mutational network
From a genocentric to a modular view of disease
“Data don’t make any sense, we will have to resort to statistics.”
Not all deleterious mutations are deleterious
An old splinter

PolyPhen said Benign!

Zebrafish assay said Pathogenic!

WHY???

 Were these observations “ciliary” oddities?

- **HumVar “Disease”** (22,207 variants):
  - 16,544
  - 3,250
  - 313

- **ClinVar “Pathogenic”** (10,596 variants):
  - 6,503
  - 530

- **24,304,185**
  - Found in MultiZ 100-Way alignment
  - (24,307,128 variants)

- **Pie chart**:
  - Largest estimate (12%)
  - Smallest estimate (3.0%)
Compensatory Mutations

- Secondary mutations that alleviate deleterious effect of primary mutations

Functional wild-type protein

Primary mutation (U) disrupts protein structure/function

Compensatory mutation (V) restores structure/function

Sayuri & Tetsuro, Frontiers in Microbiology, Vol 2, #267, 2012
Finding Potential Compensatory Mutations

<table>
<thead>
<tr>
<th>Human Protein</th>
<th>Mutated to -A- in human disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>IYKQALIFRLEGNIPESELLELFQTCAVLSPQSN</td>
<td></td>
</tr>
<tr>
<td>--V--F--Q--P--A--A--K--</td>
<td></td>
</tr>
<tr>
<td>--V--Q--Q--T--A--A--K--</td>
<td></td>
</tr>
<tr>
<td>--V--Q--H--A--I--K--</td>
<td></td>
</tr>
<tr>
<td>--V--Q--Y--A--K--</td>
<td></td>
</tr>
<tr>
<td>--------------------------N---E---K--</td>
<td></td>
</tr>
<tr>
<td>N--------------------------D----------N-----K--</td>
<td></td>
</tr>
<tr>
<td>N--------------------------D----------N-----K--</td>
<td></td>
</tr>
<tr>
<td>N--------------------------D----------N-----K--</td>
<td></td>
</tr>
<tr>
<td>--Q--------------------------D----------N-----K--</td>
<td></td>
</tr>
<tr>
<td>--Q--------------------------D----------N-----K--</td>
<td></td>
</tr>
</tbody>
</table>
Clinical features
Global developmental delay
microcephaly
feeding issues
failure to thrive
abnormal muscle tone
low immunoglobulins
frequent respiratory infections

Clinical testing
normal female microarray
metabolic testing – negative
extensive genetic testing – negative

**Prospective Utility for WES analysis**

- **BTG2**: Involved in the G1/S transition of the cell cycle
- **NOS2**: Nitric oxide synthase 2, inducible
- **TTN**: Titin
**BTG2 is the disease driver**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Nucleotide</th>
<th>All Allele #</th>
<th>All Allele MAF</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTG2</td>
<td>c.421G&gt;A</td>
<td>A=0/G=12996</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>PolyPhen-2³</th>
<th>SIFT⁴</th>
<th>Mutation Taster⁵</th>
</tr>
</thead>
<tbody>
<tr>
<td>p.V141M</td>
<td>Benign</td>
<td>Tolerated</td>
<td>Polymorphism</td>
</tr>
</tbody>
</table>

**A**
- **Control**
  - Bilateral expression
- **btg2 MO**
  - Unilateral expression
  - No expression
  - Bilateral expression

**B**

**C**
- **Control**
- **btg2 MO**

**D**

- **Microcephaly**
- **↓Neuronal expr.**
- **↓cell proliferation**
Conservation of BTG2 141M in multiple species
BTG2 has two compensatory mutations

Number of cells stained with phospho Histone H3

*P<0.01 vs V141M rescue alone

Jordan, Frangakis et al, Nature 2015
What is the biochemical basis of this phenomenon?
A really scary (?) possibility
An observation

From literature harvesting, 20-50% of mouse KOs do not match knock-ins

**Hypothesis**

Patient

<table>
<thead>
<tr>
<th>Missense mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease</td>
</tr>
</tbody>
</table>

Knockout mouse

| Deleterious       |

Knockin mouse

<table>
<thead>
<tr>
<th>No phenotype = cis complementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>?</td>
</tr>
</tbody>
</table>
# Exemplar: EIF2B5

<table>
<thead>
<tr>
<th>Gene</th>
<th>EIF2B5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human phenotype</td>
<td>Childhood ataxia with CNS hypomyelination</td>
</tr>
<tr>
<td></td>
<td>progressive loss of myelin in the CNS,</td>
</tr>
<tr>
<td></td>
<td>leading to neurological motor and cognitive</td>
</tr>
<tr>
<td></td>
<td>deficits</td>
</tr>
<tr>
<td>Mutation</td>
<td>R136H</td>
</tr>
<tr>
<td>Mouse Knockout</td>
<td>decrease of mean corpuscle volume</td>
</tr>
<tr>
<td></td>
<td>decrease of corpuscular hemoglobin</td>
</tr>
<tr>
<td></td>
<td>concentration</td>
</tr>
<tr>
<td></td>
<td>decreased circulating glucose level</td>
</tr>
<tr>
<td></td>
<td>decreased heart weight</td>
</tr>
<tr>
<td>Mouse Knockin</td>
<td>No significant effects on global protein</td>
</tr>
<tr>
<td></td>
<td>synthesis</td>
</tr>
<tr>
<td></td>
<td>Under normal conditions, the mutant mice</td>
</tr>
<tr>
<td></td>
<td>exhibited a normal life span and did not</td>
</tr>
<tr>
<td></td>
<td>develop severe clinical symptoms</td>
</tr>
<tr>
<td>Zebrafish phenotype</td>
<td>cerebellum ataxia (3 dpf)</td>
</tr>
<tr>
<td></td>
<td>microcephaly (3 dpf)</td>
</tr>
</tbody>
</table>
Human *EIF2B5* p.R136H is a CPD

![Graph showing area of the cerebellum (µm²) for different conditions: Uninjected, 3ng sbMO, 3ng sbMO + mouse wt RNA, 3ng sbMO + human wt RNA, 3ng sbMO + mouse R132H RNA, 3ng sbMO + human R136H RNA.](image)

- **Human**:
  - Control
  - 3ng sbMO + wt
  - 3ng sbMO + R136H

- **Mouse**:
  - 3ng sbMO + wt
  - 3ng sbMO + R132H
Back to TTC21B

P209L is hypomorphic in the human context

P209L: drives isolated renal phenotypes in ciliopathies

Human TTC21B-P209L protein results in reduced cilia length

Hypothesis: \( Ttc21B^{P209L/P209L} \) mice should be viable

\( Ttc21b \) is mutated in the alien mouse: lethal by E18.5


\( Ttc21b^{P209L/P209L} \) is also embryonic lethal

Unpublished.
Assay of cis-compensatory effects in human vs. mouse

<table>
<thead>
<tr>
<th>Condition</th>
<th>Class I</th>
<th>Class II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uninjected</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>10ng ttc21b MO</td>
<td>42%</td>
<td></td>
</tr>
<tr>
<td>+ 100pg WT human RNA</td>
<td>27%</td>
<td></td>
</tr>
<tr>
<td>+ 100pg WT mouse RNA</td>
<td>23%</td>
<td></td>
</tr>
<tr>
<td>10ng ttc21b MO</td>
<td>27%</td>
<td></td>
</tr>
<tr>
<td>+ 100pg P209L human RNA</td>
<td>38%</td>
<td></td>
</tr>
<tr>
<td>+ 100pg P209L mouse RNA</td>
<td>n.s.</td>
<td></td>
</tr>
</tbody>
</table>

P-value
vs uninjected: ****
vs MO: * ****

n=108 n=106 n=111 n=98 n=97 n=104
Fitness-independent CPDs?

A53T transgene $\alpha$-synuclein

Lee et al, PNAS 2002
A-syn structure and variants from the *felis catus* genome

- **Amphipatic domain**
- **Autosomal dominant PD**
- **Highly hydrophobic domain**
- **Percursor of the non amyloidogenic componente of NAC**
- **Promotes protein aggregation**
- **Acidic character**
- **Anti-amyloidogenic properties**
Testing for CPD in the most common PD-causing allele
Big Picture

- Allele effect informs burden
- Allele burden informs expressivity
- Not all deleterious alleles are deleterious
- Must study effect of alleles in cognate context

Keep studying the same patients with new technologies
Thank you