Understanding The Genetics of Diamond Blackfan Anemia

Jason Farrar, MD
jefarrar@uams.edu
About Me

• Assistant Professor of Pediatrics at University of Arkansas for Medical Sciences & Arkansas Children’s Hospital
• I’ve studied DBA genetics most of my professional life but I’m a hematologist not a geneticist
• This is a new talk
• I’ve never given a professional presentation in flip-flops before today
What Do We Mean By “Genetic Basis”?  

- Behavior or trait is passed to progeny by way of genes  
  - cell to cell (somatic or germline changes)  
  - Parent to child (germline)  
- How do we know DBA has a genetic basis?
An Autosomal Dominant Pedigree

I

II

III

IV
# Inheritance Patterns

- **Autosomal Dominant:** One gene alteration is sufficient to cause disease (*most DBA*)

- **Autosomal Recessive:** Both copies (maternal and paternal) of gene must be altered to manifest

- **X-linked Recessive:** Disorder generally only seen in males (*DBA with a couple of specific gene alterations*)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Aa    |     |     | 50% chance of offspring with the disorder

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aa</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aa</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| aa    |     |     | 25% chance of offspring with the disorder
The Human Genome
Our ‘Recipe Book’

- 46 total volumes (chromosomes)
  - Occur in 23 pairs
    - 2 copies of each volume (2 number 1’s etc.)
    - Last set = X,Y
- A total of 22,000 recipes (genes)
- Each recipe (gene) is composed of words (nucleotide triplets)
- Each nucleotide (A,G,C,T) is a letter
The Human Genome

- Organized as 23 chromosome pairs
  - Think of a 46 volume collection of 22,000 recipes.

- Cytogenetic (karyotypes) studies
  - Tell you if the right number of volumes are there, and if – from the outside – the books are intact
Karyotypes Are Almost Always Normal in DBA
(but this is how DBA gene discovery got started)

800 – 1000 bands
~ 25 genes / band
1 band ~ 3 – 5 Mb

Gustavsson et al, J Med Genet, 1997:34, 779-82
### What Are the DBA Genes?

<table>
<thead>
<tr>
<th>Small Ribosomal Subunit</th>
<th>Large Ribosomal Subunit</th>
<th>Non-Ribosomal Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RPS7</strong></td>
<td><strong>RPL5</strong></td>
<td><strong>GATA1</strong>*</td>
</tr>
<tr>
<td><strong>RPS10</strong></td>
<td><strong>RPL11</strong></td>
<td><strong>TSR2</strong>*</td>
</tr>
<tr>
<td><strong>RPS15A</strong></td>
<td><strong>RPL15</strong></td>
<td>* X-linked recessive</td>
</tr>
<tr>
<td><strong>RPS17</strong></td>
<td><strong>RPL18</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS19</strong></td>
<td><strong>RPL26</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS20</strong></td>
<td><strong>RPL27</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS24</strong></td>
<td><strong>RPL31</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS26</strong></td>
<td><strong>RPL35</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS27</strong></td>
<td><strong>RPL35A</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS28</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RPS29</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Possible:
- **RPS15**
- **RPL19**
The Ribosome: Our Cellular Chef

“recipe” (gene) is read here
“dish” (protein) produced here

Small (40S) Subunit
-33 proteins (RPS genes)
-1 ribosomal RNA

Large (60S) Subunit
-47 proteins (RPL genes)
-3 ribosomal RNAs

Front
Side
How Are DBA Genes Abnormal?

- **Pages (chromosomes) stuck together**
  - Translocations; very, very rare & almost always associated with DBA along with unusual additional findings
- **Recipes (genes) missing**
  - Mom’s version is there, but not Dad’s (or vice versa)
  - Gene deletions; can be small (e.g., just a DBA gene or part of the gene) or large (e.g., along with all the neighboring recipes)
  - Common
- **Recipes (genes) illegible**
  - “Nonsense” mutations effectively make the rest of the recipe unusable
  - End-effect is similar to gene deletions
  - Very common
- **Recipes (genes) altered**
  - “Missense” mutations
  - Changing the “word” specifying aspartic acid with one for proline is a bit like substituting cayenne pepper for flour.
  - Common
How Do We Determine Which Gene is Responsible in Your Child?

• **Copy Number Detection**
  – Make sure that all of the recipes that should be there are actually there
  – Genome-wide: SNP microarray
  – Targeted: MLPA (multiplex probe dependent ligation and amplification); quantitative PCR

• **Sequencing**
  – Reading and comparing the recipes to what we think is a standard
  – Detects altered words (nucleotides) that lead to either missense or nonsense changes
    • Single gene
    • Gene panels
    • Whole exome (all the coding genes (recipes))
    • Whole genome
Chromosomal (or SNP-) Microarray

• Analogous to opening up all the volumes and checking in multiple (several million places out of 3.2 billion) to see if the right letters are in each place
Single Nucleotide Polymorphisms (SNPs)

- SNPs – variations in single nucleotides scattered (non-randomly) across the genome
- Within the approximate 3.2 billion nucleotide human genome sequence, there are an estimated 10 million SNPs
- The identity of these SNPs, along with other less common forms of genetic variation, are essentially responsible for making us genetically unique
Single Nucleotide Chromosomal Microarray (SNP CMA)

• Using single nucleotide polymorphisms (SNPs) as probes
  – Usually 1-2 million SNPs on a chip
• Detects broad range of deletions of various sizes on the chromosomes
Array-based Copy Number Detection

The pattern in DBA essentially always looks like this.

Array based (genomewide) methods are used to diagnose many, many genetic conditions and so are widely available, relatively inexpensive, and easily ordered in most centers.
Deletions of *RPS19* and *RPS26*


- 2 copies
- 1 copy

**RPS19**

**RPS26**
Targeted Copy Number Detection

• Spot checking the recipe book, but only at specific genes rather than across the genome

• Multiple methods
  – MLPA (most widely available in DBA panels in Europe and North America)
  – quantitative PCR

• More sensitive than array methods for identifying very small deletions
  – Implication: you can have a normal SNP-array or CGH study and still have a tiny deletion that can’t be found with sequencing
Gene Sequencing

- Chromosomal microarrays are very helpful in identifying deletions of specific genes
- However, they won’t detect small changes, point mutations, etc
- Often the only method to make a diagnosis is to sequence the gene
- Historically it has been very expensive and time consuming to sequence large genes
Gene Sequencing

- Sequencing a **gene** is like going into one of the 46 **volumes**, finding one **recipe** (chapter) and reading it in its entirety.
Sanger Sequencing

**RPL35A**

Both a C and T at this position
-Mutation “T” changes the normal word from arginine (CGA) to stop (TGA). This makes an “illegible” recipe so nothing can be made from it.
NextGen Sequencing

• High throughput sequencing was faster than the generation before based on advances in automation and multi-sample processing.
  – However, the process still involved longer runs of sequencing, but doing each sequence just once.
• "Next Gen" sequencers do much shorter sequences but do thousands of overlapping copies.
  – Massively parallel sequencing
Whole Exome Sequencing (WES)

• The exome is the part of the genome that codes for proteins
• The exome is made up of 22,000 genes that comprise about 2% of the entire genome
• WES is like going through all of the recipe books and reading all 22,000 recipes letter by letter
Whole Exome Sequencing Data

- Letter changed from C to G, but no words are in this region, so this is not a disease-causing change, rather a common SNP.
Whole Exome Sequencing
Clinical Application

• Currently whole exome sequencing is available as a clinical test.
• Pricing is decreasing rapidly as technology becomes more ubiquitous
  – Costs down to $4000 for singleton cases, $~6000 for trios
  – Turn around times of 3-4 months
Panel-Based NGS Sequencing

- Like exome sequencing, enrich/sequence a tiny fraction of the genome, but only those known to cause DBA
  - Cheaper, faster, easier to interpret, easier to get insurance coverage than clinical WES
  - Many commercial tests can be ordered with concomitant deletion analysis
  - The test has to be specific for DBA; just ordering a “23 and Me” test won’t give the data you need!
Whole Genome Sequencing (WGS)

• As the name implies, sequencing the entire human genome
  – all 3.2 billion base pairs
  – x 2 copies (from mother & father)
  – x ~50-100 time at each letter
  – enormous amount of data!

• About 94% of the entire genome is ‘accessible’ (able to be uniquely sequenced and mapped)

• Reading all 22,000 recipes plus all sorts of extra pages inserted between the recipe pages in the books
  – Including the ‘junk DNA’

• Not currently used clinically, but is being used in some research studies
Why Is Establishing a Genetic Diagnosis Important?

- Confirmation of a clinical diagnosis of DBA
- To identify at risk family members (incomplete penetrance and variable expressivity)
  - Reproductive choices and family genetic counseling
- To identify suitable family members for stem cell transplantation (incomplete penetrance and variable expressivity, again...)
- For gene therapy opportunities (Dr. Bodine talk)
Challenges in DBA Genetics

• In only ~3/4 of cases, even with extensive genetic evaluation, can a genetic cause be identified
  – what are the additional genes?
  – Regulatory regions?
• Clinical testing panels keeping pace with research
### Panel Sequencing: GeneDx

#### Small Ribosomal Subunit
- RPS7
- RPS10
- RPS15A
- RPS17
- RPS19
- RPS20
- RPS24
- RPS26
- RPS27
- RPS28
- RPS29

#### Large Ribosomal Subunit
- RPL5
- RPL11
- RPL15
- RPL18
- RPL26
- RPL27
- RPL31
- RPL35
- RPL35A

#### Non-Ribosomal Protein
- GATA1*
- TSR2*

* X-linked recessive

**Possible**
- RPS15
- RPL19
### Panel Sequencing: Fulgent

<table>
<thead>
<tr>
<th>Small Ribosomal Subunit</th>
<th>Large Ribosomal Subunit</th>
<th>Non-Ribosomal Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RPS7</strong></td>
<td><strong>RPL5</strong></td>
<td><strong>GATA1</strong></td>
</tr>
<tr>
<td><strong>RPS10</strong></td>
<td><strong>RPL11</strong></td>
<td><strong>TSR2</strong></td>
</tr>
<tr>
<td><strong>RPS15A</strong></td>
<td><strong>RPL15</strong></td>
<td>* X-linked recessive</td>
</tr>
<tr>
<td><strong>RPS17</strong></td>
<td><strong>RPL18</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS19</strong></td>
<td><strong>RPL26</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS20</strong></td>
<td><strong>RPL27</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS24</strong></td>
<td><strong>RPL31</strong></td>
<td><strong>Possible</strong></td>
</tr>
<tr>
<td><strong>RPS26</strong></td>
<td><strong>RPL35</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS27</strong></td>
<td><strong>RPL35A</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS28</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RPS29</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* X-linked recessive
# Panel Sequencing: Ambry

<table>
<thead>
<tr>
<th>Small Ribosomal Subunit</th>
<th>Large Ribosomal Subunit</th>
<th>Non-Ribosomal Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RPS7</strong></td>
<td><strong>RPL5</strong></td>
<td><strong>GATA1</strong></td>
</tr>
<tr>
<td><strong>RPS10</strong></td>
<td><strong>RPL11</strong></td>
<td><strong>TSR2</strong></td>
</tr>
<tr>
<td><strong>RPS15A</strong></td>
<td><strong>RPL15</strong></td>
<td>* X-linked recessive</td>
</tr>
<tr>
<td><strong>RPS17</strong></td>
<td><strong>RPL18</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS19</strong></td>
<td><strong>RPL26</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS20</strong></td>
<td><strong>RPL27</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS24</strong></td>
<td><strong>RPL31</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS26</strong></td>
<td><strong>RPL35</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS27</strong></td>
<td><strong>RPL35A</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS28</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RPS29</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Possible

- **RPS15**
- **RPL19**
## Panel Sequencing: MarrowSeq

<table>
<thead>
<tr>
<th>Small Ribosomal Subunit</th>
<th>Large Ribosomal Subunit</th>
<th>Non-Ribosomal Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RPS7</strong></td>
<td><strong>RPL5</strong></td>
<td><strong>GATA1</strong>*</td>
</tr>
<tr>
<td><strong>RPS10</strong></td>
<td><strong>RPL11</strong></td>
<td><strong>TSR2</strong>*</td>
</tr>
<tr>
<td><strong>RPS15A</strong></td>
<td><strong>RPL15</strong></td>
<td>* X-linked recessive</td>
</tr>
<tr>
<td><strong>RPS17</strong></td>
<td><strong>RPL18</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS19</strong></td>
<td><strong>RPL26</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS20</strong></td>
<td><strong>RPL27</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS24</strong></td>
<td><strong>RPL31</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS26</strong></td>
<td><strong>RPL35</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS27</strong></td>
<td><strong>RPL35A</strong></td>
<td></td>
</tr>
<tr>
<td><strong>RPS28</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RPS29</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Possible:

- **RPS15**
- **RPL19**
Challenges in DBA Genetics

- VUS: variants of undetermined significance
  - Deletions (missing pages) and nonsense sequence changes (illegible recipes) in the known DBA genes are relatively straightforward to interpret
  - Some missense (word) changes can also be easily interpreted—either because they recur again and again in DBA or because enough is known about the gene to make a strong inference (e.g., red pepper in place of flour will not make edible bread!)
  - Other missense changes can be hard to interpret:
    - omitting salt from spaghetti; substituting butter for vegetable oil; using mustard as a base for BBQ sauce can all potentially yield a palatable dish.
Classifying Missense Variants – \textit{RPS19}

- Nonsense mutations across gene
- Hotspots for missense mutations

Willig et al. \textit{Blood} 94(12): 2000
Pathogenic or not?

Yes, this is a well-described change in a mutational hot-spot. Seen in many DBA patients.

Test(s) Requested: Diamond-Blackfan Anemia Panel
Genes Evaluated: RPS19, RPL5, RPS26, RPL11, RPL35A, RPS10, RPS24, RPS17, RPS7, RPL15, RPS29, RPL26, GATA1

Result:

<table>
<thead>
<tr>
<th>Gene</th>
<th>Coding DNA</th>
<th>Variant</th>
<th>Zygosity</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPS19</td>
<td>c.184 C&gt;T</td>
<td>p.Arg62Trp (R62W)</td>
<td>Heterozygous</td>
<td>Pathogenic Variant</td>
</tr>
</tbody>
</table>

Concurrent targeted array CGH analysis with exon-level resolution (ExonArrayDx) did not reveal a deletion or duplication of the included genes.

Interpretation:

This individual is heterozygous for a pathogenic variant in the RPS19 gene, consistent with the diagnosis of Diamond-Blackfan Anemia.

Pathogenic or not?

Yes, this is a well-described change in a mutational hot-spot. Seen in many DBA patients.

Test(s) Requested: Diamond-Blackfan Anemia Panel
Genes Evaluated: RPS19, RPL5, RPS26, RPL11, RPL35A
Result: POSITIVE

<table>
<thead>
<tr>
<th>Gene</th>
<th>Coding DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPS19</td>
<td>c.184 C&gt;T</td>
</tr>
</tbody>
</table>

Concurrent targeted array CGH analysis w Sbg2 codons duplication of the included genes.

Interpretation: This individual is heterozygous for a p Diamond-Blackfan Anemia.

Pathogenic or not?

• ????: unclear
  – We can make educated guesses (sometimes)
  – In some cases we can perform functional studies (*research based testing)

DNA sequence analysis of the RPL11 gene demonstrated a three base pair deletion in exon 4, c.286_298del. This in-frame deletion is predicted to result in the deletion of an amino acid residue, p.Phe99del. This sequence change does not appear to have been previously described in patients with RPL11-related disorders and has also not been described as a known benign sequence change in the RPL11 gene. The functional significance of this sequence change is not known at present and its contribution to this patient’s disease phenotype cannot definitively be determined. Heterozygous pathogenic variants in RPL11 have been identified in patients with Diamond-Blackfan anemia [OMIM#612562].

• Mutations in RPL11 can be de novo or inherited. Incomplete penetrance may be observed in some cases. Testing of this patient's parents and affected family members, and close clinical correlation is suggested to help clarify the nature of this sequence change.
Challenges in DBA Genetics

- Variable expression and incomplete penetrance
- Who has DBA?
Variable Expression and Incomplete Penetrance
Variable Expression and Incomplete Penetrance

• Variable Expression: you can carry a DBA gene **without severe anemia** but still have:
  – A child who has severe anemia
  – DBA associated physical malformations
  – Risk of cancer (Dr. Vlachos’ talk)

• Incomplete penetrance: you can carry a DBA mutation in your genes, but appear **completely normal**. Same implications as above

• Contextual aspects to diagnosis of DBA (Dr. Lipton)
  – Is it “DBA” with large red cells, no anemia?
  – Is it “DBA” with a VSD or small thumbs but normal CBC?
Challenges in DBA Genetics

• Genotype/Phenotype relationships
  – Gold standard here would be to predict based on which gene or which specific mutation who is going to respond to steroids, enter a clinical remission, or is at especially high risk of a cancer
  • We’re not there yet
    – Oral-facial malformations: *RPL11, RPL5, RPS26*
    – Thumb: *RPL11*
    – Neutropenia: *RPL35A*
Thank You!
X-Linked Recessive Pedigree