Translating Next Generation Sequencing into the Clinical Diagnostic Arena

Karl V. Voelkerding M.D.
Associate Professor of Pathology
University of Utah, Salt Lake City
Medical Director for Advanced Technology
ARUP Laboratories
voelkek@aruplab.com

CHI XGEN Congress
March 18, 2010
Presentation Outline

• Genetic Diagnostic Targets for NGS

• Hypertrophic Cardiomyopathy Experience

• Perspective on Path Forward
Presentation Outline

- Genetic Diagnostic Targets for NGS
- Hypertrophic Cardiomyopathy Experience
- Perspective on Path Forward
Targeted Resequencing for Diagnostics

Multiple Genes ↔ Multiple Mutations

Overlapping Phenotype
Targeted Resequencing for Diagnostics

Cardiomyopathies

- Hypertrophic
- Dilated
- Channelopathies

10-30 Genes
Targeted Resequencing for Diagnostics

Cardiomyopathies

- Hypertrophic
- Dilated
- Channelopathies

10-30 Genes

Mitochondrial Disorders

- Mitochondrial Genome (Heteroplasmcy)
- Nuclear Genes
Targeted Resequencing for Diagnostics

Cardiomyopathies
- Hypertrophic
- Dilated
- Channelopathies (10-30 Genes)

Mitochondrial Disorders
- Mitochondrial Genome (Heteroplasmy)
- Nuclear Genes

X-Linked Mental Retardation (~80 Genes)
Enrichment Strategies

Gene Panel

Amplification Based
- PCR or LR-PCR
- RainDance ePCR
- Fluidigm
- MIP/Selector

Array Based
- Solid Surface
- In Solution

Next Generation Sequencing
Presentation Outline

• Genetic Diagnostic Targets for NGS

• Hypertrophic Cardiomyopathy Experience

• Perspective on Path Forward
Hypertrophic Cardiomyopathy

Normal

Right Ventricle

Left Ventricle

Hypertrophic Cardiomyopathy


Hypertrophic

Right Ventricle

Left Ventricle

> 13mm Wall Thickness
Clinical Features of HCM

• Childhood to Adult Presentation

• Symptoms

  Dyspnea
  Angina
  Arrhythmias/Syncope
Clinical Features of HCM

- Childhood to Adult Presentation

- Symptoms
  - Dyspnea
  - Angina
  - Arrhythmias/Syncope
  - Sudden Death
Differential Diagnosis

• Primary Hypertrophic Cardiomyopathy

• Secondary Left Ventricular Hypertrophy
  – Congenital syndromes
  – Metabolic (e.g., glycogen storage disease)
  – Amyloidosis
Treatment

- Beta and calcium channel blockers

- Antiarrhythmics and cardioversion
  - Implantable defibrillators

- Surgical myectomy to increase outflow

- Transplantation
Genetics of Primary HCM

• Autosomal Dominant Inheritance

• Incidence of 1 in 500-1,000

• Variable Penetrance
Cardiac Sarcomere

Myofibril

Sarcomere

Kamisago et al. NEJM 343(23):1688
## Mutational Spectrum of HCM

<table>
<thead>
<tr>
<th>Protein</th>
<th>Gene</th>
<th>Mutations</th>
<th>Gene Size bp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myosin, heavy chain 7</td>
<td>MYH7</td>
<td>193</td>
<td>32,628</td>
</tr>
<tr>
<td>Myosin binding protein C</td>
<td>MYBPC3</td>
<td>138</td>
<td>28,280</td>
</tr>
<tr>
<td>Troponin T type 2</td>
<td>TNNT2</td>
<td>33</td>
<td>25,673</td>
</tr>
<tr>
<td>Troponin I type 3</td>
<td>TNNI3</td>
<td>32</td>
<td>12,963</td>
</tr>
<tr>
<td>Cysteine and glycine-rich protein 3</td>
<td>CSRP3</td>
<td>12</td>
<td>27,024</td>
</tr>
<tr>
<td>Tropomyosin 1, α</td>
<td>TPM1</td>
<td>11</td>
<td>36,274</td>
</tr>
<tr>
<td>Myosin, light chain 2</td>
<td>MYL2</td>
<td>10</td>
<td>16,758</td>
</tr>
<tr>
<td>Actin</td>
<td>ACTC</td>
<td>7</td>
<td>14,631</td>
</tr>
<tr>
<td>Myosin, light chain 3</td>
<td>MYL3</td>
<td>5</td>
<td>12,617</td>
</tr>
<tr>
<td>Protein kinase, AMP-activated, γ2</td>
<td>PRKAG2</td>
<td>4</td>
<td>328,114</td>
</tr>
<tr>
<td>Phospholamban</td>
<td>PLN</td>
<td>2</td>
<td>19,112</td>
</tr>
<tr>
<td>Troponin C type 1</td>
<td>TNNC1</td>
<td>1</td>
<td>9,041</td>
</tr>
<tr>
<td>Titin</td>
<td>TTN</td>
<td>2</td>
<td>281,434</td>
</tr>
<tr>
<td>Myosin, heavy chain 6</td>
<td>MYH6</td>
<td>2</td>
<td>32,628</td>
</tr>
<tr>
<td>Titin-cap</td>
<td>TCAP</td>
<td>2</td>
<td>9,361</td>
</tr>
<tr>
<td>Caveolin 3</td>
<td>CAV3</td>
<td>1</td>
<td>20,199</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>455</strong></td>
<td><strong>906,737</strong></td>
</tr>
</tbody>
</table>
HCM Gene Distribution

1. TNNT2
2. TTN
3. CAV3
4. MYL3
5. TNNC1
6. PLN
7. PRKAG2
8. CSR3
9. CSRP3
10. MYBPC3
11. MYL2
12. MYH6
13. MYH7
14. ACTC1
15. TPM1
16. TNC1
17. TCAP
18. TNI3
19. MYH6
20. MYH7
21. ACTC1
22. TPM1
23. TNI3
24. TPX1
25. TCAP
Hypertrophic Cardiomyopathy

Value of Genetic Testing

• Confirm Genetic Cause in Patient

• Family Counseling
  – Specific Mutation Screening
Molecular Diagnostic Options

- Individual Gene Analysis
- Multi-Gene Resequencing Array
- Next Generation Sequencing
Hypertrophic Cardiomyopathy

Phase One Project

Single Human Genomic DNA – Normal Cardiac Hx
Fourteen “Full Length” Cardiomyopathy Genes
Compare Roche 454 and Illumina Technologies
Hypertrophic Cardiomyopathy Phase One

Gene Enrichment by Long Range PCR

344,082 bp

Roche 454 Sequencing

Illumina GA Sequencing
LR-PCR of Cardiomyopathy Genes

Pool Amplicons → Fragment → NGS Library Prep

5 Kb
# Sequencing Run Parameters

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Read Length</th>
<th>Base Quality (avg)</th>
<th>Aligned Reads</th>
<th>Coverage (avg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illumina</td>
<td>36</td>
<td>31</td>
<td>7,081,505 (89.9%)</td>
<td>798</td>
</tr>
<tr>
<td>Roche 454</td>
<td>235 (avg)</td>
<td>32</td>
<td>265,155 (94.1%)</td>
<td>198</td>
</tr>
</tbody>
</table>

**Illumina Single Lane**

**Roche 454 Single Plate**
Variant Plot - *MYH7* Region A

Reference Sequence Position

Read Percentage

Heterozygote

Homozygote

Exon

Illumina

Intron

454
Variant Plot - *MYH7* Region B

**Initial Variant Selection Criteria**

- Coverage $\geq 30X$
- Read Percentage $\geq 20\%$

**Variant?**

Reference Sequence Position

Read Percentage

- Exon
- Intron
- Illumina
- 454
## Exon Variant Discovery

<table>
<thead>
<tr>
<th></th>
<th>Illumina</th>
<th>Roche 454</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met selection criteria</td>
<td>31</td>
<td>40</td>
</tr>
<tr>
<td>Platform concordants</td>
<td>27</td>
<td>27</td>
</tr>
</tbody>
</table>
Exon Variant Discovery

<table>
<thead>
<tr>
<th></th>
<th>Illumina</th>
<th>Roche 454</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met selection criteria</td>
<td>31</td>
<td>40</td>
</tr>
<tr>
<td>Platform concordants</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sanger Verified</td>
</tr>
</tbody>
</table>
Concordant Variant g.5871488T>C in ACTC1

Roche 454 50%

Illumina 52%
# Exon Variant Discovery

<table>
<thead>
<tr>
<th></th>
<th>Illumina</th>
<th>Roche 454</th>
<th>Sanger Verified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met selection criteria</td>
<td>31</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Platform concordants</td>
<td>27</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Discordants</td>
<td>4</td>
<td>13</td>
<td></td>
</tr>
</tbody>
</table>
Exon Variant Discovery

<table>
<thead>
<tr>
<th></th>
<th>Illumina</th>
<th>Roche 454</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met selection criteria</td>
<td>31</td>
<td>40</td>
</tr>
<tr>
<td>Platform concordants</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Discordants</td>
<td>4</td>
<td>13</td>
</tr>
</tbody>
</table>

Sanger Verified

- **Illumina and GS FLX**: Coverage < 30 fold → 5
- **Illumina**: Misalignments - Cross homology → 2
- **Roche 454**: Homopolymer sequencing errors → 10
Homopolymer Error

g.11968221delC

PRKAG2

Roche 454
32%

Illumina
2%

Poly C_7
Misalignment Error

\[ g.4892783C>T \]

**MYH7**

Roche 454

0%

Illumina

22%

**MYH7**

**MYH6**
## Exon Variant Discovery

<table>
<thead>
<tr>
<th>Met selection criteria</th>
<th>Illumina</th>
<th>Roche 454</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform concordants</td>
<td>31</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>(27)</td>
<td>= 27</td>
</tr>
<tr>
<td>&lt;30 fold coverage</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Sanger Verified

- **32 Variants**
### Exon Variant Discovery

<table>
<thead>
<tr>
<th></th>
<th>Illumina</th>
<th>Roche 454</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met selection criteria</td>
<td>31</td>
<td>40</td>
</tr>
<tr>
<td>Platform concordants</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>&lt;30 fold coverage</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

32 Variants

<table>
<thead>
<tr>
<th></th>
<th>Illumina</th>
<th>Roche 454</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterozygous Variants Read %</td>
<td>21-67%</td>
<td>33-86%</td>
</tr>
<tr>
<td>Homozygous Variants Read %</td>
<td>80-94%</td>
<td>100%</td>
</tr>
</tbody>
</table>
Exon Variant Functional Evaluation

<table>
<thead>
<tr>
<th></th>
<th>Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Variants</td>
<td>32</td>
</tr>
<tr>
<td>Synonymous</td>
<td>16</td>
</tr>
<tr>
<td>Nonsynonymous</td>
<td>2 *</td>
</tr>
<tr>
<td>Untranslated regions</td>
<td>14</td>
</tr>
<tr>
<td>In Databases or Literature</td>
<td>32</td>
</tr>
</tbody>
</table>

Comment:

* Nonsynonymous variants not associated with HCM
Phase One Lessons

- Next Generation Sequencing Promising ....
  - Front End Technically Complex
  - Bioinformatics Time Intensive

- Sanger Confirmation Essential
Phase One Lessons

• Next Generation Sequencing Promising ....
  – Front End Technically Complex
  – Bioinformatics Time Intensive

• Sanger Confirmation Essential

False Positives: Address by Platform/Bioinformatic Improvements

False Negatives: None in ~ 20,000 Bases Sanger Sequenced
Current Efforts

• Scaling Up for Clinical Translation
• Focus on Enrichment
• Platform Agnostic
Enrichment Strategies

Gene Panel

Amplification Based
- PCR or LR-PCR
- RainDance ePCR
- Fluidigm
- MIP/Selector

Array Based
- Solid Surface
- In Solution

Next Generation Sequencing
Hypertrophic Cardiomyopathy Phase Two

1. Genomic DNA
2. Illumina Library Prep
3. Hybridize Library to Cardiac Gene Panel Capture Array
4. Elute Enriched Fraction
5. Illumina Sequencing – 76 base
Hypertrophic Cardiomyopathy Phase Two

Capture and Coverage Example: *ACTC1* Gene

Gene Panel Variants Identified: Control DNA
**Febit Preliminary Data**

<table>
<thead>
<tr>
<th>Gene</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MYH7</td>
<td></td>
</tr>
<tr>
<td>ACTC1</td>
<td></td>
</tr>
<tr>
<td>TNNT2</td>
<td></td>
</tr>
<tr>
<td>TPM1</td>
<td></td>
</tr>
<tr>
<td>CSRP3</td>
<td></td>
</tr>
<tr>
<td>MYBPC3</td>
<td></td>
</tr>
<tr>
<td>MYL2</td>
<td></td>
</tr>
<tr>
<td>MYL3</td>
<td></td>
</tr>
<tr>
<td>MYL3</td>
<td></td>
</tr>
<tr>
<td>TNNI3</td>
<td></td>
</tr>
</tbody>
</table>
# Febit Preliminary Data

<table>
<thead>
<tr>
<th>Gene</th>
<th>LR-PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MYH7</td>
<td>7</td>
</tr>
<tr>
<td>ACTC1</td>
<td>5</td>
</tr>
<tr>
<td>TNNT2</td>
<td>1</td>
</tr>
<tr>
<td>TPM1</td>
<td>1</td>
</tr>
<tr>
<td>CSRP3</td>
<td>-</td>
</tr>
<tr>
<td>MYBPC3</td>
<td>-</td>
</tr>
<tr>
<td>MYL2</td>
<td>-</td>
</tr>
<tr>
<td>MYL3</td>
<td>-</td>
</tr>
<tr>
<td>TNNI3</td>
<td>-</td>
</tr>
</tbody>
</table>
## Febit Preliminary Data

<table>
<thead>
<tr>
<th>Gene</th>
<th>LR-PCR</th>
<th>Febit</th>
</tr>
</thead>
<tbody>
<tr>
<td>MYH7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>ACTC1</td>
<td>5</td>
<td>17*</td>
</tr>
<tr>
<td>TNNT2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>TPM1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>CSRP3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MYBPC3</td>
<td>-</td>
<td>2*</td>
</tr>
<tr>
<td>MYL2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MYL3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TNNI3</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*Co-Capture*
<table>
<thead>
<tr>
<th>Gene</th>
<th>LR-PCR</th>
<th>Febit</th>
<th>Bioinformatics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MYH7</strong></td>
<td>7</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td><strong>ACTC1</strong></td>
<td>5</td>
<td>17*</td>
<td>7</td>
</tr>
<tr>
<td><strong>TNNT2</strong></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>TPM1</strong></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>CSRP3</strong></td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td><strong>MYBPC3</strong></td>
<td>-</td>
<td>2*</td>
<td></td>
</tr>
<tr>
<td><strong>MYL2</strong></td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td><strong>MYL3</strong></td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td><strong>TNNT13</strong></td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

*Co-Capture

Sanger Confirmed
# Febit Preliminary Data

<table>
<thead>
<tr>
<th>Gene</th>
<th>LR-PCR</th>
<th>Febit</th>
<th>Bioinformatics</th>
</tr>
</thead>
<tbody>
<tr>
<td>MYH7</td>
<td>7</td>
<td>7</td>
<td>MYH6*</td>
</tr>
<tr>
<td>ACTC1</td>
<td>5</td>
<td>17*</td>
<td>7</td>
</tr>
<tr>
<td>TNNT2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>TPM1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>CSRP3</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>MYBPC3</td>
<td>-</td>
<td>2*</td>
<td></td>
</tr>
<tr>
<td>MYL2</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>MYL3</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>TNNI3</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

* Co-Capture

Sanger Confirmed
Phase Two Lessons

• Febit Array Capture
  – Reduces Technically Complex
  – Co-Capture Problematic
Presentation Outline

- Genetic Diagnostic Targets for NGS
- Hypertrophic Cardiomyopathy Experience
- Perspective on Path Forward
Targeted Resequencing for Diagnostics

Cardiomyopathies
- Hypertrophic
- Dilated
- Channelopathies
  - 10-30 Genes

Mitochondrial Disorders
- Mitochondrial Genome (Heteroplasmy)
- Nuclear Genes

X-Linked Mental Retardation
  - ~ 80 Genes
Targeted Resequencing for Diagnostics

- Enrichment
- Read Lengths
- Bioinformatics

- Specificity
- Amplification
- Target Dependent
- Customize
Acknowledgements

**Advanced Technology Group**
- Shale Dames
- Jacob Durtschi
- Jack Stephens
- Katherine Geiersbach

**Huntsman Cancer Institute**
- Brad Cairns
- David Nix
- Brian Dalley

**University of Colorado**
- David Pollack
- Todd Castoe
- Alex Poole

**ARUP Institute for Clinical and Experimental Pathology**