Next Generation Sequencing in Leukemia: From Discovery to Clinical Care

Keyur Pravinchandra Patel MD PhD
(kppatel@mdanderson.org)
Medical Director, Molecular Diagnostics Laboratory
(Hematologic Malignancies)
Assistant Professor,
Department of Hematopathology,
Division of Pathology and Laboratory Medicine
The University of Texas M.D. Anderson Cancer Center

No conflict of interest to disclose
The Clinical Molecular Diagnostics Laboratory (MDL)
Overview

- NGS as a genotyping platform
- Implementation of NGS in clinical diagnostics
- NGS as a discovery tool in clinical care
- Future directions
Next Generation Sequencing for Clinical Diagnostics-Principles and Application to Targeted Resequencing for Hypertrophic Cardiomyopathy

A Paper from the 2009 William Beaumont Hospital Symposium on Molecular Pathology

Karl V. Voelkerding,*† Shale Dames,† and Jacob D. Durschi†

Next generation sequencing (NGS) refers to high-throughput sequencing technologies that have emerged...
Next Generation Sequencing as a Discovery Tool

Recurring Mutations Found by Sequencing an Acute Myeloid Leukemia Genome
Elaine R. Mardis, Ph.D., Li Ding, Ph.D., David J. Dooling, Ph.D.,
NEJM 2009, 361;1058-66

DNMT3A Mutations in Acute Myeloid Leukemia
Timothy J. Ley, M.D., Li Ding, Ph.D., Matthew J. Walter, M.D.,
NEJM 2010, 363;2424-33

BRAF Mutations in Hairy-Cell Leukemia
Enrico Tiacci, M.D., Vladimir Trifonov, Ph.D., Gianluca Schiavoni, Ph.D.,
NEJM 2011, 364;2305-15

Oncogenic CSF3R Mutations in Chronic Neutrophilic Leukemia and Atypical CML
Julia E. Maxson, Ph.D., Jason Gotlib, M.D., Daniel A. Pollyea, M.D.,
NEJM 2013, 368;1781-90
NGS-Based Discoveries in Hematologic Malignancies

Mutations in Hematologic Malignancies

• **Integral part of diagnostic work-up** for myeloid neoplasms (AML, MDS, MPN, MDS/MPN)
• Diagnostic, prognostic and predictive value
• **Growing list of mutated genes with clinical utility:** \(NPM1, FLT3, RAS\) (\(KRAS, NRAS\)), \(KIT, CEBPA, WT1, IDH1, IDH2, DNMT3A, EZH2, JAK2, MPL\), several new genes in lymphoid tumors
• **New associations of known mutations:** \(BRAF\) in hairy cell leukemia
Hematology Diagnostics Timeline

- **Cytogenetics**
  - Discovery of t(9;22) CML
  - Discovery of t(15;17) AML
  - First identification of mutations in oncogenes
  - First identification of mutations in tumor suppressors
  - Completion of Human Genome Project

- **Sanger sequencing**

- **FISH**

- **GEP**
  - DLBCL subtype classification
  - Copy-number characterization of ALL, AML, CLL, CML, DLBCL, MM

- **SNP/CGH arrays**
  - Complete genomes of CLL, DLBCL, ETP-ALL, FL, HCL, MM, WM

- **Massively parallel sequencing**

---

### New Clinical Tests by Year

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ABL Seq</td>
<td>CEBPA</td>
<td>Luminex</td>
<td>KIT exon 17</td>
<td>MPL</td>
<td>IDH1</td>
<td>Leukemia Panel</td>
<td>CMS53</td>
<td></td>
</tr>
<tr>
<td>FIP1L1-PDGFR</td>
<td>RT JHF</td>
<td>ABL p.T315I</td>
<td>NPM1</td>
<td>IDH2</td>
<td>DNMT3A</td>
<td>ATM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH (CLL)</td>
<td>JAK2</td>
<td>TCRB</td>
<td>aCGH</td>
<td>TP53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Genotyping Platforms at MDL

Sanger Sequencing
- *CEBPA*
- Confirmation of *IDH1, IDH2, KIT* exon 17, *DNMT3A*

Pyrosequencing
- *KRAS, NRAS, JAK2, BRAF*

Capillary Gel Electrophoresis
- *FLT3, NPM1*

Real-Time PCR
- *KIT* p.D816V
- Translocations [(t(15:17), inv(16), t(9;22), t(8;21) etc.]

High Resolution Melting Analysis
- *IDH1, IDH2, DNMT3A, KIT* exon 17

Challenges:
Maintaining multiple platforms, workflows
High sample requirement (DNA and RNA)
Results not available simultaneously for clinical decision making
# Sequencing Technologies

<table>
<thead>
<tr>
<th>Platform</th>
<th>Sensitivity (for clinical use)</th>
<th>Sample Requirement</th>
<th>Multiplexing Capability</th>
<th>Throughput</th>
<th>Type of mutations</th>
<th>Quantitative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sanger</td>
<td>20%</td>
<td>High (ug)</td>
<td>None</td>
<td>Low</td>
<td>Point, indel</td>
<td>No</td>
</tr>
<tr>
<td>Pyro-</td>
<td>5%</td>
<td>Intermediate</td>
<td>None</td>
<td>Low</td>
<td>Point</td>
<td>Yes</td>
</tr>
<tr>
<td>Sequenom / ABI SNaPshot</td>
<td>10%</td>
<td>Intermediate</td>
<td>Intermediate</td>
<td>Medium</td>
<td>Point</td>
<td>Yes</td>
</tr>
<tr>
<td>NGS</td>
<td>5%</td>
<td>Low (ng)</td>
<td>High (amplicons and samples)</td>
<td>High</td>
<td>Point, indel, translocation, copy number change</td>
<td>Yes</td>
</tr>
</tbody>
</table>

NGS= next-generation sequencing
Next Generation Sequencing

- 2nd generation or massively parallel sequencing
- Simultaneous amplification and sequencing of a large number of amplicons

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Challenges for Clinical Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>High throughput</td>
<td>Selection of test platform</td>
</tr>
<tr>
<td>Better sensitivity</td>
<td>Evaluation of multi-gene panels</td>
</tr>
<tr>
<td>Efficient use of limited tissue</td>
<td>Establishment of assay characteristics</td>
</tr>
<tr>
<td>Consolidation of platforms</td>
<td>Validation on different sample types</td>
</tr>
<tr>
<td>Wider range of mutation detection</td>
<td>Results interpretation and reporting</td>
</tr>
<tr>
<td></td>
<td>Legal and ethical issues</td>
</tr>
<tr>
<td></td>
<td>Billing and compliance</td>
</tr>
</tbody>
</table>
Next Generation Sequencing Platforms
MDACC Clinical Laboratory

- Ion Torrent (Life Technologies)
  - Semiconductor based detection of pH change

- Illumina
  - Flow cell based, 4-color optical imaging of fluorescent labeled nucleotides

**PGM**
Small Gene Panels (1.5 – 2.0 Gbases/run)

**Proton**
Large gene panels and Whole Exome Sequencing (10 Gbases/run)

**MiSeq**
Small Gene Panels (1.5 – 2.0 Gbases/run)

**HiSeq**
For Whole Exome and Genome Sequencing (600 Gbases/run)
Human Genome: 3 Gb
Whole Exome: 30 Mb
Simplified NGS Workflow
Ion Torrent: Semiconductor based sequencing

Sequencing Wells
(3.5 \(\mu\)M Diameter)

Metal-Oxide Semiconductor
DNA immobilized on Ion Spheres
(2.0 \(\mu\)M Diameter)

314 Chip
(1.2 million)

316 Chip
(6 million)

318 Chip
(11 million)

Proton 1
(165 million)
Ion PGM Ampliseq Cancer Panel Work Flow

DNA extraction (3 hrs for PB&BM/15 hrs for FFPE)

Quantitation by Qubit Fluorometer

10ng Genomic DNA as Template

196 Amplicons (Hotspots in 46 genes)

Adapter & Barcode Ligation

ISP

Load on Ion Semiconductor Chip (1 hr)

Clonal Amplification on to ionspheres By Emulsion PCR (17 hrs)

3 Day Turn Around Time

Data Analysis (1 hr)

Sequencing 100bp 2.5 hr
Cluster Generation For Sequencing (Illumina)

- Hybridization of amplicon templates to a ‘lawn’ of oligos
- Synthesis of complementary strands
- Copying of immobilized strands by ‘Bridge Amplification’.
- Clonal amplification of a single strand into dense clusters
- Removal of strands attached by P7 adaptor & blocking of 3’ ends
- Sequencing by hybridizing sequencing primer
Illumina Platform: Sequencing By Synthesis: SBS

Optical Imaging of Fluorescence from Clusters

- DNA strands in the cluster act as templates for DNA polymerization
- Sequencing based synthesis using reversible terminator chemistry
- Incorporation of each fluorescently labeled nucleotide detected using optical imaging
- True paired end sequencing

Template Clusters in the Flowcell Lane
( ~ 1000 DNA strands/cluster)

(0.8 – 1.0 million clusters/mm²)
MiSeq TruSeq Cancer Panel Work Flow

DNA extraction (3 hrs for PB, BM)

Quantitation by Qubit Fluorometer

250ng Genomic DNA as Template

4 Day Turn Around Time

Sequencing and base calling (27 hrs)

Adapter & Barcode Ligation

Data Analysis (1 hr)

Cluster Generation for Sequencing (9 hrs)
Integrative Genome Browser (IGV, Broad Institute)

Chromosomal Position

Coverage Depth

Sequencing Reads

Gene

COSMIC
Mutation Detection by NGS

Point Mutation

- Sanger
- Conventional
- NGS

Insertion

- Capillary GE
- 4bp Ins
- NGS

Deletion

- Sanger
- NGS
Detection of Internal Tandem Duplications (Using Pindel)

FLT3 In-Tandem Duplication of 18 bp
Missing Pieces For Clinical Use

• Adequacy assessment for all amplicons
• Sample annotation system
• Easy visualization of results
• Inclusion or exclusion of variants for reporting
• Ability to monitor frequency of a specific mutation in our samples
• Conversion of base calls into standardized nomenclature
• Direct linking to genomic databases (COSMIC, dbSNP etc.)
• Internal annotation system for variants detected for future reference
• Generation of clinical reports
Variant Caller 2.0 Output (Ion PGM)

- Requires high performance servers for data analysis
- Variant caller file does not contain amino acid and coding sequence information

- Chromosome
- Genomic position
- Gene symbol
- Target id
- Type
- Ploidy
- Reference nucleotide
- Variant nucleotide
- Variant frequency
- P-value
- Coverage
- Reference coverage
- Variant coverage
- Hotspot id (COSMIC)
MiSeq Reporter Output

- Performs the data analysis
- Requires high performance computers
- Variant caller file contains amino acid and coding sequence correlates of mutations

- Sample id
- Sample name
- Manifest file
- Clusters PF
- Clusters aligned R1
- Clusters aligned R2
- Mismatch
- No call
- Coverage
- HetSNPs
- HomSNPs
- Insertions
- Deletions
MDACC Informatics Pipeline for Clinical Reporting

DAT Files Processing
- Base Calling
  - Output: BAM & FASTQ/SFF
- Alignment
  - Output: VCF
- MiSeq Reporter
  - Sample annotation
  - Suboptimal coverage filtering
  - Knowledgebase mapping (e.g., COSMIC)
  - Intersample comparison for germline/somatic variants
  - Population Analysis
  - Amplicon viewer with the integrated Genomics Viewer (IGV)

Courtesy: Mark Routbort, MD, PhD
Practical Challenges

• Re-engineering of workflow
• Reporting of larger scale genomic information
• Billing and compliance
Reporting of the Mutations

- Standardized nomenclature added to allow portability and inter-laboratory comparison
- **BRAF mutation analysis:**
  - Mutation detected in codon 600, exon 15 (GTG to GAG) of the BRAF gene that would change the encoded amino acid from Valine to Glutamate (p.Val600Glu)

![Diagram of mutation nomenclature](http://www.hgvs.org/mutnomen/)
Medico-Legal and Ethical Issues

- Reporting of all versus selected genes
- Mutations in unordered genes
- Germline variants
- Integration in clinical management
**Billing and Compliance**

- Billing based on AMA CPT codes

---

<table>
<thead>
<tr>
<th>Old Codes (pre-2013)</th>
<th>New Codes (2013)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Stacking codes</td>
<td>• Gene-specific codes</td>
</tr>
<tr>
<td>• Procedure based e.g. DNA extraction, sequencing, interpretation etc.</td>
<td>• Limited number of genes on the current code list</td>
</tr>
<tr>
<td>• Easy to implement for a new gene</td>
<td>• NO codes for gene panels</td>
</tr>
<tr>
<td></td>
<td>• Lengthy application process for new genes</td>
</tr>
</tbody>
</table>
III. Clinical Experience with NGS

• 53 gene panel live since October, 2012
  – Total unique patients: 871
  – Total samples: 906
  – Average Turnaround time: <6 business days

• 28 gene panel live since August 19, 2013
Clinical Utility of Test Results

- Diagnostic work-up
- Prognosis
- Monitoring of minimal residual disease
- Comparison of two lesions in the same patient
- Identification of targetable markers
- Detection of novel biomarkers
- Insights into tumor biology
Thank You

Together, we will end cancer.