NGS Applications on the Illumina NextSeq 500

April 6th 2017

Mehdi Keddache PhD
Sequencing Specialist
Illumina Instruments available at UTHSCSA

**MiSeq**
- 1M, 4M, 15M, 25M reads per runs
- Up to 300 bases per read
- As low as $102 /GB
- 5.5 to 56 h per run

**NextSeq 500**
- 130M, 400M reads per runs
- Up to 150 bases per read
- As low as $37 /GB
- 12 to 30 h per run

**HiSeq 3000**
- 2,500M reads per runs
- Up to 150 bases per read
- As low as $21 /GB
- 18 to 84 h per run
The NextSeq® 500 Delivers on Three Key Aspects

1. Flexibility
2. Speed
3. Simplicity
One System, Two Output Modes

- **High Output**
  - 400M Clusters
  - 120 Gigabases | PE150

- **Mid Output**
  - 130M Clusters
  - 40 Gigabases | PE150
Flexible applications

**High-Output**
- Up to 120 Gb
- 400M clusters PF
- 1x75 bp, 2x75 bp, 2x150 bp

**Mid-Output**
- Up to 40 Gb
- 130M clusters PF
- 2x75 bp, 2x150 bp

- 30x genome
- 6-12 exomes RNA-Seq
- 20 GEX profiles
- 2-3 exomes
- 2-4 samples RNA-Seq
- 6-36 panels

illumina
NextSeq 500: Faster 2-channel chemistry

4-channel

2-channel

Red Image

Green Image
NextSeq 500: Faster 2-channel chemistry

50% REDUCTION IN IMAGING TIME
NextSeq 500: Faster 2-channel chemistry

24 h

25% faster

18 h

MiSeq v3

NextSeq 500

2x75bp run
Fast Applications

Human Genome 30 | HOURS

Exome | Transcriptome 18 | HOURS

Gene Expression Profile 12 | HOURS

2 x 150bp 2 x 75bp 1 x 75bp
NextSeq 500: Designed for Simplicity

CARTRIDGE FORMAT | AUTOMATED WASH | SELF-CLEANING
NextSeq 500: Simplicity and Efficiency

Simple Load and Go 10min Setup

Chemistry
Flow cell
Buffer
Waste
Sequencing Performance in Standard Test Runs

2x150, High Output FC, PhiX library
RNA-Seq on NextSeq

Gene-level FPKM Comparisons

A

\[ R^2 = 0.99 \]
\[ \text{slope} = 0.98 \]

NextSeq 500 gene FPKM

HiSeq gene FPKM

14277 genes

B

\[ R^2 = 0.99 \]
\[ \text{slope} = 1.02 \]

NextSeq 500 gene FPKM

MiSeq gene FPKM

13999 genes

Log₂ Fold Change Comparisons

A

\[ R^2 = 0.99 \]
\[ \text{slope} = 1.0 \]

HiSeq \[ \log_2 \text{fold change Brain to UHRR} \]

NextSeq 500 \[ \log_2 \text{fold change Brain to UHRR} \]

13684 genes

B

\[ R^2 = 0.98 \]
\[ \text{slope} = 0.99 \]

MiSeq \[ \log_2 \text{fold change Brain to UHRR} \]

NextSeq 500 \[ \log_2 \text{fold change Brain to UHRR} \]

11599 genes

GAPDH coverage

NextSeq

HiSeq

MiSeq
Error profile comparison

NextSeq 500 2x150  PCR-free WGS

HiSeq 3000 2x150  PCR-free WGS

1.5% raw error rate

1.5% raw error rate
A wide range of applications

- RNA-seq
- miR-seq
- ChIP-seq
- Methylation sequencing
- WG meta genomics
- Genome Sequencing
- Exome
- Genomic Enrichment
- Targeted panels

NextSeq 500

MiSeq

HiSeq 3000
Sequencing costs

<table>
<thead>
<tr>
<th></th>
<th>30X Genome (120 GB)</th>
<th>100X Exome (10 GB)</th>
<th>RNA-seq (30M reads)</th>
<th>GEx (10M reads)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NovaSeq 5/6000 S2</td>
<td>$1,890</td>
<td>$205</td>
<td>$90</td>
<td>$30</td>
</tr>
<tr>
<td>HiSeq 4000/3000</td>
<td>$2,464</td>
<td>$264</td>
<td>$123</td>
<td>$24</td>
</tr>
<tr>
<td>HiSeq 2500 (v4)</td>
<td>$3,806</td>
<td>$381</td>
<td>$160</td>
<td>$31</td>
</tr>
<tr>
<td>HiSeq 2500 Rapid</td>
<td>$5,840</td>
<td>$553</td>
<td>$245</td>
<td>$48</td>
</tr>
<tr>
<td>NextSeq HO</td>
<td>$4,240</td>
<td>$442</td>
<td>$199</td>
<td>$35</td>
</tr>
</tbody>
</table>

For Research Use Only. Not for use in diagnostic procedures.
Illumina Instruments available at UTHSCSA

**Long read length**
Low read counts

- **MiSeq**
  - 1M, 4M, 15M, 25M read per runs
  - Up to 300 bases per read
  - As low as $102 /GB
  - 5.5 to 56 h per run

**Fast runs**
Medium read counts

- **NextSeq 500**
  - 130M, 400M reads per runs
  - 12 to 30 h per run

**Cost effective**
High read counts

- **HiSeq 3000**
  - 2.500M reads per runs
  - Up to 150 bases per read
  - As low as $21 /GB
  - 18 to 84 h per run

15 micro RNA samples that would need to wait several weeks to be loaded on HiSeq or NextSeq.
Illumina Instruments available at UTHSCSA

Long read length
Low read counts

Fast runs
Medium read counts

Cost effective
High read counts

MiSeq
1M, 4M, 15M, 25M reads per runs
Up to 300 bases per read
As low as $ 102 /GB
5.5 to 56 h per run

NextSeq 500
130M, 400M reads per runs
12 to 30 h per run

HiSeq 3000
2,500M reads per runs
Up to 150 bases per read
As low as $ 21 /GB
18 to 84 h per run

Quick test of new library preparation method before running a HiSeq full lane or a full NextSeq flowcell
Illumina Instruments available at UTHSCSA

**Long read length**
Low read counts

- **MiSeq**
  - 1M, 4M, 15M, 25M reads per run
  - Up to 300 bases per read
  - As low as $102 /GB
  - 5.5 to 56 h per run

**Fast runs**
Medium read counts

- **NextSeq 500**
  - 130M, 400M reads per run
  - Up to 150 bases per read
  - As low as $37 /GB
  - 12 to 30 h per run

**Cost effective**
High read counts

- **HiSeq 3000**
  - 2,500M reads per run
  - Up to 150 bases per read
  - As low as $21 /GB
  - 18 to 84 h per run

- **8 - 15 ChIP-seq samples for a grant submission**
Illumina Instruments available at UTHSCSA

Long read length
Low read counts

Fast runs
Medium read counts

Cost effective
High read counts

10-12 additional RNA-seq samples requested by a reviewer before accepting manuscript for publication

MiSeq

1M, 4M, 15M, 25M reads per run
Up to 300 bases per read
As low as $102/GB
5.5 to 56 h per run

NextSeq 500

130M, 400M reads per run
Up to 150 bases per read
As low as $37/GB
12 to 30 h per run

HiSeq 3000

2,500M reads per run
Up to 150 bases per read
As low as $21/GB
18 to 84 h per run
Illumina Instruments available at UTHSCSA

- **Long read length**
  - Low read counts
- **Fast runs**
  - Medium read counts
- **Cost effective**
  - High read counts

**MiSeq**
- 1M, 4M, 15M, 25M reads per runs
- Up to 300 bases per read
- As low as $102/GB
- 5.5 to 56 h per run

**NextSeq 500**
- 130M, 400M reads per runs
- Up to 300 bases per read
- 12 to 30 h per run

**HiSeq 3000**
- 2,500M reads per runs
- Up to 150 bases per read
- As low as $21/GB
- 18 to 84 h per run

- 24 Exomes in 1 lane for a pilot project that is not funded
Questions?
Targeted Methylation Sequencing
Methylation affects

- Infection
- Aging
- Cancer
- Addiction
- Alzheimer’s Disease
- Development

Development
Methylation changes can precede mutagenesis in tumors ...
Common Studies Using DNA Methylation

Epigenome Wide Association Studies (EWAS)
Discover / Screen for Biomarkers

Gene Expression & Genetics Studies
Expand Understanding
Epigenomic Wide Association Studies (EWAS)

Discover biomarkers associated with a population

Advances in epigenome-wide association studies for common diseases

Dirk S. Paul, Stephan Beck
UCL Cancer Institute, University College London, London, WC1E 6BT, UK
Enhancing Gene Expression Data with Methylation Analysis

What

Gene Expression
- Which genes are expressed

How / Why

Methylation
- How or why the genes are being regulated

From CORRELATIVE toward CAUSATIVE publications
Tools for Methylation Analysis

From screening to validation

- Sequencing Solutions (WGBS)
  - Deep dive

- Microarray Solutions
  - Cost effective

- Large Scale Screens
  - Low Cost/Sample

- Validate Critical Samples

For Research Use Only. Not for use in diagnostic procedures.
Challenges with Methylation Tools

*Epigenome-wide association & regulation of RNA*

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>WGBS \nExpensive \n(Small sample size)</td>
<td>Arrays \nLimited Targets \n(Discovery Gaps)</td>
</tr>
</tbody>
</table>

Result: Single base resolution for a fraction of the cost

For Research Use Only. Not for use in diagnostic procedures.
Illumina Methylation Portfolio

From screening to validation

- **Infinium® MethylationEPIC Kit**
  - 450K core content + enhancers

- **TruSeq® Methyl Capture EPIC**
  - EPIC content at increased density, low input, fast

- **Validate Critical Samples or Large Scale Screens**
  - Large Scale Screens
    - Low Cost/Sample

- **TruSeq® DNA Methylation WGBS**

For Research Use Only. Not for use in diagnostic procedures.
TruSeq® Methyl Capture EPIC

Regions Covered Compared to WGBS

CpG Islands 97%
CpG Shores 99%
CpG Shelves 86%
TS200 89%
Promoters 95%
Open Chromatin 66%
Transcription Factor Binding Sites (TFBS) 79%
Multifunctional Epigenetic Domains (MEDs) 96%
FANTOM5 enhancers 99%
DNAse Hypersensitive Site 33%

>3M Methylation Sites

Percent coverage compared to whole genome bisulfite sequencing (WGBS), as referenced in Illumina TruSeq® Methyl Capture EPIC data sheet

For Research Use Only. Not for use in diagnostic procedures.
Illumina Targeted Methyl Capture Solution

Content
Increase Power of Discovery

Workflow
2 day Library Prep
4 days DNA to results

Full Kit Configuration
Includes: library prep, indices, bisulfite conversion, and enrichment reagents

Data Analysis
Dedicated Apps
**TruSeq® Methyl Capture EPIC Workflow**

<table>
<thead>
<tr>
<th>DAY 1</th>
<th>DAY 2</th>
<th>DAY 3</th>
<th>DAY 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA to Analysis in less than 4 days</td>
<td><strong>LIBRARY PREP</strong> = 2 Days*</td>
<td><strong>SEQUENCING</strong> = ~20 hours**</td>
<td><strong>ANALYSIS</strong> = 12 hours</td>
</tr>
</tbody>
</table>

*12 hour assay time with overnight hybridization  
** NextSeq® or HiSeq® Rapid Run 2x100 cycles

For Research Use Only. Not for use in diagnostic procedures.
Technology and Workflow

1. Shear DNA
   - Ligate indexed adapters

2. Hybridize with biotinylated probes

3. Pull-down target sequences with SA beads

4. Bisulfite conversion
   - Amplification

For Research Use Only. Not for use in diagnostic procedures.
Flexible Configuration

Samples per run

Mid Throughput

8*

830/sample**

High Throughput

6 – 48*

590/sample*

*Samples calculated for ~55M PE100 reads
** Pricing includes library prep and sequencing reagents, USD, AMR List price as of April 2017

For Research Use Only. Not for use in diagnostic procedures.
MethylSeq / MethylKit Apps
Enables answers to hard questions

MethylSeq App 2.0
Aligns reads to the genome and calls methylation at all cytosines

MethylKit App 2.0
Enables differential methylation calling between 2 or more samples

For Research Use Only. Not for use in diagnostic procedures.
Targeted Methylation Sequencing

- Increased Discovery Power
- Simple DNA to Analysis Solution
- Up-to-date Methylation Content
Questions?