Harness the Power of MiSeq—Simplify Next Gen Sequencing

Michael G Smith, PhD
Sequencing Specialist
mismith@illumina.com

Carolyn Neiss
Territory Account Manager
cneiss@illumina.com

Alissa Nolan
MiSeq Support Specialist
anolan@illumina.com
A Sequencer for Every Need. Every Budget. Every Lab.


<table>
<thead>
<tr>
<th>Sequencer</th>
<th>Storage Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>HiSeq 2000</td>
<td>540 – 600Gb</td>
</tr>
<tr>
<td>HiSeq 1000</td>
<td>270 – 300Gb</td>
</tr>
<tr>
<td>HiScanSQ</td>
<td>135 – 150Gb</td>
</tr>
<tr>
<td>GA\textsubscript{IIx}</td>
<td>95Gb</td>
</tr>
<tr>
<td>MiSeq</td>
<td>1.5 – 2Gb</td>
</tr>
</tbody>
</table>
MiSeq
The World’s Most Widely Adopted Sequencing Technology Just Got Personal

UNPRECEDENTED PERFORMANCE
Complete workflow for some applications in a single day
Throughput >2.0 Gb
Scalable run time vs. output: 36bp to 2x150bp

UNMATCHED COST EFFECTIVENESS
<$125,000 list price “all in”
$695-$965 per run (amplification + sequencing)

EASIEST TO USE WORKFLOW
On-board cluster generation and data analysis
Plug and play reagents with RFID

MOST ACCURATE SEQUENCING
Uses proven TruSeq SBS reversible terminator chemistry

CE AND NGS APPLICATIONS
MiSeq Performance Improvements

Next Six Months

- 3.4 M clusters/run
- 2 x 150 bp read length
- 5 min cycle time

100's Mb

1-2 Gb
- 5-7 M clusters/run
- 2 x 150 bp read length
- 5 min cycle time

7 Gb
- ~15 M clusters/run
- 2 x 250 bp read length
- 4 Gb at 2 x 150
- 6 - 7 Gb at 2 x 250
- <5 min cycle time

• Read lengths get longer - 66% increase
• Output grows - 3 fold improvement in number of clusters run
Illumina Experiment Manager

Intuitive Experimental Setup

Offline wizard to help organize samples and experiment design

Creates set of instructions for the instrument on a per run basis

Connects sample prep information to downstream automated analysis
MiSeq Control Software

Fast and Efficient Run Setup

Interactive menu driven run setup
Starting a MiSeq run only takes a few clicks
Contextual help videos demonstrate proper consumables loading
At a glance run summary metrics
MiSeq Reporter - On-Instrument Alignment & Variant Calling

Unprecedented Walk-Away Informatics Solution

Automated secondary analysis for key applications:

- Resequencing
- Amplicon sequencing
- 16S Metagenomics
- *de novo* assembly
- Small RNA
- Library QC

Integrated analysis hardware

Output in standard formats:

- Fastq
- BAM
- Vcf
- txt
BaseSpace
The Best Place to Store Your NGS Data

- Eliminates need for onsite storage and compute
- Results available anywhere, anytime
- Browse the results via web-based graphical environment
- Access to a growing suite of analysis tools
- Tools for collaboration and sharing

Reads and qualities
Sample and experiment descriptions
Analysis results
variants
contigs
metagenomes
coverage statistics
miRNA counts
more…
BaseSpace Partners

Announcing Initial App Partners!

- Key Vendors in clinical interpretation, annotation, and visualization.
- Differing data models (thick client, web services, hybrid)
- Differing data usage (.fastq, .vcf)
- Twenty additional vendors working to deploy in BaseSpace soon.
Illumina Sample Prep Solutions
Integrated workflow from sample to analyzed data

- Nextera Exome Enrichment
- Nextera Custom Enrichment
- TruSeq Exome Enrichment
- TruSeq Custom Enrichment
- TruSeq DNA
- TruSeq RNA
- TruSeq Small RNA
- Nextera/Nextera XT
- TruSeq Custom Amplicon
- TruSeq Amplicon Cancer Panel
- TruSeq Chemistry Clustering & Sequencing
Illumina’s DNA Sample Prep Portfolio

End-to-end solutions for any NGS prep need

Nextera XT
- Small genomes, amplicons, plasmids
- 1ng DNA input
- 90 min workflow
- 96 indexes
- $30 USD/Sample

TruSeq DNA LT/HT
- WGRS, TGRS, Exome
- 1ug DNA input
- ~8 hr workflow
- Index-adapter plate
- 96 indexes
- $54 USD/Sample

Nextera
- Speed!
- Larger, more complex genomes
- 50ng DNA input
- 90 min workflow
- 96 indexes
- $75 USD/Sample
Library Preparation

- Prepares sample nucleic acid for sequencing
  - Fragmenting
  - Generates double-stranded DNA (if necessary)
  - Flanks with Illumina adapters

- All preparation ends with the same product
  - Double-stranded DNA with insert to be sequenced flanked by adapters

- Same protocols all both platforms
Items included in this launch
96 barcodes for TruSeq DNA are here!

- **TruSeq DNA Sample Preparation HT Kits**
  - Simple, scalable, and fully supported end-to-end solution from DNA to data
  - 96 ILMN-developed, extensively pre-screened and fully QC’d barcodes
  - Pre-filled index-adapter plates (oh, yeah!)

- **TruSeq DNA Sample Preparation LT Kits**
  - Former TruSeq DNA “v2” kits (just a name change)
  - Option for lower-throughput studies
  - 24 barcodes

- **Sales & Support Materials**
  - Comprehensive FAQs
  - User Guides
  - Datasheet
  - Web Updates
  - Product Inserts
TruSeq DNA HT Sample Preparation Kits

Most widely-adopted DNA sample prep; now with 96 barcodes!

- Pre-loaded 96-well plate w/96 barcodes
- Leverages established TruSeq DNA workflow
- Illumina-developed, fully validated and QC’d barcodes
  - Tested on all ILMN platforms
  - Fully integrated with ILMN sequencing & analysis software
- Compatible with TruSeq Exome/Custom Enrichment
## TruSeq DNA LT versus HT

<table>
<thead>
<tr>
<th></th>
<th>TruSeq DNA LT (formerly v2)</th>
<th>TruSeq DNA HT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Barcodes</strong></td>
<td>24</td>
<td>96</td>
</tr>
<tr>
<td><strong>Index IDs</strong></td>
<td>Set A: 2, 4, 5, 6, 7, 12, 13, 14, 15, 16, 18, 19&lt;br&gt;Set B: 1, 3, 8, 9, 10, 11, 20, 21, 22, 23, 25, 27</td>
<td>D701 thru D712&lt;br&gt;D501 thru D508</td>
</tr>
<tr>
<td><strong>Barcode delivery</strong></td>
<td>Individual tubes</td>
<td>Index-adapter plate</td>
</tr>
<tr>
<td><strong>Min sample order</strong></td>
<td>48</td>
<td>96</td>
</tr>
<tr>
<td><strong>Batches</strong></td>
<td>4 batches of 12</td>
<td>4 batches of 24</td>
</tr>
<tr>
<td><strong>List price per sample (USD)</strong></td>
<td>$54</td>
<td>$54</td>
</tr>
<tr>
<td><strong>Best for</strong></td>
<td># of samples processed</td>
<td>Recommended Kit</td>
</tr>
<tr>
<td></td>
<td>&lt;24</td>
<td>LT</td>
</tr>
<tr>
<td></td>
<td>24-48</td>
<td>LT or HT</td>
</tr>
<tr>
<td></td>
<td>&gt;48</td>
<td>HT</td>
</tr>
</tbody>
</table>
Step 1: Tagmentation of template DNA
Step 2: PCR to add adapters and indices

~ 300 bp

Tagmentation

Reduced-Cycle PCR Amplification
Step 3: Cleanup and Sequence

- Tagmentation
- Reduced-Cycle PCR Amplification
- Sequencing-Ready Fragment
Nextera or TruSeq DNA Sample Preparation?

- **When to use Nextera?**
  - Speed and ease
  - Low sample input
  - High manual throughput
  - Supported indexing up to 96 samples per lane
  - **Perfect for MiSeq** = e.g. small genomes and amplicons

- **When to use TruSeq DNA Sample Prep?**
  - Optimal for tight insert size distribution for long read PE sequencing
  - Human genome sequencing - highest coverage uniformity and diversity
  - Enables TruSeq Exome /Custom Enrichment
  - Cost
Long Jump: High Diversity Mate-pair libraries from sub-microgram amounts of genomic DNA

Matthew Hims, Ole Schulz-Trieglaff, Helen Bignell, Niall Gormley and Geoffrey Smith
Illumina Cambridge Ltd, Saffron Walden, UK

Simplified improved workflow
A modified Nextera reaction introduces a biotinylated adapter into DNA. After circularization of the fragments, the original junction ends are selected and prepared for sequencing by ligation of Illumina adapters.

Two protocols
One microgram of E. coli DNA was processed through the gel free protocol (left), and four micrograms was processed through the size selected protocol using the Pipelin Prep (Sage Sciences) targeting either a 5kb or 8kb library (right). Sequencing on MiSeq for 2x100bp

Key advantages
✓ Shorter and simpler protocol
✓ Low input DNA
✓ No gels for standard method
✓ High diversity
✓ Longer read lengths (2x150bp)
✓ Gap sizes up to 20kb
✓ Defined junction to mark fragment ends
✓ Low % incorrect pairs
✓ Automatable

Tight size range
More DNA required
Lower inputs proven

Gel free
Standard method
- Low input DNA
- Short protocol
- Broad sizes
- High diversity

Size selected
Optional method
- More DNA required
- Longer protocol
- Tighter size range
- Lower diversity

Trout genome
A series of size selected mate pair libraries (A) and (B) were used to provide a scaffold for the assembly of pooled BAC clones that had been sequenced previously from regular TruSeq libraries. Short-read data alone produced poor assemblies (N50<10kbp). By contrast, the addition of the Long Jump data generated complete scaffolding for almost every BAC clone (N50 of 150kbp).

Human genome
One microgram of human DNA was processed with the gel free protocol to make a high diversity mate pair library. The complexity of the library is sufficient to allow approx. 200-fold physical coverage per lane of a HiSeq2000 run with 2x100bp reads.

De novo assembly
A comparison of de novo assemblies from either (A) the standard 300-400bp TruSeq library and (B) the gel free mate pair protocol. In each case, the assemblies were performed with a single library type. In each case, the new protocol generated more complete assemblies and resulted in a single long scaffold spanning the majority of the genome. Assemblies were generated with the Velvet software4 with standard settings for coverage and mean insert size of 4kb.

Future perspectives
- Kits available: Summer 2012
- Support one microgram input DNA with gel free or sized protocols
- Potential to generate ‘close-to-finished’ reference genomes from a single library type (ie removing the necessity for a mixed assembly with regular paired reads)
- High diversity of human libraries and broad size distribution in gel free preps allows full physical coverage on a single lane of a HiSeq2000 and potential application for finding structural variants
- Further improvements to de novo scaffolding will be possible with optimisation of assembly software
- Lower inputs proven successful in R&D with modified workflow

<table>
<thead>
<tr>
<th>Library</th>
<th>5 kbp</th>
<th>10-20 kbp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diversity</td>
<td>257M</td>
<td>40M</td>
</tr>
<tr>
<td>Correct pairs: 99.4%</td>
<td>99.2%</td>
<td></td>
</tr>
</tbody>
</table>

We thank Mike Miller from HHMI and collaborators, University of Oregon for sharing preliminary data on the sequencing of the Salmonid genome.

<table>
<thead>
<tr>
<th>Library</th>
<th>N50</th>
<th>Max Scaffold</th>
<th>Total size</th>
<th>Genome Coverage</th>
<th>Input Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. Coli MG1655</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TruSeq (A)</td>
<td>149,102 (11)</td>
<td>285,913</td>
<td>4,586,407 (122)</td>
<td>98%</td>
<td>62x</td>
</tr>
<tr>
<td>Mate-Par (B)</td>
<td>4,635,607 (1)</td>
<td>4,635,607</td>
<td>4,664,330 (15)</td>
<td>99%</td>
<td>48x</td>
</tr>
<tr>
<td>R. raphanoides</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TruSeq (A)</td>
<td>279,376 (6)</td>
<td>551,032</td>
<td>4,161,361 (73)</td>
<td>82%</td>
<td>47x</td>
</tr>
<tr>
<td>Mate-Par (B)</td>
<td>2,590,288 (1)</td>
<td>2,590,288</td>
<td>4,591,572 (94)</td>
<td>94%</td>
<td>31x</td>
</tr>
<tr>
<td>B. cereus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TruSeq (A)</td>
<td>485,654 (3)</td>
<td>1,197,558</td>
<td>5,371,324 (70)</td>
<td>92%</td>
<td>66x</td>
</tr>
<tr>
<td>Mate-Par (B)</td>
<td>3,611,328 (1)</td>
<td>3,611,328</td>
<td>5,444,030 (20)</td>
<td>99%</td>
<td>40x</td>
</tr>
</tbody>
</table>


Input [DNA] 1µg 4µg 4µg
Correct pairs 98.4% 98.6% 97.6%
Chimeras 1.9% 1.7% 3%
Diversity | 510M | 265M | 110M |
Of MiSeq, Microbes, and Man

<table>
<thead>
<tr>
<th>Organism</th>
<th>Genome size</th>
<th>n</th>
<th>depth</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Staphylococcus aureus</em> (MRSA)</td>
<td>2.8 Mb</td>
<td>9</td>
<td>50x</td>
</tr>
<tr>
<td><em>Mycobacterium tuberculosis</em> (TB)</td>
<td>4.4 Mb</td>
<td>6</td>
<td>50x</td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td>4.6 Mb</td>
<td>6</td>
<td>50x</td>
</tr>
<tr>
<td><em>Plasmodium falciparum</em></td>
<td>22.9 Mb</td>
<td>2</td>
<td>30x</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Human</th>
<th>Target size</th>
<th>n</th>
<th>depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 exons</td>
<td>3 kb</td>
<td>225</td>
<td>2000x</td>
</tr>
<tr>
<td>Targeted region</td>
<td>0.5 Mb</td>
<td>54</td>
<td>50x</td>
</tr>
<tr>
<td>All coding exons</td>
<td>25 Mb</td>
<td>1</td>
<td>40x</td>
</tr>
<tr>
<td>RNA, miRNA, ChIP-Seq, etc</td>
<td>6M tags</td>
<td>1</td>
<td>n.a.</td>
</tr>
</tbody>
</table>
Performance Comparison of Benchtop Sequencers
Loman et al. (Nature Biotechnology, April 2012)

Goal:

- Assess the current slate of benchtop sequencers across the same bacterial sample

Conclusion:

- “The MiSeq generated the highest throughput per run and lowest error rate of the instruments, without significant indel errors and the lowest rate of substitution errors”

- The MiSeq workflow has the fewest manual steps as template amplification is done directly on the instrument without manual intervention in contrast to the Ion Torrent PGM and 454 GS Junior, which require preparation of amplified sequence libraries through emulsion PCR and enrichment stages off the instrument.
A pilot study of rapid benchtop sequencing of *Staphylococcus aureus* and *Clostridium difficile* for outbreak detection and surveillance. 
*Eyre et al. BMJ Open. 2012*

- MRSA clusters were identified as outbreaks, with most sequences in each cluster indistinguishable.
- In *C. difficile* clusters, closely epidemiologically linked cases were genetically distinct.
- Rapid sequencing in *C. difficile* surveillance provided early outbreak detection and identified previously undetected community transmission.
MRSA Outbreak and Sequencing

• Rapid whole-genome sequencing of microorganisms using the MiSeq system was used to distinguish neonatal patients infected with MRSA outbreak strains and non-outbreak on a clinical time scale

• The MRSA isolate was identified after construction of a phylogenetic tree comparing SNPs in the core genome to a reference genome
• NGS can be used to prevent transmission events between patients carrying different strains

• From genome DNA to data took 1.5 days
• WGS can provide clinically relevant data in a time frame that can influence patient care
• WGS can provide improved resolution to define transmission pathways and characterize outbreaks in a hospital setting

Rapid Whole-Genome Sequencing for Investigation of a Neonatal MRSA Outbreak


N ENGL J MED 366;24 NEJM.ORG JUNE 14, 2012
Targeted Re-Sequencing
Options for Targeted Resequencing with Illumina

From exons to exomes......

Low-plex
100’sbp-100kb
TruSeq Custom Amplicon
Nextera XT

Mid-plex
500kb – 25Mb
TruSeq Custom Enrichment
Nextera Custom Enrichment

High-plex
62Mb+
TruSeq Exome Enrichment
Nextera Exome Enrichment
TruSeq Exome & Custom Enrichment Workflow

*Protocol includes 2 successive rounds of enrichment*

*PCR, cluster generation & sequencing*
Nextera Exome Enrichment Kits

The most comprehensive exome...now with the simplest prep workflow

- Fastest and simplest workflow
  - Prep up to 12 samples in <3 hrs
  - Total prep & enrich time ~2.5 days
  - No need for mechanical DNA fragmentation
  - 12-plex pre-enrichment pooling

- Low DNA input of 50ng

- One kit does it all
  - Sample prep & enrichment

- Expansive coding & non-coding content

- Excellent enrichment rates & coverage uniformity
Nextera Enrichment Workflow

Prepare spls in <3hrs without mechanical shearing; enriched libraries for 96 spls ~2.5 days

A. Sample Preparation

B. Denature double-stranded DNA library (for simplicity, adapters and indexes not shown)

C. Hybridize biotinylated probes to targeted regions

D. Enrichment using streptavidin beads

E. Elution from beads
Introducing TruSeq Custom Enrichment
Premier sequencing technology coupled with expertise in custom design

- Economically target 700kb – 25Mb of DNA
  - Highest coverage uniformity
  - Lowest enrichment costs
  - Lowest sequencing costs
    - Only requires 2x50 – 2x75bp read length

- Pre-enrichment sample pooling
  - Up to 12 samples per enrichment reaction
    - High throughput and saves FTE time

- Easily customize content with DesignStudio
  - Personalized and easy to use online design tool and ordering

- Complete end-to-end solution
  - Integrated with “gel-free” TruSeq DNA sample prep and backend data analysis
  - One-stop-shop for all sequencing components
**Nextera Custom Enrichment Kits**

*Fastest, easiest, lowest input sample prep workflow for custom enrichment*

- Target 500kb to 25Mb
  - High enrichment efficiency & coverage uniformity
  - Customize content with DesignStudio
- Pre-enrichment sample pooling
  - Up to 12 spls per enrichment rxn
  - Reduce hands-on time; increase throughput
- One kit does it all
  - Sample prep and enrichment
- Fastest and simplest workflow
  - Prep up to 12 samples in <3 hrs
  - No need for mechanical DNA fragmentation
- DNA input of 50ng
Nextera Enrichment Workflow
Faster, simpler, lower DNA input requirement than TruSeq

- Archive samples for future analyses
- Increase sample throughput
- Decrease FTE time
- Remove need for mechanical DNA fragmentation
- Reduce need for automation
- Make process easier!
**TruSeq Enrichment or Nextera Enrichment?**

*Nextera for low DNA input/easier workflow; TruSeq for improved coverage uniformity*

<table>
<thead>
<tr>
<th></th>
<th>TruSeq Enrichment</th>
<th>Nextera Enrichment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Library prep input</td>
<td>1ug</td>
<td>50ng</td>
</tr>
<tr>
<td>Library prep time (hrs)</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>Prep &amp; Enrich kits</td>
<td>Separate</td>
<td>Bundled</td>
</tr>
<tr>
<td>Total workflow (days)</td>
<td>3.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Pre-enrich pooling (E/C)</td>
<td>6/12</td>
<td>12/12</td>
</tr>
<tr>
<td>Dual-indexing support</td>
<td>No (Up to 24 single-indexes)</td>
<td>Yes (Up to 96 dual-indexes)</td>
</tr>
<tr>
<td>% Enrichment (aligned/+padded)</td>
<td>55%/65%</td>
<td>55%/65% (60%/70% typical)</td>
</tr>
<tr>
<td>% Region Dropout (E/C)</td>
<td>&lt;1%/&lt;5%</td>
<td>&lt;1%/&lt;5%</td>
</tr>
<tr>
<td>Cov Uniformity (0.2X Mean)</td>
<td>≥80%</td>
<td>≥80%</td>
</tr>
<tr>
<td>Cov Uniformity (0.5X Mean)</td>
<td>≥60%</td>
<td>≥60%</td>
</tr>
<tr>
<td>Exome Pricing (USD; List)</td>
<td>$54 prep + $49 enrich = $103</td>
<td>$124 prep &amp; enrich</td>
</tr>
</tbody>
</table>

E = exome; C = custom
Targeted Resequencing Methods

- Full length PCR
- Target
- Custom Enrichment
- TruSeq Custom Enrichment
- Amplicon Sequencing
- TruSeq Custom Amplicon
- 2-step PCR
  1) Target
  2) TruSeq Custom Amplicon
"Full length" PCR primers

All sequences required for Illumina sequencing are added to the end of the gene specific oligo sequences

- **Advantages**
  - Short amplicons
  - Indexing occurs in 1st step
  - Uniform coverage across each amplicon

- **Disadvantages**
  - High cost of long oligos: add 60+ bases to each gene specific oligos
  - Multiple oligos: one for each index for every amplicon
  - Potentially high DNA cost for many targets
  - Long, paired sequence reads may be required for complete amplicon coverage
  - Laborious normalization

**Flowchart:**

1. **PCR Amplify targets**
2. **PCR Purification**
3. **Quantify Libraries**
4. **Pool & run on instrument**
Can short read sequencing be used to profile complex microbial communities?

- Sequenced 2 x 100 base reads on the GAIIx
- Sequenced the V4 region of the 16S rRNA of 67 bacterial species which were assembled into 3 mock mixtures
- Developed algorithms to handle the large number of short reads (QIIME)
- Demonstrate the 2000 single-end reads are sufficient to recapture the same relationships among samples they observe with the full dataset
- Got similar results to similar experiments using longer Sanger or 454 reads
PCR Amplicon Sequencing / Metagenomics

Target gene:

3' + strand

Amplication primers with annealing sites:

Amplication products:

Sequencing primers with annealing sites:

A

B

Relative abundance (% of 16S rRNA gene sequences)

Expected

Even1

Even2

Even3

5' primer

3' primer

Bacteroidetes

Rikenellaceae

Prevotellaceae

Porphyromonadaceae

Bacteroidaceae

Actinobacteria

Coriobacteriaceae

Bifidobacteriaceae

Proteobacteria

Cedocceae

Enterobacteriaceae

Citrobacteraceae

Proteus

Providencia

Not assigned
V4 region of the 16S rRNA gene was amplified from host-associated and free-living microbial communities

Sequenced 24 samples on the HiSeq and MiSeq

Confirmed that data and drawn biological conclusions are consistent across two platforms and sequence reads

Validated a protocol now adopted by the Earth Microbiome Project

MiSeq compatible primers: http://www.nature.com/ismej/journal/vaop/ncurrent/extref/ismej20128x2.txt
Procrustes plots comparing HiSeq vs. MiSeq

HiSeq : HiSeq

HiSeq : MiSeq

MiSeq : MiSeq
2 step PCR

Target specific oligos have 20bp of adaptor sequence added to them. Adaptor sequences extended to full length in a 2nd PCR reaction.

**Advantages**
- Short amplicons
- Lower set up costs than full length
  - Shorter oligos
  - Only one set of indexing oligos required
- Uniform coverage across each amplicon

**Disadvantages**
- Indexing occurs after 1st amplification
- Processing of amplified material pre-indexing
- Potentially high DNA cost for many targets
- Long, paired sequence reads may be required for complete amplicon coverage
PCR + Nextera approach

Adaptor sequences are introduced to standard PCR products by tagmentation. Adaptors are extended to full length via PCR

- **Advantages**
  - Fewer, longer amplicons
  - Allows for shorter sequencing reads (less time on sequencing instrument)
  - No ‘special’ oligo sequences required
  - Amenable to 96 sample indexing
  - Sequence challenging regions 200-10,000bp without complex tiling of small amplicons

- **Disadvantages**
  - Reduced coverage on ends of DNA
  - Indexing occurs after 1st amplification
  - Potentially high DNA primer cost for many targets
  - Numerous clean up steps

Diagram:
- PCR Amplify targets
  - Quantify PCR amplicons
    - Nextera Prep
      - Quantify Libraries
        - Pool & run on instrument
Nextera XT DNA Sample Prep
The fastest & easiest prep for small genomes, PCR amplicons and plasmids

- **Rapid Prep**
  - 90 min prep, only 15 min of hands on time

- **Optimized for small genomes, PCR amplicons and plasmids**

- **Innovative sample normalization**
  - No library quantification needed

- **Fastest time to results**
  - DNA to analyzed data in <8 hours with MiSeq

- **Ultra low input**
  - only a single nanogram of input DNA needed

- **Enabling Price**
  - US List **$30 per sample!**
Sample Normalization is included

No library quantification or qPCR is required – go straight to MiSeq!

Nextera XT sample pooling is as simple as pipetting 5 µl!
## Which version of Nextera should I use?

<table>
<thead>
<tr>
<th></th>
<th>Nextera</th>
<th>Nextera XT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Applications</strong></td>
<td>Large/complex genomes (human, mammalian, plants, inverts)</td>
<td>Small genomes, amplicons, plasmids (bacteria, archaea, viruses, PCR amplicons, plasmids)</td>
</tr>
<tr>
<td><strong>Total Sample Prep Time for 8 samples (hands on)</strong></td>
<td>90 (15)</td>
<td>85 (10)</td>
</tr>
<tr>
<td><strong>DNA input</strong></td>
<td>50 ng</td>
<td>1 ng</td>
</tr>
<tr>
<td><strong>Post tag cleanup?</strong></td>
<td>Zymo</td>
<td>None</td>
</tr>
<tr>
<td><strong>Sample normalization</strong></td>
<td>Manual (e.g. bioanalyzer, qPCR)</td>
<td>Bead-based normalization included</td>
</tr>
<tr>
<td><strong>Indexing</strong></td>
<td>96</td>
<td>96</td>
</tr>
<tr>
<td><strong>List Price per sample</strong></td>
<td>~$75</td>
<td>~$30</td>
</tr>
</tbody>
</table>
TruSeq Custom Amplicon

*The fastest and easiest multiplexed amplicon assay optimized for MiSeq*

- **Rapid & economical**
  - Up to 384 amplicons per sample (250bp ea)
    - 1536 amplicons Coming Soon!
    - 150bp and 450bp amplicons Coming Soon!
  - Up to 96 samples per plate
  - Plate-based processing
  - <8 hrs from gDNA to sequencing-ready library
    - Coming Soon – Mouse, Rat and Cow!
  - Utilizes standard lab equipment
  - No quant needed before sequencing

- **Fully customized target probes & capture**
  - DesignStudio for interactive design and ordering
    - Improved Design Studio Coming Soon!
  - Personalized and easy to use
  - Rapid design turnaround

- **Pre-configured, automated data analysis**
TruSeq Amplicon Assay Overview

Begin with 250ng genomic DNA

250ng Genomic DNA

Hybridization

TruSeq Amplicon – Cancer Panel oligos
TruSeq Amplicon Assay Overview

*Using a pair of primers, each targeted region is amplified in each sample*
TruSeq Amplicon Assay Overview

Incorporation of indexed primers followed by normalization and sample pooling

Genomic DNA

Hybridization

Extension-Ligation

Amplification
Sample Normalization is included

No library quantification or qPCR is required – go straight to MiSeq!

Completed libraries, range of yields

Simple Bead-based Normalization

Quant in triplicate with qPCR

Calculate dilutions

Manually dilute and pool

Index CV for 20-sample pools

<table>
<thead>
<tr>
<th></th>
<th>Pool A</th>
<th>Pool B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pool A</td>
<td>15.8%</td>
<td></td>
</tr>
<tr>
<td>Pool B</td>
<td>18.2%</td>
<td></td>
</tr>
</tbody>
</table>

TruSeq Amplicon sample normalization is as simple as pipetting 5 µl!
TruSeq® Amplicon – Cancer Panel
Hundreds of loci. Rapid prep. FFPE-ready.

Comprehensive Content
- >35 kb total including oncogenes such as BRAF, KRAS & EGFR
- 212 amplicons in one tube; 48 genes

Unrivaled Multiplexing
- Up to 96 sample pooling on MiSeq
- >90% specificity and uniformity
- Detect low frequency variants (<5%)

Unparalleled Workflow
- FFPE-enabled with sample QC Kit
- No qPCR quant needed for normalization
- Automated paired end sequencing with MiSeq
- Pre-configured, automated data analysis

For research use only
Detection of low prevalence somatic mutations in solid tumors with ultra-deep targeted sequencing

Harismendy O et al. Genome Biology 2011, 12:R124

Targeted 71kb of mutational hotspots in 42 cancer genes using Ultra Deep Targeted Sequencing (UDT-Seq) assay

PCR – Amplicon strategy using microfluidics

Using a GAIIx:
Sensitivity for variants present at >1% was 89.1%
Sensitivity for variants present at 5% was 94%

Using a MiSeq:
Significant reduction of the substitution rate
Improved sensitivity of ~5%
Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing

Exome Sequencing on HiSeq
Validation of mutated genes
Performed on MiSeq
MiSeq Offers an NGS “Ecosystem”

- **BaseSpace™**
- **BaseSpace Data Storage, Sharing**
- **Integrated, Optimized Sample Prep**
- **Simplest NGS Workflow**
- **Brodest Applications Base**
- **Proven, Well-Published Technology**
- **Established Community of Users**

A fully integrated NGS sequencing ecosystem.
Thank you!