SPRIworks – A Simplified, Automated Sample Preparation System for Next Generation Sequencing

William Donahue Ph.D.
Manager, Molecular Biology

March 17, 2011
Presentation Overview

- Introducing the SPRIworks Fragment Library Systems I, II & III (for Illumina GA/HiSeq, Roche/454 and Life Tech SOLiD)

- Additional Applications

- SPRIworks HT Sequencing Workflows

- Questions and Answers
Liability Statements

- The PCR process is covered by patents owned by Roche Molecular Systems, Inc., and F. Hoffman-La Roche, Ltd.

- All trademarks are property of their respective owners.

- Beckman Coulter, the stylized logo, Agencourt, AMPure, CleanSEQ, SPRI and SPRlworks are registered trademarks of Beckman Coulter, Inc.

- The SPRlworks products are for research use only.
Beckman Coulter Announces the Creation of Beckman Coulter Genomics

New Company Offers Comprehensive Genomic Services

BEVERLY, Mass. (August 13, 2009) – Beckman Coulter, Inc. (NYSE: BEC) today announced the creation of Beckman Coulter Genomics. The new company, which combines Agencourt Bioscience and the newly acquired Cogenics, offers researchers a single genomic services resource with expanded global capabilities.
Sequencing Experience

- Human Genome Project: *Nature, 2001*
- Mouse Genome: *Nature, 2002*
- Human Chr.10 Finishing: *Nature, 2004*
- Brown Norway Rat Genome: *Nature, 2004*
- Kinase Mutation Discovery: *Science, 2004*
- Dog Genome: *Nature, 2006*
- >100 Microbial & Fungal Genomes: 1998-2010
- CDNA Sequencing: MGC +, 2003-2008
- Cat Genome: *Genome Research, 2007*
- Five Drosophila Genomes: *Nature, 2007*
- Human Genome Structural Variation: *Nature, 2008*
- >100 Cancer Exomes: *Science, 2006-2010*
- Atlantic Salmon Genome: 2010
- Next-Generation Sequencing: *Genome Research 2008*
Simplifying Next Generation Sequencing Workflows
SPRIworks Product Concept

Manual Protocols

DNA Fragmentation

- Enzymatic Reactions and Column Purification
- Enzymatic Reactions and SPRI Purification

Automation

- Agarose Gel Size Selection
- SPRI-Based Size Selection

8 hours manual

Manually Process – Laborious Low Throughput

No More Gels
No More Columns

SPRI Does It All

5 minutes hands on
SPRIworks: Automated Library Construction

- Easy to operate bench-top system
- Self-contained reagent cartridge
- Built in SPRI size selection chemistry—No more gels
- Run 1–10 Samples at a time
- Increases sample prep capacity
SPRIworks System Components

SPRI-TE Nucleic Acid Extractor + Method Card + SPRIworks Fragment Library Kit I Cartridges (For Illumina Genome Analyzer)
SPRIworks Library Construction Solutions

System I for Illumina
Fragment Library Construction
GA II / HiSeq

System 2 for Roche/454
Fragment Library Construction
GS FLX TTN / GS JR

System 3 for SOLiD
Fragment Library Construction
V. 3, 4, 5500, 5500xl
SPRIworks: Basic Setup Steps
Load reagent cartridge into rack
Load labware, adapters and elution tube
Transfer input sample: sheared DNA, cDNA or amplicons
Insert method card and power up
Choose size range and press “GO”
SPRIworks Fragment Library System

Performance
SPRIworks System I for Illumina
Consistent and Reproducible

Small
200-400 bp

Large
300-600 bp

Three size selection options:
No size selection, 200–400bp, 300–600bp
### Multiplex Sequence Data SPRIworks vs Manual

**SPRIworks System I for Illumina Sequencing**

<table>
<thead>
<tr>
<th>Organism</th>
<th>Lane Yield (kb)</th>
<th>Clusters (PF)</th>
<th>% PF Clusters</th>
<th>% Align (PF)</th>
<th>Align Score (PF)</th>
<th>% Error Rate (PF)</th>
<th>Tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>5793</td>
<td>9466 +/- 980</td>
<td>95.66 +/- 0.43</td>
<td>81.34 +/- 0.31</td>
<td>142.08 +/- 1.04</td>
<td>0.24 +/- 0.02</td>
<td>ATCACG</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>6074</td>
<td>9926 +/- 1061</td>
<td>96.43 +/- 0.30</td>
<td>71.55 +/- 0.74</td>
<td>135.69 +/- 1.38</td>
<td>0.66 +/- 0.00</td>
<td>CGATGT</td>
</tr>
<tr>
<td>E.coli</td>
<td>6591</td>
<td>10771 +/- 1095</td>
<td>95.87 +/- 0.21</td>
<td>88.78 +/- 0.40</td>
<td>194.96 +/- 0.77</td>
<td>0.75 +/- 0.02</td>
<td>TTAGGC</td>
</tr>
<tr>
<td>Human</td>
<td>6271</td>
<td>10247 +/- 1137</td>
<td>96.13 +/- 0.33</td>
<td>80.84 +/- 0.42</td>
<td>140.83 +/- 0.95</td>
<td>0.24 +/- 0.00</td>
<td>TGACCA</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>6474</td>
<td>10578 +/- 1162</td>
<td>96.16 +/- 0.43</td>
<td>71.18 +/- 0.39</td>
<td>134.51 +/- 1.06</td>
<td>0.66 +/- 0.01</td>
<td>ACGATG</td>
</tr>
<tr>
<td>E.coli</td>
<td>8165</td>
<td>13341 +/- 1628</td>
<td>95.61 +/- 0.35</td>
<td>88.73 +/- 0.30</td>
<td>194.66 +/- 0.67</td>
<td>0.76 +/- 0.00</td>
<td>GCCAAT</td>
</tr>
<tr>
<td>Human</td>
<td>7918</td>
<td>12938 +/- 1478</td>
<td>96.36 +/- 0.30</td>
<td>80.93 +/- 0.40</td>
<td>140.83 +/- 0.56</td>
<td>0.23 +/- 0.01</td>
<td>CAGATC</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>8037</td>
<td>13132 +/- 1591</td>
<td>96.60 +/- 0.25</td>
<td>71.65 +/- 0.20</td>
<td>135.89 +/- 0.85</td>
<td>0.64 +/- 0.02</td>
<td>ACTTGAT</td>
</tr>
<tr>
<td>E.coli</td>
<td>6081</td>
<td>9936 +/- 1159</td>
<td>95.90 +/- 0.29</td>
<td>89.30 +/- 0.35</td>
<td>196.02 +/- 0.73</td>
<td>0.76 +/- 0.02</td>
<td>GATCAG</td>
</tr>
<tr>
<td>Human</td>
<td>6963</td>
<td>11377 +/- 1323</td>
<td>96.29 +/- 0.38</td>
<td>81.32 +/- 0.55</td>
<td>141.90 +/- 1.18</td>
<td>0.24 +/- 0.02</td>
<td>TAGCTT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Organism</th>
<th>Lane Yield (kb)</th>
<th>Clusters (PF)</th>
<th>% PF Clusters</th>
<th>% Align (PF)</th>
<th>Align Score (PF)</th>
<th>% Error Rate (PF)</th>
<th>Tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>10700</td>
<td>11042 +/- 740</td>
<td>96.24 +/- 0.39</td>
<td>81.41 +/- 0.27</td>
<td>141.75 +/- 0.71</td>
<td>0.23 +/- 0.01</td>
<td>ATCACG</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>10891</td>
<td>11240 +/- 794</td>
<td>96.73 +/- 0.34</td>
<td>73.22 +/- 0.41</td>
<td>139.84 +/- 1.24</td>
<td>0.65 +/- 0.02</td>
<td>CGATGT</td>
</tr>
<tr>
<td>E.coli</td>
<td>10820</td>
<td>11167 +/- 663</td>
<td>96.01 +/- 0.34</td>
<td>89.56 +/- 0.31</td>
<td>196.97 +/- 0.73</td>
<td>0.74 +/- 0.02</td>
<td>TTAGGC</td>
</tr>
<tr>
<td>Human</td>
<td>11822</td>
<td>12200 +/- 807</td>
<td>96.53 +/- 0.34</td>
<td>81.86 +/- 0.17</td>
<td>143.36 +/- 0.57</td>
<td>0.23 +/- 0.01</td>
<td>TGACCA</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>11943</td>
<td>12325 +/- 807</td>
<td>96.44 +/- 0.47</td>
<td>72.18 +/- 0.33</td>
<td>137.44 +/- 0.90</td>
<td>0.64 +/- 0.03</td>
<td>ACGATG</td>
</tr>
<tr>
<td>E.coli</td>
<td>11552</td>
<td>11922 +/- 767</td>
<td>95.76 +/- 0.39</td>
<td>89.54 +/- 0.29</td>
<td>196.71 +/- 0.65</td>
<td>0.75 +/- 0.02</td>
<td>GCCAAT</td>
</tr>
<tr>
<td>Human</td>
<td>10999</td>
<td>11350 +/- 818</td>
<td>96.49 +/- 0.36</td>
<td>81.97 +/- 0.33</td>
<td>143.33 +/- 0.87</td>
<td>0.22 +/- 0.01</td>
<td>CAGATC</td>
</tr>
<tr>
<td>Arabidopsis</td>
<td>11097</td>
<td>11452 +/- 763</td>
<td>96.58 +/- 0.29</td>
<td>73.04 +/- 0.37</td>
<td>139.56 +/- 0.95</td>
<td>0.64 +/- 0.02</td>
<td>ACTTGAT</td>
</tr>
<tr>
<td>E.coli</td>
<td>8855</td>
<td>9138 +/- 620</td>
<td>95.88 +/- 0.32</td>
<td>88.51 +/- 0.43</td>
<td>194.75 +/- 0.97</td>
<td>0.72 +/- 0.01</td>
<td>GATCAG</td>
</tr>
<tr>
<td>Human</td>
<td>10268</td>
<td>10596 +/- 746</td>
<td>96.34 +/- 0.43</td>
<td>81.52 +/- 0.50</td>
<td>142.59 +/- 0.98</td>
<td>0.22 +/- 0.02</td>
<td>TAGCTT</td>
</tr>
</tbody>
</table>

Barcoded adapters were used to construct manual and SPRIworks libraries

Libraries were sequenced in 2 lanes 10 MPX / lane
SPRlworks System II for Roche/454

Three size selection options:

No size selection, 350bp+, 500bp+
Run time: 3.5hr
SPRIworks Fragment Library System III
For Applied Biosystems SOLiD

Three size selection options:

No size selection, 150–250bp, 150–350bp

Run time = 3.5 hrs
Additional Applications
RNA-Seq Workflow

Day 1
- Poly(A) mRNA Capture
- Chemical Fragmentation
- Random Primed cDNA Synthesis
- Second Strand Synthesis

mRNA Capture and cDNA Synthesis

Day 2 SPRIworks
- End Repair
- A-tailing
- Adapter Ligation
- Size Selection

Library Construction

- Random Primed Full-Length cDNA sequencing method
- Day 2 of RNA-Seq process is standard fragment library prep
- Produces very small amounts of cDNA for library construction
RNA–Seq Library Construction (day 2) Manual vs SPRIworks

- Two input RNAs (UHRR and HBRR)
- Two input amounts (5µg and 1µg)
- Double stranded cDNA made using Illumina recommended protocol
- Small size selection for SPRIworks and 200-300bp gel cut for manual
- PCR amplification using Illumina protocol
Comparison of RNA-Seq Library Output

Examples of RNA-Seq libraries from SPRIworks and manual process
RNA-Seq Tag Count Analysis

Correlation between manual and SPIRiworks
UHRR (5μg)

Pearson = 0.99251
Spearman = 0.99627
N (>0 only) = 16760
N (total) = 18517

Correlation of 5μg vs 1μg input
(SPIRiworks)

Pearson = 0.99453
Spearman = 0.9957
N (>0 only) = 16738
N (total) = 18517

Summary of correlation analysis (Pearson)

<table>
<thead>
<tr>
<th></th>
<th>UHRR-5-Man</th>
<th>UHRR-1-Man</th>
<th>UHRR-5-SWs</th>
<th>UHRR-1-SWs</th>
<th>HBRR-5-Man</th>
<th>HBRR-1-Man</th>
<th>HBRR-5-SWs</th>
<th>HBRR-1-SWs</th>
</tr>
</thead>
<tbody>
<tr>
<td>UHRR-5-Man</td>
<td>0.98441</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UHRR-1-Man</td>
<td>0.98441</td>
<td>0.99251</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UHRR-5-SWs</td>
<td>0.99251</td>
<td>0.99453</td>
<td>0.99453</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UHRR-1-SWs</td>
<td>0.98377</td>
<td>0.99453</td>
<td>0.99453</td>
<td>0.97693</td>
<td>0.95477</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBRR-5-Man</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.97693</td>
<td>0.95477</td>
<td>0.98377</td>
<td>0.97771</td>
</tr>
<tr>
<td>HBRR-1-Man</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.97693</td>
<td>0.95477</td>
<td>0.98377</td>
<td>0.97771</td>
</tr>
<tr>
<td>HBRR-5-SWs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.95477</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBRR-1-SWs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.95477</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ChIP-Seq Library Construction
Manual vs SPRlworks

High sensitivity chip profile of input samples

- Two samples of ChIP DNA from p53 antibody mediated DNA capture (Vehicle control and Estradiol treated from Genpathway).
- 5ng input DNA based on pico green quantification
- Small size selection for SPRlworks and 200-300bp gel cut for manual
- PCR amplification using Illumina recommended protocol

*Note: no visible DNA in the size selection range
ChIP–Seq Library Output and Sequencing Metrics

SPRIworks vs Manual

PCR amplification of the ChIP-Seq libraries

Bioanalyzer traces of ChIP-Seq libraries

<table>
<thead>
<tr>
<th>Lane</th>
<th>Lane Yield (kbases)</th>
<th>Clusters (raw)</th>
<th>Clusters (PF)</th>
<th>1st Cycle Int (PF)</th>
<th>% PF Clusters</th>
<th>% Align (PF)</th>
<th>Alignment Score (PF)</th>
<th>% Error Rate (PF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWs-Vehicle</td>
<td>1510729</td>
<td>299074 +/- 23940</td>
<td>246851 +/- 14851</td>
<td>372 +/- 26</td>
<td>82.67 +/- 1.84</td>
<td>81.99 +/- 0.06</td>
<td>86.41 +/- 0.16</td>
<td>0.17 +/- 0.01</td>
</tr>
<tr>
<td>SWs-Estradiol</td>
<td>1476528</td>
<td>294123 +/- 23599</td>
<td>241263 +/- 14136</td>
<td>355 +/- 25</td>
<td>82.18 +/- 2.07</td>
<td>82.62 +/- 0.11</td>
<td>87.23 +/- 0.19</td>
<td>0.17 +/- 0.00</td>
</tr>
<tr>
<td>Man-Vehicle</td>
<td>1426621</td>
<td>295957 +/- 23178</td>
<td>233108 +/- 14116</td>
<td>325 +/- 35</td>
<td>79.03 +/- 3.42</td>
<td>79.39 +/- 0.91</td>
<td>83.10 +/- 1.08</td>
<td>0.19 +/- 0.01</td>
</tr>
<tr>
<td>Man-Estradiol</td>
<td>1259831</td>
<td>273232 +/- 25253</td>
<td>205855 +/- 37675</td>
<td>333 +/- 40</td>
<td>75.41 +/- 12.92</td>
<td>78.77 +/- 5.01</td>
<td>82.20 +/- 6.51</td>
<td>0.22 +/- 0.07</td>
</tr>
</tbody>
</table>
Examples of p53 Binding Sites

Integrated Genome Browser view of ChIP-Seq experiments*

Known p53 binding genes MDM2 and CDKN1A

*Genpathway ChIP-Seq analysis pipeline
SPRIworks Target Capture Library

Library Construction

End Repair

A-tailing

Adapter Ligation

Size Selection

Post LC PCR

Hybridization capture

Post Capture PCR

Library Quant

Sequencing

SPRIworks Illumina/SureSelect Library

from 3µg DNA

PE library with no size selection

*Adapted from Agilent SureSelect protocol
Illumina Sample Prep Workflow

Fragment Library Sequencing
Automating the workflow for any throughput

1-200 libraries/wk

>200 lib/wk
Scale up to Biomek user 96 well plates

Starting now with the SPRiworks chemistry will provide continuity and consistency in your results as you scale up and increase your sequencing throughput.
Roche 454 Sample Prep Workflow

Fragment Library Sequencing
Automating the workflow for any throughput

DNA Extraction → DNA Shearing → Library Construction → qPCR Norm → emPCR Setup → Oil Breaking → Bead Enrichment → Sequence

1-200 libraries/wk

>200 lib/wk
Biomek user
96 well plates

A REMe module can be incorporated onto a Biomek NxP, simplifying the emulsion PCR and oil breaking steps.

Starting now with the SPRlworks chemistry will provide continuity and consistency in your results as you scale up and increase your sequencing throughput.
SOLiD Sample Prep Workflow

Fragment Library Sequencing
Automating the workflow for any throughput

DNA Extraction → DNA Shearing → Library Construction → PCR Setup/cleanup → qPCR Norm → emPCR Setup → Oil Breaking → Bead Enrichment → Sequence

1-200 libraries/wk

SPRIworks System III

Biomek NXP- Span 8

A REMe module can be incorporated onto a Biomek NxP, simplifying the emulsion PCR and oil breaking steps

>200 lib/wk
Biomek user
96 well plates

Biomek FXP Hybrid Pre-PCR LC & Set up

Biomek NXP- Span 8

Starting now with the SPRIworks chemistry will provide continuity and consistency in your results as you scale up and increase your sequencing throughput
Conclusions

- SPRImworks enables automated reaction cleanup and size selection for all major NGS platforms
- Significantly reduces hands on prep time
- Supports several applications (RNA–Seq, ChIP–Seq, Capture)
- Sequencing performance equivalent to manual process
- Highly reproducible output
- Automation solutions available for low to high throughput