SureSelect
The Leader in Target Enrichment

Genomic Solutions Division

Yong Yi
Marketing Director

Doug Roberts
Manager, DNA Applications
1. SureSelect Technology

2. Whole Exome Sequencing

3. Summary
Maximize Sequencing Efficiency

Required Throughput = Genome Size x Average Coverage

Human Genome
3Gb x 30 = 90Gb

Target
- 50Mb x 100 = 5Gb
- 5Mb x 100 = 500Mb
- 0.5Mb x 100 = 50Mb
- 50Kb x 100 = 5Mb

Develop designs for any sequencing capacity:
- High Throughput or Desktop
SureSelect – Pioneer in NGS Target Enrichment

- Capture fragments with longest, most efficient 120-mer cRNA baits
- Probes can be designed to any regions of interest, samples can be multiplexed
- Easy to implement and compatible with validated automation solution

Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing

Andreas Gnirke¹, Alexandre Melnikov¹, Jared Maguire¹, Peter Rogov¹, Emily M LeProust², William Brockman¹,², Timothy Fennell¹, Georgia Giannoukos¹, Sheila Fisher¹, Carsten Russ¹, Stacey Gabriel¹, David B Jaffe¹, Eric S Lander¹,²,³ & Chad Nusbaum¹

Targeting genomic loci by massively parallel sequencing requires new methods to enrich templates to be sequenced. We developed a capture method that uses biotinylated RNA ‘baits’ to fish targets out of a ‘pond’ of DNA fragments. The RNA is transcribed from PCR-amplified oligodeoxynucleotides originally synthesized on a microarray, generating sufficient bait for multiple captures at concentrations high enough to drive the hybridization. We tested this method with 170-mer baits that target >15,000 coding exons (2.5 Mb) and four regions (3.7 Mb total) using Illumina sequencing as read-out. About 90% of uniquely aligning bases fell on or near bait sequence; up to 50% lay on exons proper. The uniformity was such that ~60% of target bases in the exonic ‘catch’, and ~80% in the regional catch, had at least half the mean coverage. One lane of Illumina sequence was sufficient to call high-confidence genotypes for 89% of the targeted exon space.

The development and commercialization of a new generation of increasingly powerful sequencing methodologies and instruments⁴ have lowered the cost per nucleotide of sequencing data by several orders of magnitude. Within a short time, several individual human genomes have been tested on target sets complex enough to match the scale of current next-generation sequencing instruments. The first method, microarray capture⁵,¹²,¹³, uses hybridization to arrays containing synthetic oligonucleotides that match the target DNA sequences. A second method, solution hybrid selection, uses capture of the DNA target by RNA probes selected for optimizing capture efficiency. The development of both methods has been driven by the need to efficiently sequence entire genomes or large numbers of genes in a single round of sequencing.
SureSelect – The Leader in Target Enrichment

COMPLETE Solution

Best PERFORMANCE

FLEXIBLE Designs
SureSelect - Most Complete Enrichment Solution

**All Exon Designs**
- All Exon V5
- All Exon V5+UTRs
- Non Human Exomes

**Custom Solutions**
- Custom DNA
- Custom RNA

**Targeted Panels**
- Human Methyl-Seq
- Chrm. X
- Human Kinome

**Post-capture**
- SureSelect XT Illumina
- SureSelect XT SOLiD

**Pre-capture**
- SureSelect XT2 Illumina

+ Library Prep

Agilent SureSelect Platform
For Research Use Only
SureSelect - Most Complete Workflow Solution
From Sample to Analysis

2100 Bioanalyzer

SureCycler 8800

Agilent NGS Automation

GeneSpring NGS

2200 TapeStation

MX3000

Agilent NGS Automation

Agilent Technologies
SureSelect – The Leader in Target Enrichment

COMPLETE Solution

Best PERFORMANCE

FLEXIBLE Designs
# SureSelect Technology

## The Best Performance

<table>
<thead>
<tr>
<th>Core Technology</th>
<th>Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ultra-Long RNA Baits</strong> <em>(120-mer)</em>&lt;br&gt;Binding strength&lt;br&gt;RNA:DNA &gt; DNA:DNA</td>
<td><strong>Better Sensitivity</strong>&lt;br&gt;Detect more SNP, InDels, CNV, fusions&lt;br&gt;<strong>Better Workflow</strong>&lt;br&gt;16hr hybridization, other require up to 72hrs&lt;br&gt;<strong>Better Allelic Balance</strong>&lt;br&gt;Equal representation of both alleles</td>
</tr>
</tbody>
</table>
SureSelect: Enabling Scientific Discovery

Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing

Tom Walsh\textsuperscript{a}, Ming K. Lee\textsuperscript{a}, Silvia Casadei\textsuperscript{a}, Anne M. Thornton\textsuperscript{a}, Sunday M. Stray\textsuperscript{a}, Christopher Pennin\textsuperscript{b}, Alex S. Nord\textsuperscript{b}, Jessica B. Mandell\textsuperscript{b}, Elizabeth M. Swisher\textsuperscript{b}, and Mary-Claire King\textsuperscript{b,\textdagger}

\textsuperscript{a}Agilent Technologies, \textsuperscript{b}Agence Nationale de Recherche, \textsuperscript{\textdagger}Pfizer

>350 Publications...
SureSelect – The Leader in Target Enrichment

- COMPLETE Solution
- Best PERFORMANCE
- FLEXIBLE Designs
SureSelect - Most Flexible Enrichment Solution

Custom or Catalog Solutions for DNA and RNA targets

- Proven, Refined Design Algorithms
- ISO 13485 Reagents
- Validated, Scalable Automation Solution

SureDesign – New, intuitive software for custom capture designs
SureSelect – Flexible Solution
From High Throughput to Desktop Sequencers
SureSelect All Exon V5

- Design
- Performance
- Cancer Applications
Human All Exon V5 Design and Synthesis

- Updated Content: Refseq, Gencode, UCSC, TCGA, CCDS, miRBase
- Proprietary probe selection/placement algorithms based on GC, Tm, Entropy, Mapability
- Long 120mer probes printed in Agilent’s microarray fabrication facility
  - V5 ~ 544K individual probes spanning 50.4 Mb
  - V5 +UTR ~ 882K individual probes spanning 74.6 Mb
- High GC content printed in higher replication to boost performance
- Matched for best performance with Agilent’s library preparation
- DNA probes are converted into biotinylated RNA bait
Human All Exon V5 and V5+UTRs

The Best Performing Exome
Ready for Sequencing the Next Day
The Fastest Workflow
From sample to sequencing the next day

Library Preparation 6 hours

Capture 16 hours

Wash and PCR 4 hours

Overnight Hybridization Enables Sequencer Ready Samples the Next Day
Human All Exon V5: Distribution of Read Depths

V5 (4Gb)

V5+UTR (6Gb)
Human All Exon V5 Performance Statistics

- Duplicates: 2.37%
- On-target: 65.38%
- On-target +/- 100bp: 79.40%
- 1X coverage: 99.28%
- 10X coverage: 94.07%
- 20X coverage: 81.51%
- 100x coverage: 7.49%
## The Best Performance and Fast Workflow

<table>
<thead>
<tr>
<th>Parameter</th>
<th>All Exon V5</th>
<th>All Exon V5+UTRs</th>
<th>Competitor N</th>
<th>Competitor I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Size</td>
<td>50Mb</td>
<td>75Mb</td>
<td>64Mb</td>
<td>62Mb</td>
</tr>
<tr>
<td># Genes</td>
<td>21,522</td>
<td>21,522</td>
<td>&gt;20,000</td>
<td>20,794</td>
</tr>
<tr>
<td># Targeted Exons</td>
<td>357,999</td>
<td>359,555</td>
<td>-</td>
<td>201,121</td>
</tr>
<tr>
<td>CCDS</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>RefSeq</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>GENCODE</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>miRBase</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>TCGA</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCSC</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amount of Seq. (total)</td>
<td>4Gb</td>
<td>6Gb</td>
<td>6Gb</td>
<td>10Gb</td>
</tr>
<tr>
<td>Overall Workflow</td>
<td>1.5 days</td>
<td>1.5 days</td>
<td>4.5 days</td>
<td>3.5 days</td>
</tr>
<tr>
<td>Add Custom Content</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>
The Best Performance
Outstanding coverage with less sequencing

% Reads On-Target +/-100bp

% bases >10x cov.

% bases >20x cov.

Agilent V5
Competitor N
Competitor I
Human All Exon V5: Coverage in Tumor Suppressor Genes

- Agilent displays higher depth of coverage with the same amount of sequencing (4Gb)
Human All Exon V5: Coverage in Tumor Suppressor Genes

- Agilent displays a higher depth of coverage with the same amount of sequencing (4Gb)
Example Studies: Target Enrichment in Cancer

- **Target Enrichment for Novel Discovery**
  - Diffuse Large B cell Lymphoma

- **Intratumor Heterogeneity**
  - Renal Cell Carcinoma
Diffuse Large B cell Lymphoma (DLBCL)

- Aggressive Non-Hodgkin lymphoma
- Affects 30,000 new patients in the US every year
- Current Standard of Care – R-CHOP
  - Rituximab (anti CD20), cyclophosphamide, doxorubicin, vincristine, and prednisone
- 3 year event-free survival ~60%, remaining 40% die of disease
Exome sequencing of Non-Hodgkin Lymphoma

Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing

Jens G. Lohrb, Peter Stojanovb, Michael S. Lawrencea, Daniel Audaira, Bjoern Chapuyb, Carrie Sougnezc, Peter Cruz-Gordilloa, Birgit Knoechela,b,c, Yan W. Asmanna, Susan L. Slagerb, Anne J. Novakd, Ahmet Dogana,d, Stephen M. Anseilla, Brian K. Linka, Lihua Zoub, Joshua Goulda, Gordon Saksena, Nicolas Stranskya, Claudia Rangel-Escarehod, Juan Carlos Fernandez-Lopeza, Alfredo Hidalgo-Mirandaa, Jorge Melendez-Zajglad, Enrique Hernández-Lemusa, Angela Schwarzcruz y Celsia, Ivan Imaz-Rosshandera, Akinyemi I. Olajosina, Joonil Jungb, Chandra S. Pedamallua, Eric S. Landerf,g,h,1, Thomas M. Harberrn, James R. Cerhand, Margaret A. Shippb, Gad Getza, and Todd R. Golubf,gh,i

PNAS 109 3879-3884 (2012)
Exome sequencing of DLBCL and Follicular Lymphoma (FL)

Frequent mutation of histone-modifying genes in non–Hodgkin lymphoma

Ryan D. Morin1*, Maria Mendez–Lago1*, Andrew J. Mungall1, Rodrigo Goya1, Karen L. Mungall1, Richard D. Corbett1, Nathalie A. Johnson2, Tesa M. Severson1, Readman Chiu1, Matthew Field1, Shaun Jackman1, Martin Krzywinski1, David W. Scott2, Diane L. Trinh1, Jessica Tamura–Wells1, Sa Li1, Marlo R. Fierme1, Sanja Rogic2, Malachi Griffith1, Susanna Chan1, Oleksandr Yakovenko1, Irmitraud M. Meyer3, Eric Y. Zhao1, Duane Smallus4, Michelle Moksa1, Suganthi Chittaranjan1, Lisa Rimsza4, Angela Brooks–Wilson1,5, John J. Spinelli6,7, Susana Ben–Neriah2, Barbara Meissner2, Bruce Woolcock2, Merrill Boyle1, Helen McDonald1, Angela Tam1, Yongjun Zhao1, Allen Delaney1, Thomas Zeng1, Kane Tse1, Yaron Butterfield1, Inanç Birö1, Rob Holt1, Jacqueline Schein1, Douglas E. Horsman2, Richard Moore1, Steven J. M. Jones1, Joseph M. Connors2, Martin Hirst1, Randy D. Gascoyne2,8 & Marco A. Marra1,9

Follicular lymphoma (FL) and diffuse large B-cell lymphoma (DLBCL) are the two most common non–Hodgkin lymphomas (NHLs). Here we sequenced tumour and matched normal DNA from 13 DLBCL cases and one FL case to identify genes with mutations in B-cell NHL. We analysed RNA–seq data from these and another 113 NHLs to identify genes with candidate mutations, and then re-sequenced tumour and matched normal DNA from these cases to confirm 109 genes with multiple somatic mutations. Genes with roles in histone modification were frequent targets of somatic mutation. For example, 32% of DLBCL and 89% of FL cases had somatic mutations in MLL2, which encodes a histone methyltransferase, and 11.4% and 13.4% of DLBCL and FL cases, respectively, had mutations in MEF2B, a calcium-regulated gene that cooperates with CREBBP and EP300 in acetylating histones. Our analysis suggests a previously unappreciated disruption of chromatin biology in lymphomagenesis.

Nature 476 298-303 Aug 2011
DLBCL Exome Sequencing Details

- *Lohr et al.*
  - 55 patients paired tumor and germline (normal) DNA
  - Captured with Agilent’s exome capture
  - 150-fold mean sequence coverage of targeted exons
- *Morin et al.*
  - 117 patients FL and DLBCL tumor DNA
  - Captured with Agilent’s exome capture
DLBCL mutation frequency

A

B

C

PNAS 109 3879-3884 (2012)
Agreement between two studies

- Common Pathways and Genes Identified in Both Studies
  - Known contributing genes: **CD79B, TP53, CARD11, MYD88, EZH2**
  - Suggested Pathogenic role: **MLL2, TNFRSF14, BTG1, MEF2B, GNA13**
  - Selected Common Novel findings: **HIST1H1C, PIM1**
  - Selected Novel Findings (Lohr et al): **PCLO, TMSL3**,
  - Selected Novel Findings (DLBCL, Morin et al): **SGK1, BCL2**

  Frequent nonsynonymous mutations in chromatin remodeling genes, suggesting potential cooperation in lymphomagenesis

  - HDAC inhibitors under investigation (phase I/II)

Nature 476 298-303 Aug 2011
PNAS 109 3879-3884 (2012)
Conclusions

- High Coverage of exonic content in two large cohorts of DLBCL
- Confirmed mutations in known and novel sites
- Excellent agreement between studies
- Spurred new efforts to understand potential previously unknown drivers
- Progress toward new therapeutic avenues and potentially increased survival rates
Discovery Using Target Enrichment: Renal Cell Carcinoma

**Aim:**
To characterize *intra*-tumor heterogeneity that may foster tumor evolution/adaptation

**Method:**
SureSelect Exome sequencing of:
- 4 patient samples, 7-15 biopsies
- Pre-treatment (Primary vs. Metastasis)
- Intratumor sections (R1-R9)
- Multiple metastases

Intratumor (IT) Heterogeneity Discovery Using Target Enrichment: Renal Cell Carcinoma

Results: Id and validated ubiquitous, shared, private mutations, 34% of regional shared across tumor, 1.4% false negative rate due to high depth, plot the origin of IT variants by mapping mutations that were shared or private.
Discovery Using Target Enrichment: Renal Cell Carcinoma

Results/Conclusions:

- Identified driver mutations ($VHL$, $KDM5C$, $SETD2$, $MTOR$, $PTEN$)
- Only $VHL$ was mutated ubiquitously in all analyzed regions
- Different regions of the tumor have different mutations in the very same driver gene, e.g. $MTOR$, $SETD2$
- Intra-tumor heterogeneity may contribute to failure of targeted chemotherapy
- Suggests potential need for multiple assessments of tumor
Summary

• SureSelect exome target enrichment
  • Optimized probe design/production for improved capture
  • High coverage of exons with only 4Gb of sequencing
  • Samples ready for sequencing the next day

• Enables large cancer sequencing studies
  • Allows in-depth determination of nucleotide variations
  • Driving discovery of new potential targets and therapies
  • Increasing understanding of tumor evolution
Thank You!