Instructions for Viewers

- Change the size of any window by dragging the lower left corner. Use controls in top right corner to close or maximize each window.

- What each widget does:
  - Shows the video screen
  - Opens the Ask a Question box
  - Download slides and more info
  - Facebook login
  - Twitter login (#ScienceWebinar)
  - Shows slide window
  - Shows speaker bios
  - Search Wikipedia
  - LinkedIn login
  - If you need help
Participating Experts:

Stephen C. Peiper, M.D.
Thomas Jefferson University
Philadelphia, PA

Steven M. Wolinsky, M.D.
Northwestern University
Chicago, IL
A Pathologist’s View of Genomically Informed Medicine

Stephen C. Peiper, MD
Peter A. Herbut Professor & Chair
Department of Pathology, Anatomy & Cell Biology
Thomas Jefferson University
Disclosures

• No financial interests
• Member of Roche Molecular Advisory Board
• Participated in the early access program for the 454 Jr
• Clinical applications for 454 Jr are not recommended by Roche
• Analyses were performed with IRB approval (expedited review for de-identified specimens)
Oncology Clinical Care

• Transformational role of NextGen sequencing in discovery
  – Explosive growth in genomics: insight into pathologic signaling pathways
  – Intense pharma response to generate targeted therapeutic compounds
  – Clinical trials to translate pathway inhibition to patient responses

• Critical tool in clinical (CLIA-certified) diagnostics
  – Companion diagnostics for targeted therapeutics
  – Mutational profiling to demonstrate complex signatures & tumor variation
## FDA-Approved Companion Diagnostics
(04/25/2012)

<table>
<thead>
<tr>
<th>Target (#)</th>
<th>Technology</th>
<th>Malignancy/Therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>HER2 Amplification (10)</td>
<td>In Situ (n=6)/IHC (n=4)</td>
<td>Breast Cancer/trastuzumab Gastric &amp;GE-Junction/</td>
</tr>
<tr>
<td>ALK Fusion (1)</td>
<td>In Situ</td>
<td>NSCLC/crizotinib</td>
</tr>
<tr>
<td>EGFR* (1)</td>
<td>IHC</td>
<td>Adenocarcinoma of Colon/cetuximab &amp; panitumumab</td>
</tr>
<tr>
<td>cKIT* (1)</td>
<td>IHC</td>
<td>Gastrointestinal Stromal Tumor/imatinib</td>
</tr>
<tr>
<td>BRAF V600 (1)</td>
<td>RT-PCR</td>
<td>Melanoma/vemurafinib</td>
</tr>
</tbody>
</table>
# Current Practice: Non-FDA-Approved, Commonly Used Companion Diagnostics

<table>
<thead>
<tr>
<th>Target</th>
<th>Technology</th>
<th>Malignancy/Therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCR/ABL</td>
<td>In Situ, RT-PCR</td>
<td>Chronic Myelogenous Leukemia/imatinib</td>
</tr>
<tr>
<td>KRAS Codon 12,13 Mutation</td>
<td>RT-PCR &amp; Sequencing</td>
<td>Metastatic Adenocarcinoma of Colon/cetuximab &amp; panitumumab</td>
</tr>
<tr>
<td>EGFR Exon18-20 Mutation</td>
<td>RT-PCR &amp; Sequencing</td>
<td>Non-Small Cell Lung Carcinoma/erlotinib &amp; gefitinib</td>
</tr>
<tr>
<td>BRAF V600 Mutation</td>
<td>RT-PCR &amp; Sequencing</td>
<td>Metastatic Adenocarcinoma of Colon/cetuximab &amp; panitumumab</td>
</tr>
</tbody>
</table>
### Future Prospects for Non-FDA-Approved Companion Diagnostics

<table>
<thead>
<tr>
<th>Target</th>
<th>Technology</th>
<th>Malignancy/Therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCR/ABL</td>
<td>In Situ, RT-PCR</td>
<td>Chronic Myelogenous Leukemia/imatinib</td>
</tr>
<tr>
<td>KRAS Codon 12,13</td>
<td>RT-PCR &amp; Sequencing</td>
<td>Metastatic Adenocarcinoma of Colon/cetuximab &amp; panitumumab</td>
</tr>
<tr>
<td>EGFR Exons 18-21 Mutation</td>
<td>RT-PCR &amp; Sequencing</td>
<td>Lung: NSCLC/erlotinib &amp; gefitinib</td>
</tr>
<tr>
<td>BRAF V600</td>
<td>RT-PCR &amp; Sequencing</td>
<td>Metastatic Adenocarcinoma of Colon/cetuximab &amp; panitumumab</td>
</tr>
<tr>
<td>Many</td>
<td>NextGen Sequencing</td>
<td>Multiple</td>
</tr>
</tbody>
</table>
Companion Diagnostics: Present & Future

• Present
  – Individual drugs: individual tests
  – Low throughput & expensive
  – Panels/profiling: limited clinical utility

• Future: Targeted NextGen sequencing panels
  – High throughput
  – Decreasing price
  – Profiling detects mutation signature for eligibility for (combinations of) experimental therapies
Targeted NextGen Sequencing in Diagnostic Pathology

• The testing menu for patient care
  – Companion diagnostics/Theranostics
  – Codified guidelines
  – Direct clinical/diagnostic utility

• Use of routinely processed specimens
  – FFPE: DNA from 4-6 10 um sections

• Microdissection
  – Enrichment of tumor content
  – Tumor heterogeneity

• NextGen Sequencing
  – Amplicon sequencing
  – Targeted sequencing (275 exons from 51 genes: 0.1% of exome)
62% Tumor Cells => 31% mutant KRAS
Metastatic Colon Carcinoma: KRAS Mutation

62% Tumor Cells => 31% mutant KRAS
National Comprehensive Cancer Network Guidelines for Biomarker Testing in Colorectal Carcinoma

• KRAS (c 12,13) mutation (Fall 2008)
  Non-Response to anti-EGFR Rx (Met)

• BRAF V600E mutation (Jan 2010)
  Non-Response to anti-EGFR Rx (Met)

• Mismatch Repair (Jan 2010)
  Resistance to 5FU Rx (Stage II)
Companion Diagnostics for Metastatic Colorectal Carcinoma

Prediction of poor response to anti-EGFR therapy

- *KRAS* mutation (30-40%)
- *BRAF* mutation (5-10%)
- *NRAS* mutation (3%)
- *PIK3CA* mutation, exon 20 (3%)
Familial Colorectal Cancer

- Hereditary non-polyposis colorectal carcinoma (broad clinical diagnosis of familial cases)
- **Lynch syndrome** (subset of HNPCC)
  - ~2-7% of cases (onset at early age)
  - Other malignancies (endometrial carcinoma)
  - Germline mutations, autosomal dominant
  - DNA mismatch-repair protein deficiency (IHC)
  - Microsatellite instability (MSI): functional assay
DNA Mismatch Repair Pathway

Loss in Lynch Syndrome

• MLH1 (~30%)
• PMS2 (<5%)
• MSH2 (~60%)
• MSH6 (7-10%)

IHC

• MLH1-, PMS2-
• PMS2-
• MSH2-, MSH6-
• MSH6-

• In sporadic loss may be epigenetic silencing or somatic mutation
Colorectal Adenocarcinoma: Clinical Genomic Testing

<table>
<thead>
<tr>
<th>Targeted Gene</th>
<th>Indication</th>
</tr>
</thead>
<tbody>
<tr>
<td>KRAS (Exons 2,3)</td>
<td>Resistance to anti-EGFR therapy</td>
</tr>
<tr>
<td>BRAF (Exon 16)</td>
<td>Resistance to anti-EGFR therapy</td>
</tr>
<tr>
<td>NRAS (Exons 2,3)</td>
<td>Resistance to anti-EGFR therapy</td>
</tr>
<tr>
<td>PIK3CA (Exon 20)</td>
<td>Resistance to anti-EGFR therapy</td>
</tr>
<tr>
<td>MLH1 (All Exons: 1-19)</td>
<td>Lynch Syndrome, 5FU response</td>
</tr>
<tr>
<td>PMS2 (All Exons: 1-15)</td>
<td>Lynch Syndrome, 5FU response</td>
</tr>
<tr>
<td>MSH2 (All Exons: 1-16)</td>
<td>Lynch Syndrome, 5FU response</td>
</tr>
<tr>
<td>MSH6 (All Exons: 1-10)</td>
<td>Lynch Syndrome, 5FU response</td>
</tr>
<tr>
<td>P53 (Exons 2-11)</td>
<td></td>
</tr>
<tr>
<td>APC (Exon 18)</td>
<td></td>
</tr>
<tr>
<td>RET (Exons 5, 8, 10, 11, 13, 14-6)</td>
<td>TKI?</td>
</tr>
<tr>
<td>Others</td>
<td>Low Frequency Mutations</td>
</tr>
</tbody>
</table>
Illustrated Case:
Endometrial Carcinoma in a 41 year old woman

IHC: Mismatch Repair Protein Deficiency
MLH1-/PMS2- => MLH1 mutation
MSH2+/MSH6- => MSH6 mutation
MSI: Functional Assay

Normal

Tumor

Normal

Tumor
## Targeted NextGen Sequencing

<table>
<thead>
<tr>
<th>GENE</th>
<th>RESULT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSH6</td>
<td>Deletion: Identical germline mutation detected by Sanger sequencing in CLIA laboratory</td>
</tr>
<tr>
<td>MLH1</td>
<td>Mutation: Identical germline mutation detected by Sanger sequencing in CLIA laboratory</td>
</tr>
<tr>
<td>KRAS</td>
<td>SNP, No Activating mutation</td>
</tr>
<tr>
<td>BRAF</td>
<td>No Mutation</td>
</tr>
<tr>
<td>NRAS</td>
<td>No Mutation</td>
</tr>
<tr>
<td>PI3KCA</td>
<td>No Mutation</td>
</tr>
<tr>
<td>P53</td>
<td>SNP, No Mutation</td>
</tr>
<tr>
<td>APC</td>
<td>SNP, No Mutation</td>
</tr>
<tr>
<td>RET</td>
<td>SNP, No Mutation</td>
</tr>
<tr>
<td>HSP90</td>
<td>Mutation of undetermined significance</td>
</tr>
</tbody>
</table>
# Targeted NextGen Sequencing

<table>
<thead>
<tr>
<th>GENE</th>
<th>RESULT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSH6</td>
<td>Deletion: Identical germline mutation detected by Sanger sequencing in CLIA laboratory</td>
</tr>
<tr>
<td>MLH1</td>
<td>Mutation: Identical germline mutation detected by Sanger sequencing in CLIA laboratory</td>
</tr>
<tr>
<td>KRAS</td>
<td>SNP, No Activating mutation</td>
</tr>
<tr>
<td>BRAF</td>
<td>No Mutation</td>
</tr>
<tr>
<td>NRAS</td>
<td>No Mutation</td>
</tr>
<tr>
<td>PI3KCA</td>
<td>No Mutation</td>
</tr>
<tr>
<td>P53</td>
<td>SNP, No Mutation</td>
</tr>
<tr>
<td>APC</td>
<td>SNP, No Mutation</td>
</tr>
<tr>
<td>RET</td>
<td>SNP, No Mutation</td>
</tr>
<tr>
<td>HSP90</td>
<td>Mutation of undetermined significance</td>
</tr>
<tr>
<td>MSH2</td>
<td>2 additional cases with mutations identical to germline mutations detected by Sanger sequencing in a CLIA laboratory</td>
</tr>
</tbody>
</table>
9. Identification of novel cancer mutations that are most critical to the oncogenic phenotype?
   – Single tumor types have 30-150 changes/mutations
   – Frequent/recurring mutations: Drivers
   – Low frequency mutations: identification of contributions to tumor cell biology vs bystander
     • Clinical behavior
     • Metastasis
     • Drug resistance
## Colorectal Adenocarcinoma: Mutation Signatures

<table>
<thead>
<tr>
<th></th>
<th>32</th>
<th>12</th>
<th>352</th>
<th>111</th>
<th>333</th>
<th>896</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MMR-IHC</strong></td>
<td>N/A</td>
<td>DEF</td>
<td>DEF</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>MMR-Seq</strong></td>
<td>WT</td>
<td>WT</td>
<td>MUT</td>
<td>MUT</td>
<td>MUT</td>
<td>MUT</td>
</tr>
<tr>
<td><strong>MSS/MSI</strong></td>
<td>N/A</td>
<td>N/A</td>
<td>MSI</td>
<td>N/A</td>
<td>N/A</td>
<td>MSI</td>
</tr>
<tr>
<td><strong>KRAS^</strong></td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
<td>MUT (G13D)`</td>
<td>MUT (G12D)`</td>
</tr>
<tr>
<td><strong>BRAF^</strong></td>
<td>V600E`</td>
<td>V600E`</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
</tr>
<tr>
<td><strong>APC</strong></td>
<td>WT</td>
<td>MUT</td>
<td>WT</td>
<td>MUT</td>
<td>MUT</td>
<td>MUT</td>
</tr>
<tr>
<td><strong>P53</strong></td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
<td>MUT</td>
<td>MUT (R282W)</td>
<td>MUT</td>
</tr>
<tr>
<td><strong>PI3K-CA</strong></td>
<td>WT</td>
<td>MUT*</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
</tr>
<tr>
<td><strong>Other Passenger vs Modifier</strong></td>
<td>---</td>
<td>TET2 (SS-conf)</td>
<td>MLL (SS-conf)</td>
<td>NOTCH1 (SS-conf)</td>
<td>---</td>
<td>PTEN*</td>
</tr>
</tbody>
</table>
Future Pathology Diagnosis

• Microscopic diagnosis of malignancy
• Selection of multiple tumor regions for NextGen sequencing genomic analysis
• Classification: Pathologic + Genomic
  – Tissue of origin
  – Mutational signature
  – Pathway activation
  – Target selection
  – Tumor heterogeneity
Summary

• NextGen sequencing is a high throughput technology that will revolutionize the care of patients with malignant diseases
  – Genome centers: discovery
  – Clinical laboratory: diagnosis
• Clinical study on colorectal adenocarcinomas demonstrates the power & feasibility of this approach for diagnosis
• Challenges for patient care testing:
  – CLIA-certification: validation, QA, proficiency testing
  – Bioinformatics
  – Dynamically changing menu (i.e. Lipson et al. Nat Med Feb 2012)
  – Tumor heterogeneity (i.e. Gerlinger et al. NEJM 2012 366:883)
• Mutational profiling by NextGen sequencing is ideally suited for future selection of combinations of targeted therapies
Participating Experts:

Stephen C. Peiper, M.D.
Thomas Jefferson University
Philadelphia, PA

Steven M. Wolinsky, M.D.
Northwestern University
Chicago, IL
A Systems-level Approach to Studying HIV/AIDS Susceptibility

Steven M. Wolinsky
Northwestern University
## Genome-wide association studies

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Population</th>
<th>Sample size</th>
<th>SNP location</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV RNA at set point</td>
<td>European</td>
<td>2554</td>
<td>HLA-C, HLA-B, HCP5</td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>515</td>
<td>HLA-B</td>
</tr>
<tr>
<td>HIV control</td>
<td>European</td>
<td>1712</td>
<td>HLA-C, MICA, HLA-B, HCP5, PSORS1C3</td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>1233</td>
<td>HLA-B, intergenic</td>
</tr>
<tr>
<td>Disease progression</td>
<td>European</td>
<td>1071</td>
<td>ZNRD1, RNF39</td>
</tr>
<tr>
<td>Progression to AIDS</td>
<td>European American</td>
<td>755</td>
<td>PARD3B</td>
</tr>
<tr>
<td>Long term non-progression</td>
<td>European</td>
<td>1627</td>
<td>HLA-B, HCP5</td>
</tr>
</tbody>
</table>

Pathways involved in HIV replication

Eukaryotic initiation factor

TNF/Stress Related

Double Stranded RNA Induced Gene Expression

RIG-I-like Receptor Signaling

Images From Biocarta & KEGG
A systems-level approach

- Molecular Virology
- Innate Immunity
- Mucosal Biology
- Systems-based Analysis of HIV-Host Interactions
- Mathematical Modeling
- Resequencing
Research goals

• What is the catalog of rare genetic variants and structural variations in host proteins identified by genome-wide functional screens?
• How do specific biologically relevant genomic variants influence HIV/AIDS susceptibility?
Approach

• Identification of rare variants that have putative functional consequences and therefore the largest effect size involves a strategy for:
  – judicious selection of candidate genes to examine;
  – large-scale re-sequencing at both ends of the extremes of the phenotype distribution; and
  – genotyping the selected variants across individuals.
# The natural history cohorts

## Multicenter AIDS Cohort Study (MACS)

<table>
<thead>
<tr>
<th>Category</th>
<th>Enrollment</th>
<th>Seropositive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Prevalent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Incident</td>
</tr>
<tr>
<td>Enrollment</td>
<td>6972</td>
<td>3549</td>
</tr>
<tr>
<td>Seropositive</td>
<td></td>
<td>2884</td>
</tr>
<tr>
<td></td>
<td></td>
<td>665</td>
</tr>
<tr>
<td>AIDS</td>
<td></td>
<td>1901</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1882</td>
</tr>
<tr>
<td>Highly exposed seronegative</td>
<td></td>
<td>3423</td>
</tr>
</tbody>
</table>

## Women's HIV Interagency Study (WIHS)

<table>
<thead>
<tr>
<th>Category</th>
<th>Enrollment</th>
<th>Seropositive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Prevalent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Incident</td>
</tr>
<tr>
<td>Enrollment</td>
<td>3766</td>
<td>2812</td>
</tr>
<tr>
<td>Seropositive</td>
<td></td>
<td>2791</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>AIDS</td>
<td></td>
<td>1461</td>
</tr>
<tr>
<td></td>
<td></td>
<td>715</td>
</tr>
<tr>
<td></td>
<td></td>
<td>746</td>
</tr>
<tr>
<td>Highly exposed seronegative</td>
<td></td>
<td>954</td>
</tr>
</tbody>
</table>
HLA genotyping

• Each HLA locus has many alleles, and therefore primers must amplify all alleles with the same efficiency
• HLA loci are members of multigene families, and therefore primers must be specific enough to exclude other loci in the gene family
• Allele ambiguity results when polymorphisms that distinguish alleles fall outside of the regions examined by the typing system
• Genotype ambiguity results from an inability to establish phase between closely linked polymorphisms identified by the typing system
Representative Agilent traces after Fluidigm access array amplification
High-resolution genotyping

Comparison Graph FL15_rc / Region 1 (Samples 1-48)
Strategy

• Targeted capture and massively parallel sequencing (40x coverage) of genes within the exome-containing regions identified by proteomic, genomic, and transcriptomic approaches

• Stepwise filtering approach to screen and catalog the identified genomic variants within phenotypic extremes

• Novel computational statistics to test the candidate genes

• High-throughput genotyping of the top candidates within a larger population

• Confirming the functional importance of the genetic effects
Phenotypic extremes

Rare alleles - minor allele frequency (MAF) <1%
Targeted resequencing

- Promoter variant
- Coding variant
- UTR variant
- Intrinsic variant
- Intergenic variants
- Non-coding RNA variant

synonymous SNPs, small in-frame indels, stop losses, stop-introducing SNPs, splice site disrupting SNPs, or small frameshift indels

*Nature Reviews Genetics* 12, 628 (2011)
# Targeted capture

<table>
<thead>
<tr>
<th>Design</th>
<th>Solid capture</th>
<th>Solid capture 100 bp offset</th>
<th>Optimized liquid pool 1</th>
<th>Optimized liquid pool 1 100 bp offset</th>
<th>Optimized liquid pool 2</th>
<th>Optimized liquid pool 2 100 bp offset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regions of interest</td>
<td>7531</td>
<td>7531</td>
<td>7566</td>
<td>7566</td>
<td>7566</td>
<td>7566</td>
</tr>
<tr>
<td>Length (Mb)</td>
<td>2.1</td>
<td>2.1</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
</tr>
<tr>
<td>% coverage</td>
<td>83.9</td>
<td>89.2</td>
<td>90.2</td>
<td>93.8</td>
<td>95.0</td>
<td>98.5</td>
</tr>
<tr>
<td>Targets with no coverage</td>
<td>838</td>
<td>777</td>
<td>464</td>
<td>459</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Total capture targets</td>
<td>7576</td>
<td>8301</td>
<td>8996</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total capture space (Mb)</td>
<td>3.4</td>
<td>2.7</td>
<td>2.9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sequence quality

• Depth and evenness of coverage, read length, and read quality
• Analysis of the quality score and/or the k-mer frequency
• Distribution of estimated accuracies for raw base calls (*phred*-like scores)
• Systematic error patterns
Concordance with GWAS data
Functional classification of variants

• Cross-reference the potentially functional variants with the Human Gene Mutation Database (HGMD) and the Online Mendelian Inheritance in Man (OMIM) database

• Predict the functional effects of protein coding changes with algorithms that account for evolutionary, biochemical, and structural considerations

• Use exploratory hierarchical cluster and principal component analyses to identify and quantify potential batch effects or other technical and biological artifacts—namely, population stratification and cryptic relatedness
Data analysis pipeline

Primary Data Analysis
- Raw image Processing
- Signal Processing
- Base calling

Secondary Data Analysis
- Mapping
- Assembly
- Mapping tool with references
- Mapping tool without references
- BWA
- MAQ
- SHRiMP
- Phrap
- SSAHA2
- Bowtie
- Newbler
- Cortex
- MIRA

Application
- SNP calling
- Indel
- Structural variation
- LD
- GATK
- Cortex
- ssahaSNP
- PLINK
- PyroBayes
The compute cloud

Internet Firewall

Workstations

Lab Network (GbE)

Cluster Gateway

DNA Sequencer Network (1/10 GbE)

Management Network (GbE)

Post—Analysis

DNA Sequence Processing

Login & Gateway

System Management

Disk & File System (GPFS)

Cluster Network (10GbE)
Acknowledgements

Systems biology

HINT

WIHS

MACS

Kings

Salk

NU

JHU

UCLA

Sinai

UC

USC

Pitt

UCSD

UCSD

SBI

GU

MMC

Penn

SBI

UC

UIC

USC

GU

MMC

UC

SBI

UC

UIC

USC

GU

MMC

UC

SBI
Next Generation Sequencing and Translational Research: The Express Lane from Bench to Bedside

9 May, 2012

Participating Experts:

Stephen C. Peiper. M.D.
Thomas Jefferson University
Philadelphia, PA

Steven M. Wolinsky, M.D.
Northwestern University
Chicago, IL

Q&A
To submit your questions, type them into the text box and click Submit.
Look out for more webinars in the series at:

www.sciencemag.org/webinar

To provide feedback on this webinar, please e-mail your comments to webinar@aaas.org

For related information on this webinar topic, go to:

www.gsjunior.com