

Standards in the Life Sciences: NIST Supporting NCI

Laurie E. Locascio, Ph.D.
Director, Material Measurement Laboratory

NIST
National Institute of Standards and Technology
U.S. Department of Commerce

MATERIAL MEASUREMENT LABORATORY

NIST is…

The National Metrology Institute

working toward global harmonization and traceability to the SI

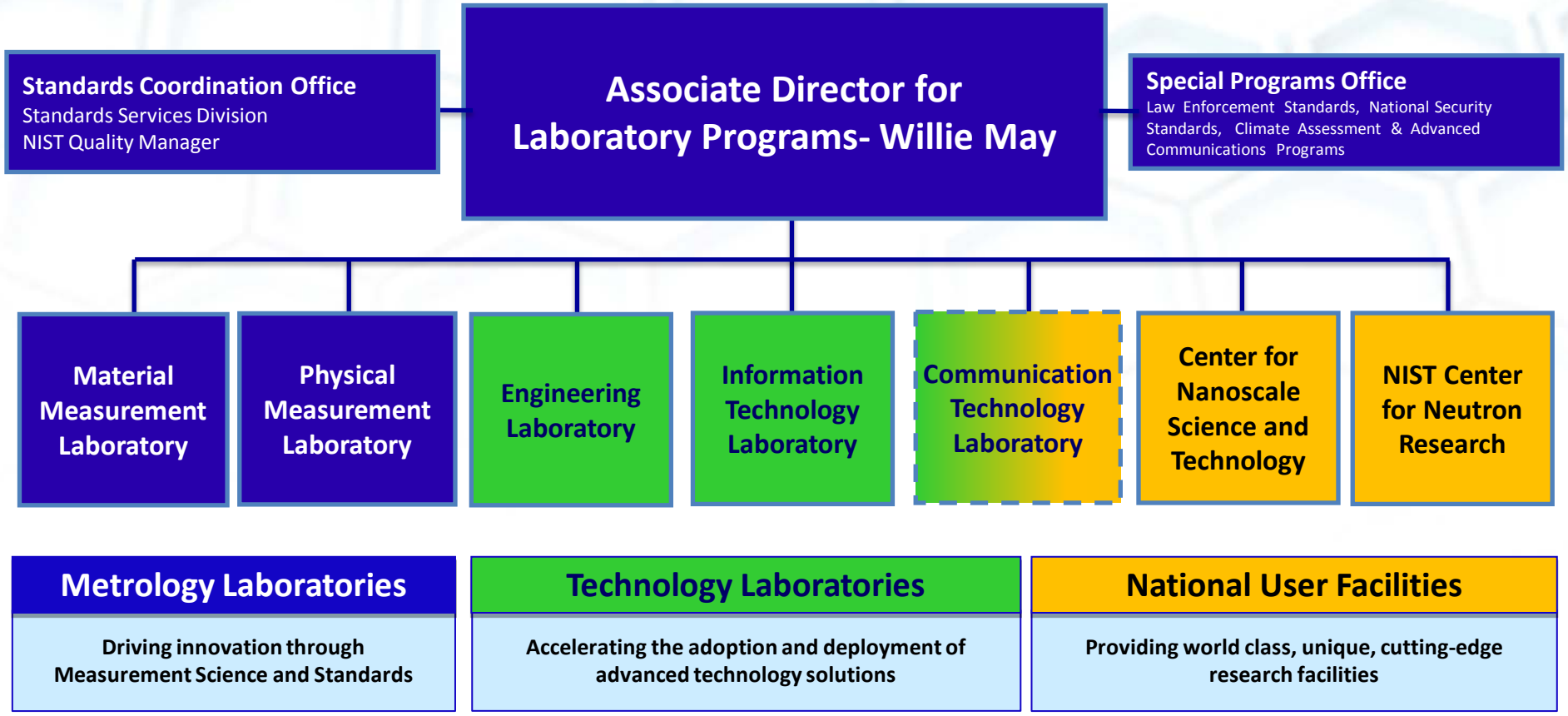
"Industry's National Laboratory"

partnering/serving industry to help maintain US leadership in science and technology products

Department of Commerce

developing standards to support international trade and commerce

NIST Laboratory Programs



MML is responsible for Biosciences, Chemical Sciences, Materials Science and Engineering with ~1000 staff on 8 campuses across the US

NIST: A Premier Scientific Institution

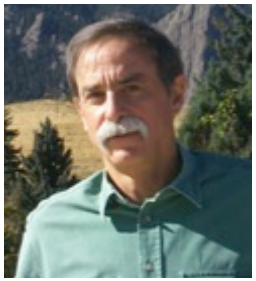
A world-leading measurement science and standards program

- **Work resulting in 4 + 1 Nobel Prizes since 1997**
- **Kyoto Prize winner in 2011**
- **MacArthur Fellowship winner in 2003**
- **National Medal of Science winners in 1998 and 2007**
- **~ 60 National Academy Members (10 current)**
- **~120 National Society Fellows**
- **~60 National/International Awards/yr**

Bill Phillips
1997 Nobel Prize

Eric Cornell
2001 Nobel Prize

John Hall
2005 Nobel Prize



David Wineland
2007 National Medal of Science and 2012 Nobel Prize

John Cahn
2011 Kyoto Prize

Dan Shechtman
2011 Nobel Prize *in Chemistry*

STANDARDS SUPPORT DISCOVERY AND INNOVATION

Standards ensure

- Validity of data
- Confidence in and comparability of data
- More rapid, integrated technology development

Reference materials

Documentary standards

Beyond Reproducibility

It is important that the public have confidence in the scientific method, and that all researchers, research reviewers and funders have a good understanding of the hallmarks of scientific investigations that produce results with a high level of confidence.

Foundations of Traditional Metrology

Systematic approaches for

- Traceability
- Measurement Uncertainty
- Method Validation

Supports measurement assurance

- confidence in quality of results
- confidence in comparability of results

Complete for physical sciences but what about biosciences?

NIST DOESN'T DRIVE THE DEVELOPMENT OF STANDARDS, THE COMMUNITY DOES

NIST Supports Confidence in Bio-Measurements with Other Agencies

- NCI, Early Detection Research Network : Interlaboratory comparisons, standards, data analysis protocols for miRNA, genetic biomarkers
- NCI, Clinical Proteomics Tumor Analysis Consortium: 'Common data analysis pipeline' (CDAP) for mass spectrometry consortium data
- NIH/NIDCR: Standard Reference Materials (SRMs) for measurement assurance of calcium phosphate based biomaterials; standardized measurements for improved dental materials and oral health care
- DHS Office of Science and Technology: Engineered yeast with known genetic sequences for testing detection of unknowns
- CDRH, FDA: genomic reference materials for personalized medicine through the Genome in a Bottle consortium; FDA used pilot NIST material to aid in approving the first next generation sequencer
- DARPA Living Foundries program: Standards for Synthetic Biology

Engaging with NIH to Address Irreproducibility

- Opportunities with NSTC LifeScience Subcommittee to develop consensus plans with other Agencies.
- NIST organizing a WH Symposium on confidence in data for innovation and data sharing (early 2016) working with Jon Lorsch (NIGMS) and Jim Olds (NSF-BIO).
- NIST engaging in discussions with NIH/NIGMS and others on collaborative work in training led by Jon Lorsch and Mike Rogers.
- NIST communicating with Phil Bourne (NIH Associate Director for Data Science) regarding quality of data and other data issues.

Advancing Measurement Assurance of Key Technologies in Biological Sciences

- Genetic/ genomic measurement confidence
  - External RNA Controls Consortium (ERCC and ERCC 2.0): standards for DNA microarray measurements
  - Genome in a Bottle Consortium: protocols and materials for whole genome sequencing
- Cell line Authentication: genomic sequence markers for unambiguous identification of non-human cell lines used for basic research and production of biopharmaceuticals, including monkey, mouse (commercial kit in development), Chinese hamster, rat, dog, others
- Cell therapy product QC: internal fluorescence reference method for quantitative comparison of expression of markers for pluripotency; stem cell qualification imaging methods with NIH and others
- First international reference cell standard for CD4+ cell counting for HIV/AIDS monitoring (WHO/BS/10.2153) for flow cytometry (with 2 other institutes)

External RNA Controls Consortium (ERCC and ERCC 2.0)

Reference methods and data tools providing confidence in genome-scale measurements

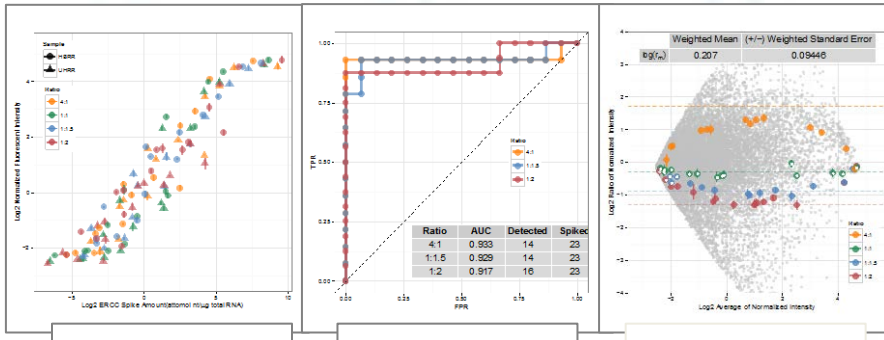
- NIST-led consortium with more than 90 public, private, academic partners

Outputs:

- SRM 2374 DNA Sequence Library for External RNA controls
- ERCC RNA Spike-In Mix-Life Technologies
- ERCC dashboard: New software tool for turnkey assessment of the technical performance of gene expression experiments
- ERCC 2.0 creating new suite of RNA controls to include: transcript isoform, small and mi RNA, long and non-coding RNA controls

NIST SRM 2374 jointly developed with the ERCC

erccdashboard: Performance Metrics



Dynamic Range | Diagnostic Performance | Ratio Bias & Variability

Contact: Marc Saliit, marc.salit@nist.gov)

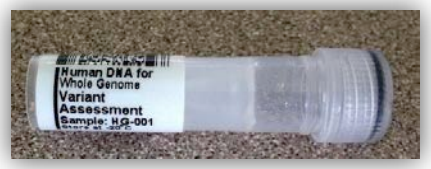
Genome in a Bottle consortium

Next generation sequencing rapidly adopted for clinical applications, but currently no established method to assess the accuracy of sequencing results

- NIST led consortium with more than 75 public, private, academic partners
- First whole human genome reference material (RM) released soon
- Using NIST-GIAB standard genotypes determined through the consortium, FDA approved the first high-throughput DNA sequencer
- High interest in March 2014 Nature Biotechnology publication: Altmetric score in 99th percentile for twitter/news/blog posts

Genome in a Bottle Consortium

Concordance needed within Whole Genome Sequencing Technologies and Bioinformatics Platforms



8300 tubes acquired for NIST RM: 1st of 5+ genomes for GIAB Consortium

Contact: Justin Zook (justin.zook@nist.gov)

Partnering with NIH through EDRN

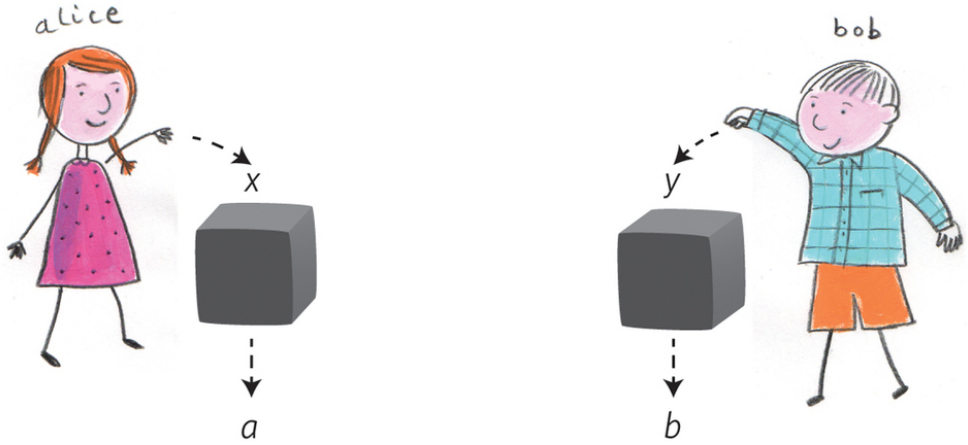
- Collaboration initiated in mid 2000's
- NCI early to consider external validation as critical for biomarker discovery
- As technologies become more quantitative (mass spectrometry, PCR, NGS…) the role of a partnership between discovery scientists and metrologists becomes even more clear

Finding our sweet spot within EDRN

Technologies rapidly changing
Measurands (analytes) rapidly changing

Why is there resistance from some in the discovery science community?

There are many sources of analytical errors and it doesn't imply 'bad science'



- Differences in analytical techniques
- Differences in statistical methods (study designs)
- Unintentional selective reporting
- Incomplete protocol reporting
- Lack of appropriate specimens and reagents
- Variations in interpretation
- Bias, chance and overfitting
- Lack of appropriate controls
- Lack of knowledge in laboratory tests into clinical tests
- Scientific misconduct

NIST and EDRN

NIST IAA objective is to increase measurement accuracy and reproducibility of measurements on new biomarkers for cancer.

Developing reference materials:

- Well-characterized materials with certified values to ensure that procedures and analytical methods are working correctly.

Conducting interlaboratory testing:

- Allows EDRN labs to learn about the measurement processes, benchmark how they compare to others.

Assay validation and identifying areas that impact reproducibility:

- Comprehensive measurement methods using proper controls, reference materials, biostatistics, and informatics.

Reference Materials: Mixed-Tissue Reference Samples

- High-content 'omics technologies measure thousands of different biomolecules simultaneously.
- This scale is impractical to calibrate.

Approach:

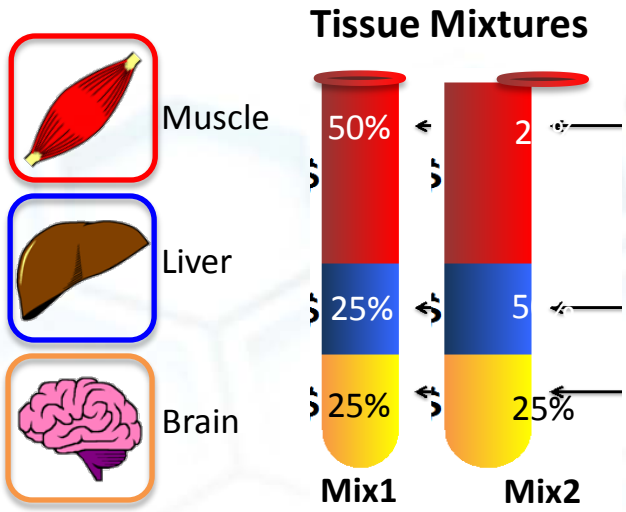
- Use tissue mixture fractions as surrogate for concentration.
- Use *in silico* modeling to predict mixture fraction from RNA-Seq signals of neat components and mixtures across the population of biomolecules.

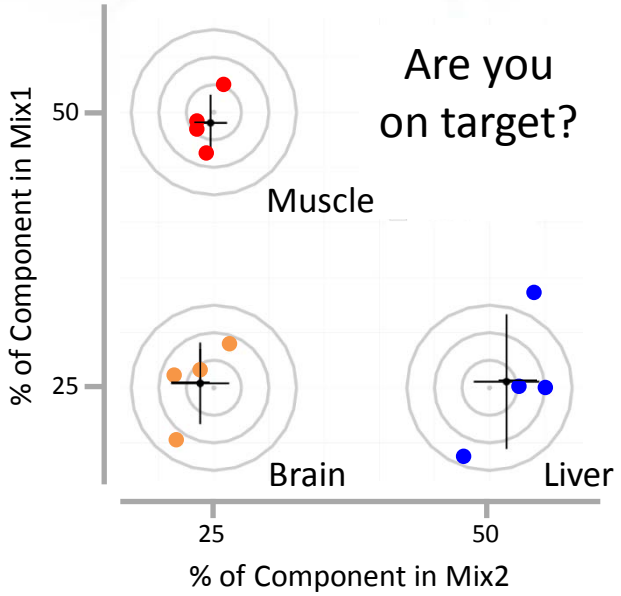
Accomplishment:

- Demonstrated suitability of mixtures for RNA-Seq
- Development of figures of merit to identify outliers, chart performance, compare methods

Future:

- Develop a standard mixture set with well-characterized RNA content for use as process controls, technology development, optimization





JR Parsons, PS Pine, SA Munro, and M Salit

Interlaboratory Studies: miRNA

- A variety of technologies have been developed for high content profiling of differentially expressed microRNA.
- Assessing repeatability and reproducibility among laboratories with various technologies allows labs to benchmark results and optimize performance.

Approach:

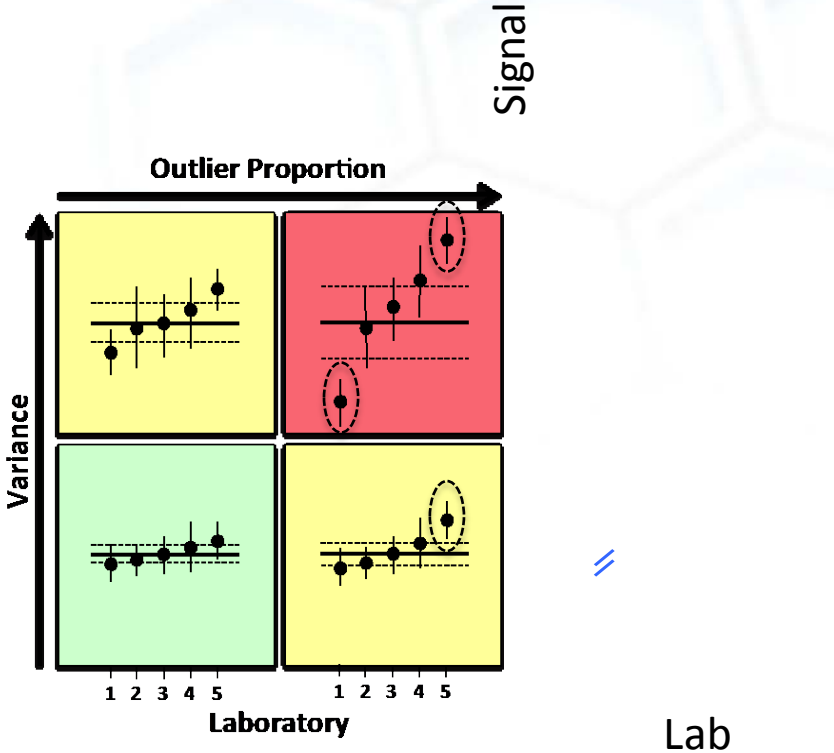
- Share a set of RNA reference samples with labs
- Enable labs to compare their results to consensus, regardless of technology

Accomplishment:

- Developed metrics and figures of merit for comparisons

Future:

- Extend paired tissue paradigm to other 'omics technologies



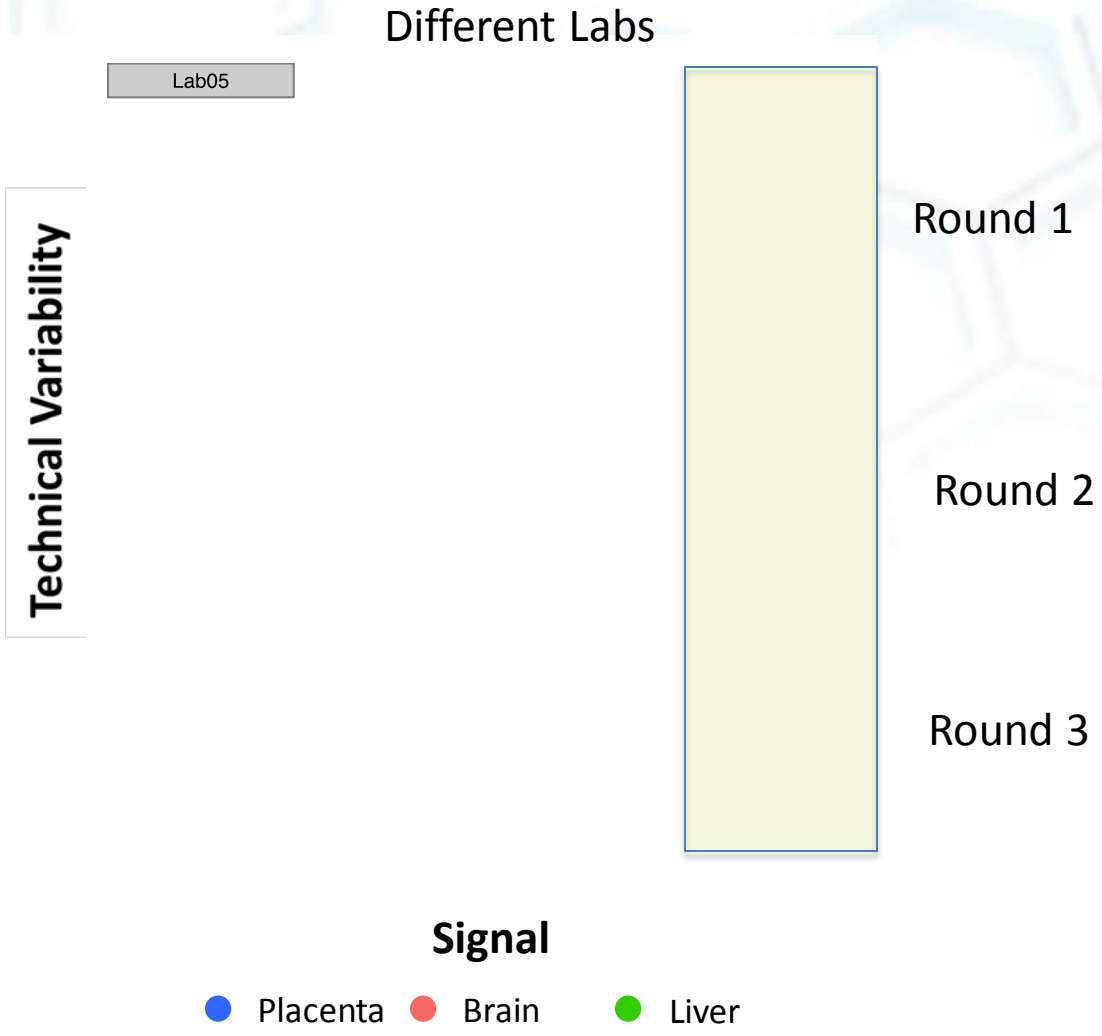
Scott Pine and Marc Salit

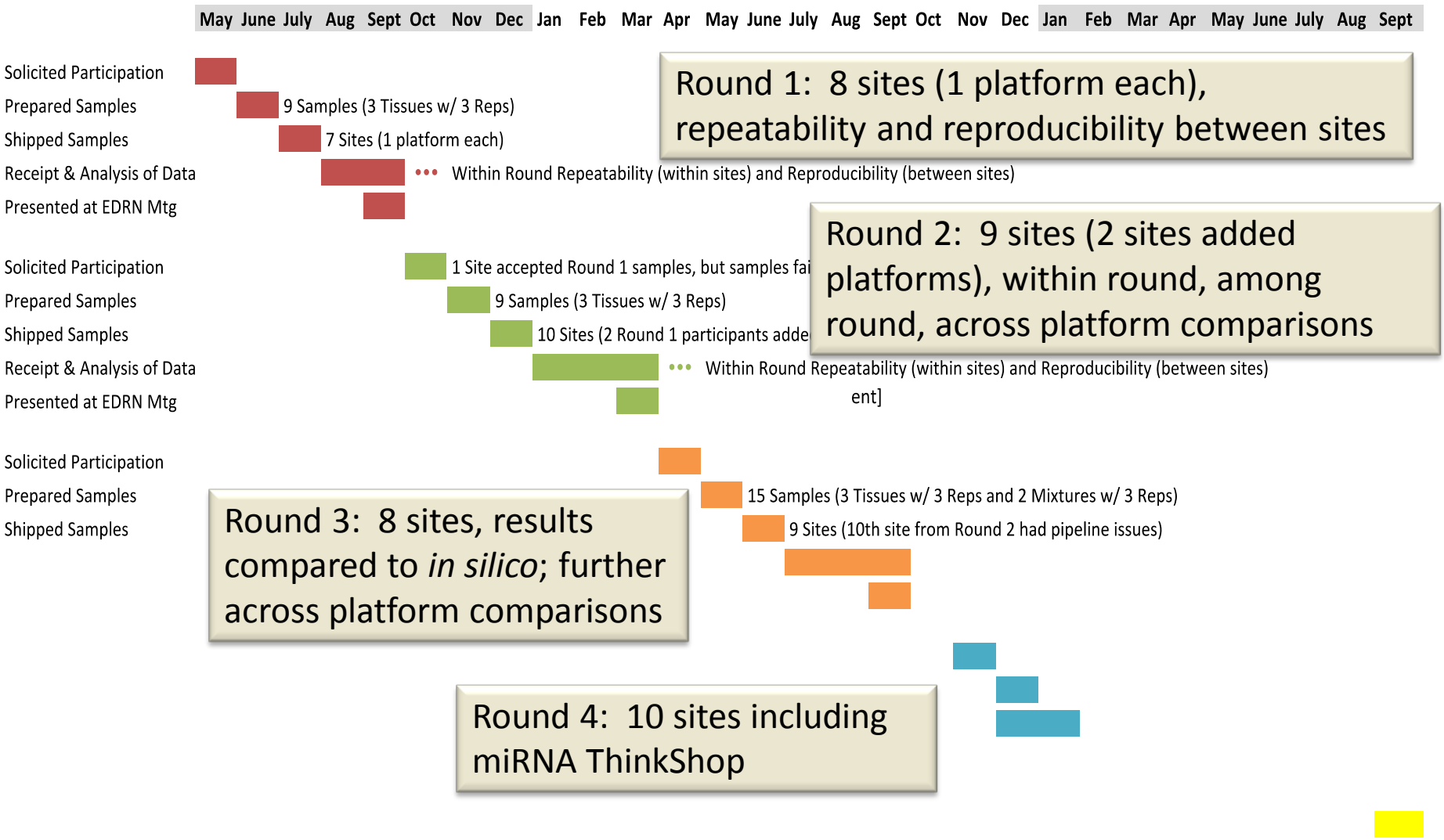
Results: Performance of Individual Labs Technical Variability (Rounds 1-3)

Evaluated technical variance as a function of signal across the dynamic range of high content platforms

Identify potential issues with sites, samples, or measurement system limits.

Labs can improve with successive rounds




May June July Aug Sept Oct Nov Dec Jan Feb Mar Apr May June July Aug Sept Oct Nov Dec Jan Feb Mar Apr May June July Aug Sept
Solicited Participation
Prepared Samples
9 Samples (3 Tissues w/ 3 Reps)
Shipped Samples
7 Sites (1 platform each)
Receipt & Analysis of Data
Within Round Repeatability (within sites) and Reproducibility (between sites)
Presented at EDRN Mtg
Round 1: 8 sites (1 platform each), repeatability and reproducibility between sites
Solicited Participation
1 Site accepted Round 1 samples, but samples fai
Prepared Samples
9 Samples (3 Tissues w/ 3 Reps)
Shipped Samples
10 Sites (2 Round 1 participants adde
Receipt & Analysis of Data
Within Round Repeatability (within sites) and Reproducibility (between sites)
ent]
Presented at EDRN Mtg
Round 2: 9 sites (2 sites added platforms), within round, among round, across platform comparisons
Solicited Participation
Prepared Samples
15 Samples (3 Tissues w/ 3 Reps and 2 Mixtures w/ 3 Reps)
Shipped Samples
9 Sites (10th site from Round 2 had pipeline issues)
Round 3: 8 sites, results compared to *in silico*; further across platform comparisons
Round 4: 10 sites including miRNA ThinkShop

Assay Validation: Mitochondrial DNA Biomarkers for Prostate Cancer

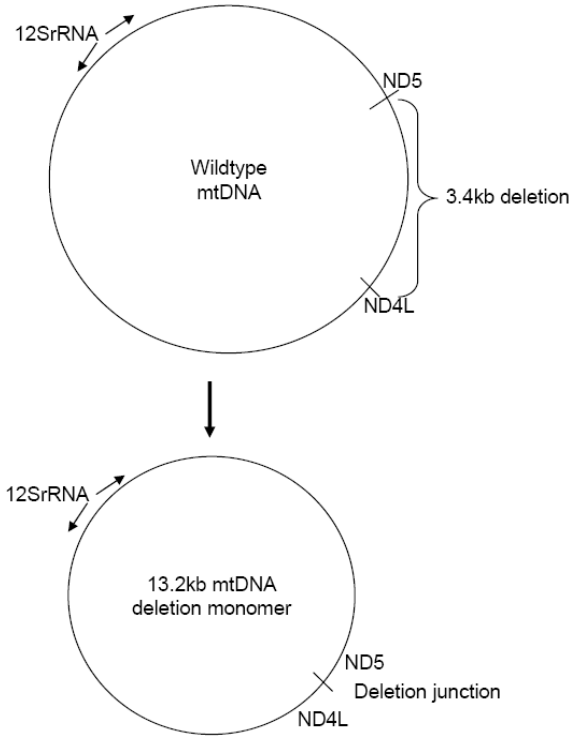
Specific biomarker validation methods are needed

Approach:

- Conduct confirmation studies for accurate, repeatable and reproducible detection of mitochondrial DNA copy number and a 3.4 kb deletion in patient samples as potential prostate cancer early detection biomarkers

Accomplishment:

- NIST modified assay for mitochondrial DNA detection to correct serious study design flaws
- Generated assay, data handling and results interpretation schemes for assessing characteristics of mitochondrial DNA as a potential prostate cancer biomarker using a qPCR assay
- First report that a 3.4 kb mitochondrial DNA deletion and increase in mitochondrial DNA copy number as potential prostate cancer biomarkers can be detected in the body fluids urine and serum



Samantha Maragh, Steve Lund

Validation of Mitochondrial DNA Biomarkers for Prostate Cancer and Assessment in Urine

NIST confirmation study

	sensitivity %	specificity %	p-value signal
mtDNA Deletion positive			
Tissue	63.2 (12/19)	65 (13/20)	0.045*
Urine	70 (14/20)	60 (12/20)	0.128
Serum	73.7 (14/19)	33.3 (5/15)	0.384
Increased fold mtDNA genomes			
Tissue	68.4 (13/19)	60 (12/20)	0.07+
Urine	64.3 (9/14)	86.7 (13/15)	0.006**
Serum	57.1 (8/14)	71.4 (10/14)	0.204
Both biomarkers combined			
Tissue	73.7 (14/19)	65 (13/20)	0.012*
Urine	64.3 (9/14)	100 (15/15)	0.020*
Serum	64.3 (9/14)	64.3 (9/14)	0.125

Biomarker performance per individual. Plus sign, significant at p-value $\leq$ 0.1; single asterisk, p-value $\leq$ 0.05; double asterisk, p-value $\leq$ 0.01. Sensitivity, number of cancer that are predicted positive / total cancer; specificity, number of non-cancer that are predicted negative / total non-cancer.

ACKNOWLEDGEMENTS

NIH NCI EDRN PROGRAM

NIH NCI CPTAC PROGRAM

NCI:
Sudhir Srivastava
Lynn Sorbara
Paul Wagner
Jo Ann Rinaudo

NIST:
Marc Salit
Scott Pine
Sarah Munro
Samantha Maragh
Steve Choquette
Hua Jun He
Ken Cole
Steve Lund
JR Parsons
Anne Plant

Other NCI Collaborations: Clinical Proteomics Tumor Analysis Consortium

J. Roth, S. Markey, S. Stein

Problem & Significance:

- What are the proteomic changes present in cancers?
- Do DNA/RNA variations correlate with protein levels and does phosphorylation correlated with tumor type.
- Results can depend greatly on data analysis details.

NIST Approach:

- Create a 'common data analysis pipeline' (CDAP) for processing mass spectrometry-based consortium data

Accomplishment:

- First version of CDAP completed and first set of tumor samples and reference standards have been analyzed and results make publicly available.
- Created a gene-based protein reporting system for compatibility with popular gene-based analysis tools.

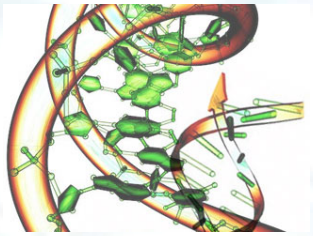
Future:

- Refine data analysis for analysis of next set of tumor results from CPTAC and affiliate labs.
- Create single, consistent gene-based protein results across all tumor types (colon, breast, ovarian) and analysis.

NATIONAL CANCER INSTITUTE

MASS SPECTROMETRY DATA CENTER

The Cancer Genome Atlas



Samples + tumor-specific DNA and RNA sequences

↓

OFFICE OF CANCER CLINICAL PROTEOMICS RESEARCH

MS-based proteomics datasets from 6 centers

↓

NIST

Data analysis to generate results for public distribution

Project 3: Cell-Free DNA Measurements for Cancer

Approach:

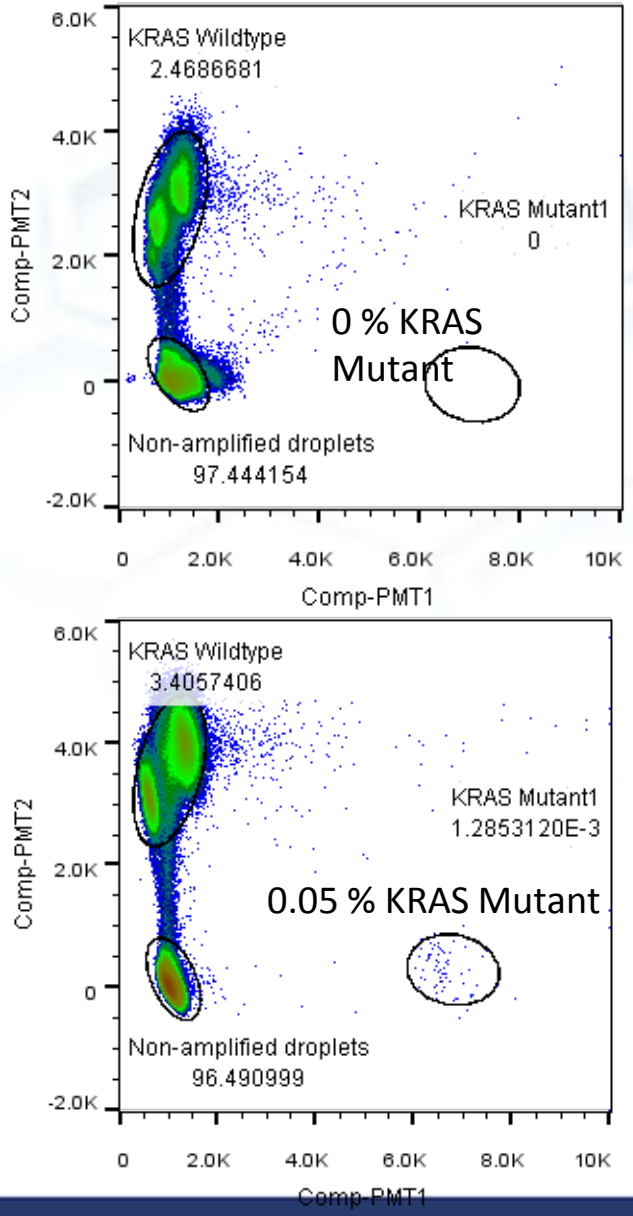
- Work with EDRN labs to improve the confidence and reliability of measurement of DNA released from cancer cells into blood or urine through interlaboratory testing and development of cancer reference materials

Accomplishment:

- Validation of cancer DNA mutations at very low concentrations in NIST labs; recruiting EDRN labs for testing samples with initial DNA biomarkers

Future:

- Work with the EDRN labs to define the relevant mutations; validate assays at low concentrations
- Refine the requirements for reference materials that will be used to ensure the accurate and reliable measurements



Hua-Jun He, Samantha Maragh, Steve Choquette, Kenneth Cole

Validation of Mitochondrial DNA Biomarkers for Prostate Cancer and Assessment in Urine

mtDNA deletion in prostate tissue original report

	Data not normalized	
	original report	(not reported)
	per biopsy	per individual
Sensitivity	83%	55.5%
Specificity	79%	37.5%

	Data normalized by NIST	
	Re-analysis of original data	
	per biopsy	per individual
Sensitivity	59.5%	77.8%
Specificity	92.5%	87.5%

NIST confirmation study

	sensitivity %	specificity %	p-value signal
mtDNA Deletion positive			
Tissue	63.2 (12/19)	65 (13/20)	0.045*
Urine	70 (14/20)	60 (12/20)	0.128
Serum	73.7 (14/19)	33.3 (5/15)	0.384
Increased fold mtDNA genomes			
Tissue	68.4 (13/19)	60 (12/20)	0.07+
Urine	64.3 (9/14)	86.7 (13/15)	0.006**
Serum	57.1 (8/14)	71.4 (10/14)	0.204
Both biomarkers combined			
Tissue	73.7 (14/19)	65 (13/20)	0.012*
Urine	64.3 (9/14)	100 (15/15)	0.020*
Serum	64.3 (9/14)	64.3 (9/14)	0.125

Biomarker performance per individual. Plus sign, significant at p-value $\leq$ 0.1; single asterisk, p-value $\leq$ 0.05; double asterisk, p-value $\leq$ 0.01. Sensitivity, number of cancer that are predicted positive / total cancer; specificity, number of non-cancer that are predicted negative / total non-cancer.

Project 4: Circulating tumor cell capture

Problem:

- CTCs are cancer cells shed by a tumor into the blood, and their abundance correlates with disease progression
- Many CTC platforms for analysis, but no agreement yet in their measurements

Approach: enable comparability

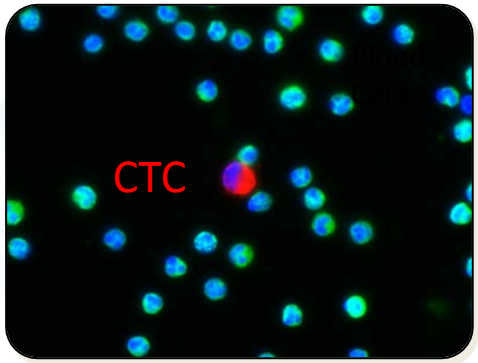
- Understand technology, needs, technology roadblocks
- Develop complementary technologies to enable comparisons between platforms

Accomplishments:

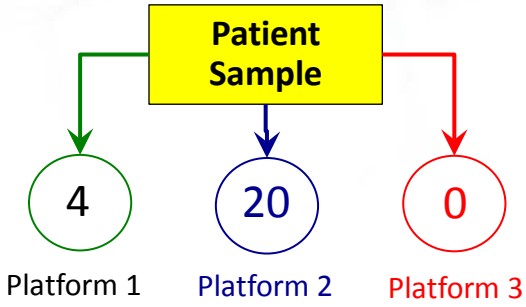
- Developed new platform using immunomagentic bead capture in microfluidics-capture efficiency improved over existing technology
- Could not entice companies to do engage in comparisons with nascent technologies in a highly competitive field

Future

- Where technology is today: interest in personalized medicine and post capture analysis



Rare CTC amongst white blood cells
Epic Sciences



Customers and Partners

Interagency Agreement 2010-2015

Project 1: microRNA measurement quality assurance

Project 2: Cancer biomarkers in urine

Project 3: Cell Free DNA measurements

Project 4: Capture and analysis of single circulating tumor cells

…for Precision Medicine

Need to work at “genome-scale”

- traceability impractical
- need new models for uncertainty for base calls

Need to address qualitative properties such as sequence where current methods to assign uncertainty are flawed

Work with Other Agencies on International Standards

- ASTM Tissue Engineered Medical Products: Documentary standards for assuring comparability and reproducibility of products for tissue engineering and regenerative medicine, with FDA.

- ISO TC 276 Biotechnology: NIST serves as Chair of US Delegation. Documentary standards for biobanking, bioanalytical methods, and bioprocessing, with FDA, USDA, USTR, NIH, DOE, ITA, and US industry reps.

Results: Performance of Individual Labs Reproducibility

Evaluated *reproducibility* among labs: detection of designed-in ratios as a function of Lab.

In silico modeling using data from pure samples allows for prediction of designed-in values (open circles)