Steps towards metrology for quantitative biology

Marc Salit and David Duewer
Materials Measurement Laboratory

SIM Metrology Course

National Institute of Standards and Technology
October 2013
Metrology for Physical and Chemical Properties...
Measuring Biological Properties...
What makes Metrology in Biology different?

What is Biology?
• Biology is a natural science concerned with the study of life and living organisms, including their structure, function, growth, evolution, distribution, and taxonomy.

What do we measure?
• Properties over many length scales
  – $\sim 10^9$ range
• Dynamic properties over many time scales
  – $10^{10}$ range
• Many properties at once
  – genomes can have $\sim 10^9$ elements
• Abundances of many things
  – proteomes can have $\sim 10^{15}$ dynamic range

From Wikipedia entry for “Biology”
What Properties?

Property Classes
- Morphology
- Physical Activity
- Metabolic Activity
  - dynamic biochemical activity
- Interactions with other entities
- Amount
- Identity
  - from organism taxonomy to nucleotide identity to sequence identity...

Over what space??
Complex Space for Metrology
Where do we start?

• One starting point is Molecular Biology
  – molecular basis of biological activity
• Interactions amongst systems in cells
• Can we measure key elements in the molecular “plan” underlying biological phenomena?

“Central Dogma” of molecular biology
‘Omics

Omics aims at the collective characterization and quantification of pools of biological molecules that translate into the structure, function, and dynamics of an organism or organisms.

• typically involves measurement of a “complete” cohort of biomolecules to characterize a biological state
• genomes began to be reported first
  – “templates” for living matter
• transcriptomes began to be measured
  – “gene expression”

Wikipedia entry for Omics accessed 31 Oct 2013
What is Gene Expression?

- A caterpillar and a butterfly have identical genomes!
  - The difference between them arises from which genes are expressed, and when and where they are expressed.
  - From a genome of billions of base pairs and tens of thousands of genes, the subset expressed in a cell confer its nature.

- Quantitative biology in the post-genomic world
How is Gene Expression profiling used?

- **Systems Biology**
  - research phase

- **Toxicogenomics**
  - appearing in clinical trials
  - environmental application

- **Pharmacogenomics**
  - appearing in clinical trials

- **Clinical Diagnostics**
  - devices being developed

- **Genomic Medicine**
  - expected to be an underlying technology

Levels of observation:
- Organism
- Organ
- Tissue
- Cell
- Protein
- RNA
- Chromosomal DNA
Measurement Problems

- **Comparability of results**
  - Traceability
  - Validation
  - Uncertainty

- **Management of variability**
  - Across platforms

- **Truth**
  - Quantitative accuracy
  - Specificity

- **Informatics**
  - High dimensionality
    - Many more genes than samples
    - Presents statistics challenges
  - Statistics alone are insufficient to deliver biological interpretation

---

**Evaluation of gene expression measurements from commercial microarray platforms**

Paul K. Tan, Thomas J. Downey, Edward L. Spitznagel Jr, Pin Xu, Dadin Fu, Dimitar S. Dimitrov, Richard A. Lempicksi, Bruce M. Raaka, and Margaret C. Cam

Microarray Core Laboratory, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health; Perkin-Elmer Incorporated; 
Department of Mathematics, Washington University, Laboratory of Experimental and Computational Biology (LECB), National Cancer Institute, NIH; Laboratory of Allergy and Infectious Diseases (NIAID), NH, SA/C-Fredrick's, Inc.; Clinical Endocrinology Branch, NIDDK, NIH, USA
Measurement Strategies

- QRT-PCR
  - only a few RNA transcripts at a time

- Microarrays
  - proposed NIST focus
  - massively multiplexed RNA measurements
    - high-throughput, combinatorial approach
    - “bio-nano-informatics”
    - develop, deploy infrastructure
  - technology also applies to
    - DNA sequencing
    - proteomics/protein sequencing

TIGR 32,000 element Human array
Microarrays

- Affinity hybridization
- Spectroscopic detection
- Image analysis
- Statistical analysis of massively multiplexed data
  - To identify gene expression patterns of interest
Microarrays

- Affinity hybridization
- Spectroscopic detection
- Image analysis
- Statistical analysis of massively multiplexed data
  - To identify gene expression patterns of interest
Microarrays

- Affinity hybridization
- Spectroscopic detection
- Image analysis
- Statistical analysis of massively multiplexed data
  - To identify gene expression patterns of interest
Microarrays

- Affinity hybridization
- Spectroscopic detection
- Image analysis
- Statistical analysis of massively multiplexed data
  - To identify gene expression patterns of interest
What does a microarray experiment look like?

- RNA isolation
- Clean-up, labeling, hybridization
- Imaging of spatially arrayed signals
- Image feature extraction
- Informatics to assign gene expression measures
- Bioinformatics to elucidate biological insight
Worldview

- Three legged stool of metrology
  - Traceability
  - Measurement Uncertainty
  - Method Validation
Traceability

• tying results to a common reference
  – usually realized with calibration
  – enables comparison of results amongst those using the common reference
    • across space and time

• biological measurement results are often traceable to the experimental control
  – enrichment of a molecular signal
VIM 2 Definition

Traceability

• "property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties."
Traceability

- "property of the result of a measurement or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties."
Comparing results

• results are only useful when compared
  – to other results
    • e.g., to observe a trend
  – to limits
    • e.g., a threshold for action
  – different results in different places or measured at different times...
    • “comparability over space-and-time”
Comparability of results

- Whole and sole goal of traceability.
  - raison d'être!
- results linked to a common reference can be compared
- scope of reference defines scope of comparability
  - global network
    - SI
Comparability of results

• Whole and sole goal of traceability.
  – raison d'être!

• In biology, we use a ‘control’ as our reference, and compare against that control to “measure” biological effects
  – experiment design establishes traceability
  – scope of traceability is to the control
Emerging Challenges in Metrology for Biological Measurements

- **Commutability**
  - measurand “mismatch”
  - typically caused by bias arising from matrix

- **Speciation**
  - specificity matters

- **Unstable measurands & dynamic processes**

- **Complex and multiplexed measurands**
  - artifacts, biostatistics

- **Qualitative properties**
  - traceability?
  - measurement uncertainty?

- **Measurement of stochastic processes**
  - population studies poorly, misleadingly representative

- **in vivo measurement**
  - imaging-based
  - molecular sensors

- **in vivo sensor logic**
  - detect condition in engineered cells
  - trigger action
    - switch metabolic pathway
Measured Data are Samples from Distributions

- The thing we’re measuring has some distribution
  - we sample it
  - from our samples we infer the underlying distribution
  - our sampling is fraught with extra dispersion and biases
  - a lot of what we try to do is to separate out the underlying stuff that gives rise to the distribution of our measured data
Basics of Measurement Uncertainty Calculation

PDF's of the influence quantities $X_1$, $X_2$, $X_3$

equation of the measurand

measurand $Y$

$g_{X_1}(\xi_1) \rightarrow Y = f(X) \rightarrow g_Y(\eta)$
How does this work for Genomics?

- Quantitative results
  - e.g., Gene Expression

- Qualitative results
  - e.g., absence/presence, identity, mixtures
    - variant calling

- Hybrids
  - e.g., gene specific methylation

- Critical Elements
  - too many measurands to calibrate
  - no existing metrology for “nominal” properties
  - variation arising from bioinformatics

- Traceability ill-posed
  - traceable to A, C, G, T?

- MU needs extension for qualitative results
  - A ± ?
Approaches for Genomics

Quantitative Results

• Use external “Spike-in” controls to validate in situ
  – External RNA Controls

• Use mixed tissue samples to validate/monitor performance of measurement process
  – FDA-developed Gene Expression controls
  – Pilot project at UMMC

Qualitative Results

• Research to develop strategy to create and apply in situ controls
  – ERCC Controls for Systematic Sequencing Errors

• Initially, establish strategies to use shared, well-characterized samples to demonstrate measurement performance
  – Controls for Whole-genome Sequencing
Approaches for Genomics

Quantitative Results

• Use external “Spike-in” controls to validate *in situ*
  – *External RNA Controls*

• Use mixed tissue samples to validate/monitor performance of measurement process
  – *FDA-developed Gene Expression controls*
  – *Pilot project at UMMC*

Qualitative Results

• Research to develop strategy to create and apply *in situ* controls
  – *ERCC Controls for Systematic Sequencing Errors*

• Initially, establish strategies to use shared, well-characterized samples to demonstrate measurement performance
  – *Controls for Whole-genome Sequencing*
Method Validation

• Demonstration by provision of objective evidence...
  – that what I’m measuring is what I intend to be measuring
  – to prove I’m not just reporting artifacts
Measurement Uncertainty

• estimated value that gives me reasonable expectation of dispersion around my result
  – given my measurement system
  – use data from validation experiment to develop estimates
Standards for Genome-scale Measurements

Sarah A. Munro, Ph.D.
Marc Salit, Ph.D.
Biosystems and Biomaterials Division
What does NIST do in Biology?
Example: Genome-scale Reference Materials

• Whole Human Genome Reference Materials
  – Cell line DNA
  – Genome in a Bottle Consortium

• Standard Reference Material (SRM) 2374
  – DNA Sequence Library for External RNA Controls
  – External RNA Control Consortium
ERCC Collaborative Study to Create SRM 2374

- ERCC members contributed candidate sequences
  - Collaborative study to test RNA controls on variety of platforms
  - Selected well-performing sequences for SRM
- SRM contains 96 unique control sequences in plasmid DNA
  - *In vitro* transcription to make RNA controls
  - Intended to mimic mammalian mRNA
Library of 96 Controls in SRM 2374

Sequence Lengths

GC Content
Creating Spike-in Mixtures from SRM 2374

SRM 2374 Plasmid DNA Library

in vitro transcription

RNA transcripts

Pooling

Mixtures with known abundance ratios
Microarray Analysis

- Total RNA Isolation
  - mRNA Amplification
  - Target Preparation
    - Hybridization of Array Content
      - Hybridized Target Detection
- Hybridized Target Detection
- Expression Measures
  - Statistical Analysis

RNA Sequencing

- mRNA Enrichment from Total RNA
  - Library Preparation
    - Sequencing by Synthesis or Ligation
      - Sequence Alignment to Reference
Method Validation with the *erccdashboard*

- Open-source R package - *erccdashboard*
- Developed as part of FDA-led SEQC interlaboratory study
- Assess technical performance of a gene expression experiment
- Compare results
  - Within a single laboratory
  - Between laboratories
Bias in ratios mRNA fraction? mRNA enrichment?

Signal Ratio vs. Spike-in Abundance
215 Dynamic Range

Observed

LOD of Spiked Transcripts

Signal vs. Spike-in Abundance

Ratio
- 4:1
- 1:1
- 1:1.5
- 1:2

Sample
- CTL
- MET

Read Depth Normalized Log2 Transformed ERCC Counts

Log2 ERCC Spike Amount (attomol nt/μg total RNA)
ROC Curves, differentiation between true positives and true negatives

Area Under the Curve (AUC) for comparison

<table>
<thead>
<tr>
<th>Ratio</th>
<th>AUC</th>
<th>Measured</th>
<th>Spiked</th>
</tr>
</thead>
<tbody>
<tr>
<td>4:1</td>
<td>1.00</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>1:1.5</td>
<td>0.95</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>1:2</td>
<td>0.97</td>
<td>16</td>
<td>23</td>
</tr>
</tbody>
</table>
Threshold P-value for chosen FDR

For 4:1 ratio ERCC controls LODR = 26 counts in this experiment

<table>
<thead>
<tr>
<th>Ratio</th>
<th>LODR Estimate</th>
<th>90% CI Lower Bound</th>
<th>90% CI Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>4:1</td>
<td>26</td>
<td>19</td>
<td>32</td>
</tr>
<tr>
<td>1:1.5</td>
<td>Inf</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>1:2</td>
<td>250</td>
<td>120</td>
<td>350</td>
</tr>
</tbody>
</table>
 LODR for 4:1 ratio
ERCC controls
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>95% CI Lower Bound</th>
<th>95% CI Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r_m )</td>
<td>0.9539</td>
<td>0.9478</td>
<td>0.96</td>
</tr>
</tbody>
</table>

This ERCC is below the 4:1 LODR.
This ERCC is below the 4:1 LODR.

ERCC results give an indication of how confident we should be in the measurement process that we used for our sample transcript measurements.
Acknowledgements

• NIST
  – Jenny McDaniel
  – David Duewer
  – Matthias Roesslein
  – Scott Pine
  – Steve Lund
  – Justin Zook
  – Margaret Kline
• Members of the ERCC
• FDA SEQC Study Participants

Postdoc Opportunities at NIST in Gaithersburg, MD and Stanford University, CA

• New joint research program “Advances in Biomedical Measurement Science” (ABMS) at Stanford

For information contact:
  – smunro@nist.gov
  – salit@nist.gov