Innovating Genomic Research via Error-Corrected Sequencing Methodologies

11 June 2019
Alexandria, Virginia

Co-chair: Bob Young, MilliporeSigma
Co-chair: Francesco Marchetti, Health Canada
Co-chair: Mark Fielden, Amgen
Agenda

- HESI and introduction of transformative new technologies
- *In vivo* mutation analysis - role in safety assessment
- Error-Corrected Next Generation Sequencing
- Case Study: Duplex Sequencing™ *in vivo* mutation data
- Next steps for Error-Corrected Sequencing Workgroup
- Applications to drug and chemical safety assessment
Goal: Validate Error-Corrected Next Generation Sequencing as an alternative methodology for evaluating *in vivo* mutagenesis

Rationale: EC-NGS can...
- Detect ultra rare, new mutants following exposure to chemicals
- Be applied to any gene, any tissue and any species, including humans
- Be integrated into existing toxicology studies to reduce animal use
- Measure mutants and mutation spectra for mechanistic information
- Be used to detect chemical-induced mutagenesis for hazard identification
- Detect oncogene mutations and clonal expansion to potentially improve the biological relevance for assessing cancer risk
Health and Environmental Sciences Institute (HESI)

- **Mission:**
  - Developing science for a safer, more sustainable world

- **Examples of success**
  - Leadership in introduction of new technologies
  - “Alternatives to Carcinogenicity Testing project”
Genetic Toxicology Technical Committee (GTTC)

- **Mission:**
  - To advance the field of genetic toxicology and human risk assessment through the *international collaboration of experts*.

- **Objectives:**
  - Integrate genetic toxicology into risk assessment and decision-making for protection of human health.
  - Improve new and existing test guidelines, strategies, and interpretation of results.
  - *Examine non-traditional modalities, including novel entities and technologies.*
In Vivo Mutation Analysis – Historically difficult to measure

1850’s | 1920’s-1950’s | 1940’s-1950’s | 1940’s-1980’s | 1990’s-Present
Why Worry About Mutations?

DNA mutations are associated with many human diseases

Somatic Cells

- Cancer
- Other Diseases

Germ Cells

- Heritable Damage (genetic damage to offspring)
- Infertility

Spermatocytes, oocytes
The Value of *In Vivo* Mutation Detection: Predictive Biomarker of Cancer Risk

Window of opportunity to detect biomarkers of cancer risk

Genetic heterogeneity (passenger and driver mutations)

http://www.ndhealthfacts.org/wiki/Oncology_%28Cancer%29
**Next Generation Sequencing (NGS)**

NGS is powerful for the detection of clonal mutations (present in most molecules in a sample) but remains poor at detecting low frequency subclonal genetic differences.

**Easy:**
- Genetic difference between two individuals or tumor vs. normal

**Hard:**
- Detecting one mutant cancer cell among 1,000 normal cells

**Impossible:**
- Identifying ultra-rare mutations induced by carcinogen exposures
Error-corrected (EC)-NGS Technology Comparison

• Measuring mutations *in vivo* is difficult

• Transgenic Rodent (TGR) Mutation assays measure Mutant Frequency (MF) in *one transgene in any tissue*

• Pig-a Mutation Assay measures MF in *one gene in one tissue*

• Next Generation Sequencing was revolutionary but has error rate of 1 in about 1000 bases sequenced

• Error-Corrected NGS drops error rate to 1 in $10^8$

• *This is below background MF of mammalian genomic DNA*
Clinical Application: Detection of Chemotherapy Resistance Mutations

- Chronic Myelogenous Leukemia genotyping is critical for effective therapy
- Sequence Exons (4-7) in Abl gene; E279K mutation known to confer Imatinib resistance
- Duplex Sequencing identified ultra rare E279K mutant tumor cells with resistance mutation
- Clinician can change chemotherapy before relapse
Duplex Sequencing™ - Example of Error-Corrected NGS

- DNA fragmented and tagged with unique double stranded DNA barcodes
- Tagged fragments amplified by PCR and sequenced by NGS
- Errors occur at fairly high levels during PCR and sequencing
- Fragments aligned in silico using fragment and strand-specific bar codes
- Mutant calls are made based on persistent variants in most reads of both strands
Brief Review of EC NGS Pilot Data

- **Bridging study: Compare Big Blue® TGR to Duplex Sequencing™**
  - Mutant analysis comparing TGR “plaque assay” to Duplex Sequencing
  - Big Blue® mouse treatments: Vehicle, Ethyl nitrosourea and Benzo(a)pyrene
  - Tissues: Liver and Bone Marrow
  - Target genes analyzed: cII transgene and 4 endogenous genes

- **Extension into non-TGR animals**
  - TgrasH2 mice; carries human HRAS oncogene
  - Vehicle and Urethane; Lung, spleen and blood
  - 7 endogenous genes including 3 murine ras genes and human HRAS transgene

- **Data from collaboration among:**
  - MilliporeSigma: Bob Young and Rohan Kulkarni
  - Amgen: Mark Fielden and Sheroy Minocherhomji
  - TwinStrand BioSciences: Jesse Salk and Clint Valentine
Similar Mutant Frequencies with TGR Plaque and Duplex Methods

**cII Mutant Frequency Using**

Duplex Sequencing™

mutants per basepair

**cII Mutant Frequency Using**

Big Blue® Plaque Assay

mutants per scorabble plaques
Plaque and Duplex Methods Detect Comparable Mutant Classes

Duplex Sequencing of cII in genomic DNA

Duplex Sequencing of cII mutant plaques

C-terminus

N-terminus
Mutagen-specific Trinucleotide Mutational Spectra

Benzo(a)pyrene

N-ethyl-N-nitrosourea

Urethane
Mouse BaP Trinucleotide Mutant Spectra aligns to human smoker lung tumor spectra

Duplex Sequencing™ Signature: Benzo(a)Pyrene in Mice after 28 days

COSMIC Database Signature 4: Human tumors from cigarette smokers

http://cancer.sanger.ac.uk/cosmic/signatures
DS can be extended into endogenous genes and other animals.
Moving beyond “Hazard Identification”

- Most common human lung tumor mutation is exon 3 codon 61 in human HRAS gene (COSMIC)
- Urethane treated TgrasH2 mouse positive controls; 100% get lung adenocarcinomas by 11 weeks
- After 29 days, increased Mutant Frequency seen across introns and exons of all genes BUT
- We also see up to 300-fold increase in T>A transversions at codon 61 in exon 3
- This may represent clonal expansion events at the earliest stages of tumorigenesis - biologically different than random mutagenesis across the genome
HESI GTTC Error-Corrected Sequencing Workgroup History

Chronology:

- **May 2017:** Early Duplex Sequencing™ data discussed with GTTC leadership
- **May 2018:** DS data presented to GTTC at Annual Meeting
- **Dec 2018:** Presentation to HESI and GTTC Leadership and voted to accept
- **Feb 2019:** Webinar to interested members, distributed survey
- **Mar 2019:** Liaison to HESI eSTAR (Genomics) Committee
- **May 2019:** Organizing session held during May GTTC Annual Meeting
Error-Corrected Sequencing Workgroup Organizing Session May 2019

- Open to all error-corrected NGS technologies - initial focus on DS
- Majority selected *in vivo* applications in initial questionnaire
- Goal is to transfer technology to second/third lab
- Biomarker qualification -- Engage FDA, letter of intent for context of use
- Phase 1: Optimize methods by running limited chemicals in a few labs
- Standardize methods and analytical tools
- Phase 2: Organize multi-lab, multi-chemical collaborative trial
Nonclinical Biomarker of Cancer Risk - i.e. FDA/PhRMA NegCarc Program

- Use DS in 6 month rat toxicity studies as an enhanced biomarker of cancer risk
- Goal is to replace 2 year rat cancer bioassays in many situations

Predictive Toxicology Applications - Add to early repeat dose tox studies

- Same endpoint to kill drugs early, work on alternatives or set safety margins

Follow-up to positive in vitro mutation or positive carcinogenicity data

- Replacement for TGR assays

In Vitro mutagenic biomarker and use for generating mutagenic spectra

Evaluation of cells for biopharmaceutical cell banking or human gene editing
Thank you.

For follow up questions:

Bob Young

MilliporeSigma

bob.young@milliporesigma.com

+1 301-610-2143