NIST "Genome Editing Standards" convened on May 2, 2016
Workshop Summary

I. Purpose of the Workshop

On May 2, 2016 NIST hosted a one day workshop in partnership with the American Society of Gene and Cell Therapy (ASGCT) to bring together key stakeholders in government, academia, industry and informatics experts to begin identifying measurement assurance and standards barriers, technology gaps, and potential solutions for the genome editing field. A focus was placed on the measurement and standards needs for those intending to use targeted genome editing tools to generate cell and gene therapies. This workshop was the first of its kind for the genome editing field. It was anticipated that this workshop would be the first in a series of meetings aimed at identifying and charting a path forward for establishing increased confidence in genome editing applications. The workshop was sponsored by Microsoft Corporation and Editas Medicine and held at Microsoft's Chevy Chase, MD location.

II. Agenda

8:30 - 9:00	**Check-in & Continental Breakfast**
9:00 - 9:15	Welcome / Introduction to the day *Samantha Maragh (NIST)*
9:15 - 9:35	Beyond Reproducibility in Biomedical Measurements *Anne Plant (NIST)*
9:35 - 9:45	Genome in a Bottle: so you've sequenced a genome, how well did you do? *Justin Zook (NIST)*
9:45 - 10:30	**Keynote address:** Defining and Improving the Specificities of Genome-Editing Nucleases *J. Keith Joung (Massachusetts General Hospital / Harvard Medical School)*
10:30 - 10:45	**Break**
10:45 - 11:15	Regulatory Perspective *Lilia Bi (FDA, CBER)*
11:15 - 11:30	Perspective from industry on measurement needs *Edward Rebar (Sangamo Biosciences)*
11:30 - 12:15	**Panel Discussion:** Measurement Assurance Needs for Genome Editing *Panelists: Lilia Bi (FDA), Vic Myer (Editas), Garrett Rettig (IDT), Martin Aryee (Harvard), J. Keith Joung (Mass. General Hospital/Harvard), Reynald Lescarbeau (Intellia)*
12:15 - 12:30	Charge for Breakout Sessions: Identifying Barriers, Gaps & Solutions *Samantha Maragh (NIST)*
12:30 - 12:45	**Lunch pick-up (provided)**
12:45 - 1:45	**Breakout Sessions Round 1 - during working lunch** 1) Needs for confidence in off-target measurements

	2) Needs for confidence in genome editing pipeline assessment and comparability
	3) Data/Informatics Standards Considerations
1:45 - 2:00	Breakout session switch
2:00 - 3:00	**Breakout Sessions Round 2**
	1) Needs for confidence in off-target measurements
	2) Needs for confidence in genome editing pipeline assessment and comparability
	3) Data/Informatics Standards Considerations
3:00 - 3:30	**Break**
	Moderator Samantha Maragh (NIST)
3:30 - 4:15	Reports from breakout sessions (~15 mins each)
4:15 – 5:00	Open Discussion: Addressing identified measurement and standards barriers, gaps and solutions
5:00 - 5:15	Follow up actions / future meetings discussed
5:15 - 6:30	**Networking Reception**

III. Introduction and Overview

Targeted Genome Editing is a technology space where there is a great need for reliable measurement methods for assuring the results of editing. Modalities for targeted genome editing include but are not limited to Zinc Finger Proteins (ZFPs), Homing Endonucleases, Transcription Activator-Like Nucleases (TALENs) and Clustered, Engineered Meganucleases, Regularly Interspaced Palindromic Repeats (CRISPR). These technologies are being actively pursued by industry, academic, government and non-profit sectors to advance medicine and bioscience in areas such as: regenerative medicine, synthetic biology, novel antimicrobials and antivirals, protein therapeutic biomanufacturing, agriculture and global food production. Utilizing these technologies for production and medicine will first require robust quantitative assays and measurements to enable high confidence in characterization of DNA alterations resulting from genome editing.

Confidence and comparability of measurements underpinning the genome editing field is a community identified need. Questions have arisen surrounding: How to reliably measure which changes were made? How to overcome the limited ability to quantify changes? Is data from studies within the same lab or across different labs comparable? What tools are needed for users to assess their confidence in various parts of the process? These publicly identified needs led to NIST investing effort in the genome editing area to understand the needs and determine how assisting with these needs was consistent with the NIST mission. To that end NIST convened this public workshop and invited subject matter experts to provide their perspectives on the measurement assurance needs of the genome editing filed.

IV. Summary of Talks

Beyond Reproducibility in Biomedical Measurements (*Anne Plant -NIST)*

Genome in a Bottle: so you've sequenced a genome, how well did you do? (Justin Zook – NIST)

Sequencing technologies and bioinformatics pipelines disagree. Who is right? Is anyone right?

Genome in a Bottle Consortium (GIAB) – Whole Genome Variant Calling (human genome)

https://www.nist.gov/programs-projects/genome-bottle http://jimb.stanford.edu/giab/

- gDNA reference materials to evaluate performance
- materials certified for their variants against a reference sequence, with confidence estimates
- established consortium to develop reference materials, data, methods, performance metrics
- Characterized Pilot Genome NA12878
- Ashkenazim Trio, Asian Trio from PGP in process
- Obtain metrics for validation, QC, QA, PT
- Determine sources and types of bias/error
- Learn to resolve difficult structural variants
- Improve reference genome assembly
- Optimization
- Enable regulated applications

Global Alliance for Genomics and Health Benchmarking Task Team (GA4GH)

- Developed standardized definitions for performance metrics like TP, FP, and FN.
- Developing sophisticated benchmarking tools
- Approaches to benchmarking Variant Calling
  - Well-characterized whole genome Reference Materials
  - Many samples characterized in clinically relevant regions
  - Synthetic DNA spike-ins
  - Cell lines with engineered mutations
  - Simulated reads
  - Modified real reads
  - Modified reference genomes
  - Confirming results found in real samples over time
- Challenges in benchmarking Variant Calling
  - It is difficult to do robust benchmarking of tests designed to detect many analytes (e.g., many variants)
  - Easiest to benchmark only within high-confidence bed file, but…
  - Benchmark calls/regions tend to be biased towards easier variants and regions
  - Some clinical tests are enriched for difficult sites
  - Always manually inspect a subset of FPs/FNs
  - Stratification by variant type and region is important
  - Always calculate confidence intervals on performance metrics

Defining and Improving the Specificities of Genome-Editing Nucleases (*J. Keith Joung- Massachusetts General Hospital / Harvard Medical School)*

- For patient safety it is critical to assess off-target genome editing
- Off-target editing isn't random, but guided by the editing nuclease design
- Defining and quantifying off-targets should be part of overall safety evaluation
- Ideal state for defining off-targets would be: genome-wide, unbiased, highly sensitive
- Summary of cell based off-target editing detection methods (IDLV Capture, HTGTS, BLESS)
- Details on studies of GUIDE-Seq method developed in Joung lab.

- *In vitro* (purified molecules in a tube) methods for off-target detection has advantages over cell based methods and are being explored and developed. Can have lower detection limit and be more scalable approach.
- FIRST description of new in vitro off-target detection method developed in Joung lab "CIRCLE-Seq" – purified DNA is shared, circularized then treated with genome editing tool, circles with cleavable target are linearized and adapters are added to linear ends for sequencing. Stated as a lower background technique to other in vitro off-target detection methods and more sequencing efficient able to obtain sufficient data from a MiSeq run.
- Comparison of GUIDE-Seq and CIRCLE-Seq data sets showed many more off-targets detected by CIRCLE-Seq- potentially higher false positive, but checking GUIDE-Seq for a few sites that were CIRCLE-Seq positive and GUIDE-Seq negative showed true edits were detectible in GUIDE-Seq DNA. Some CIRCLE-Seq identified sites could not be confirmed by targeted re-sequencing
- It is difficult but important to determine how to appropriately evaluate and compare off-target detection assays

Regulatory Perspective (*Lilia Bi -FDA, CBER)*

- Gene therapy (GT) products are the products that mediate their effects by transcription or translation of transferred genetic material, or by specifically altering host genetic sequences

GT products:
  - Plasmid
  - Viral vector (various virus type)
  - Bacterial vector (Listeria, Salmonella, etc)
  - mRNA
  - Ex vivo genetically modified cells
  - Site-specific nucleases
- Genome Editing (GE) Technologies for Gene Therapy (GT) & Regulation
  - Genome editing technology used to specifically modify the human genome is considered gene therapy and subject to FDA regulation, whether used directly or used to modify cells ex vivo.
  - Emerging class of gene therapy products that mediate their effect by specifically modifying human genome sequences.
  - Human somatic cells: CBER/OCTGT*
  - Plants: regulated by CFSAN, USDA
  - Animals: regulated by CVM, USDA
  - FDA co-sponsored National Academy of Medicine study on gene editing http://www.nationalacademies.org/hmd/Global/News%20Announcements/NAS-NAM-Human-Gene-Editing.aspx
- Potential Safety Concerns for GE Technologies used in Gene Therapies
  - Cleaving the desired target in the genome - specificity issue
  - Adverse effects of genomic DNA cleavage at on- and off-target sites
  - Adverse effects of gene mutation introduced by endogenous DNA repair activities
  - Overall chromosome instability
  - Inadequate assembly of a donor gene in the genome
  - Adverse impact of the vector delivery system (e.g., insertional mutagenesis potential)
- Current Recommendations to Assess Safety of GE Products
  - Kinetics of nuclease cleavage and persistence of cleavage activity
  - Percentage of cleavage at the on- and off-target sites

  - Identification and characterization of off-target cells/tissues
  - Evaluation of the profile of indels and types of mutations generated

Public Access to CBER

- CBER website: http://www.fda.gov/BiologicsBloodVaccines/default.htm Phone: 1-800-835-4709 or 240-402-8010
- Consumer Affairs Branch (CAB) Email: ocod@fda.hhs.gov Phone: 240-402-7800
- Manufacturers Assistance and Technical Training Branch (MATTB) Email: industry.biologics@fda.gov Phone: 240-402-8020
- Follow us on Twitter https://www.twitter.com/fdacber

OCTGT Regulatory Resources

- *OCTGT Learn* Webinar Series http://www.fda.gov/BiologicsBloodVaccines/NewsEvents/ucm232821.htm
- Regulatory Questions: CBEROCTGTRMS@fda.hhs.gov Lori.Tull@fda.hhs.gov 240-402-8361

Perspective from industry on measurement needs (*Edward Rebar- Sangamo Biosciences)*

- *Quantifying Genome Editing for Therapeutic Applications*
- Therapeutic Genome Editing: Key Quantitation Needs

Activity	Key Needs
Lead activity screens	throughput / simplicity / precision (accurate lead ranking) *
Specificity studies	assay for identifying candidate off-target sites high sensitivity (ideally $1:10^9$)
Preclinical & clinical studies	Accuracy *

- * Can be helped by standard protocols
- Checking for vector integration is important and a useful approach
- Sequencing is a valuable tool but can have high background due to: sequencing errors, PCR artifacts, differences in reference genome and “your” cells, sequencing data processing
- For therapeutic genome editing – measurement needs vary with program stage
- Deep sequencing provides a robust means for meeting most measurement needs
- A key issue in gauging specificity is finding candidate off-target sites
- A second issue in gauging specificity is achieving sufficient sensitivity
- Cellular assays provide a complementary approach for gauging editing specificity and safety

Panel Discussion

- Q1: What do you think are the steps in the genome editing pipeline that are most important to have confidence in characterizing/measuring, and which are most difficult to characterize/measure with confidence?
- Q2: What role does informatics, data handling, data storage play in confident assessment of genome editing and comparability of studies or data sets?
- Q3: Where might measurement assurance infrastructure or standards (i.e. physical test material/calibrant, documentary standard, accepted community reporting norm, standard or shared dataset, blinded comparison) make the most impact for confident assessment of genome editing?
- Q4: Where do you think the genome editing community might be able to work together “pre-competitively” to improve genome editing measurement assurance

V. Summary of Breakout Groups

Needs for confidence in off-target measurements

Topic 1- Round 1:

1. What aspects of off-target genome editing do you need to be able to make confident statements about that you are not currently able to?

- Location & frequency of edits
- Rapid patient by patient assays
- Define low probability event
- Incorporating patient specific data
- Cell type editing
- Physical type of off-target
- Documentation of conditions
- Confounding transfection effects
- Transfection kinetics & reagents
- Transfection normalization
- Sensitivity of detection increase
- Orthogonal techniques
- Characterizing the functional importance of off-targets
- Relevance of off-targets to causing disease
- Defining how comprehensive analysis needs to be
- Knowing the background mutation rate
- Distribution of mutations in population

2. What is needed to make or progress towards making these statements with confidence?

- Transfection standard
- Reference cell line for comparison of off-targets
- Publicly available testing agency or institute
- Benchmark method for determining editing is happening in a cell
- Comprehensive list of off-targets for benchmarking/comparison
- Transfection standard
- Reference cell line for comparison of off-targets
- Publically available testing agency or institute
- Benchmark method for determining editing is happening in a cell
- Comprehensive list of off-targets for benchmarking/comparison
-

Topic 1- Round2:

- Cleavage rates vs indel efficiency
- Known sensitivity of detecting off targets
- Guide-seq vs Circle seq sensitivity
- Bias from assays
- Same lab, same assays, same time
- Which off targets matter?
- Methods to follow up and determine what matters
- Differences in sequencing
- Integrating virus vs edit comparison of data

- Parental vs. mutant control
- Proficiency scheme of gene editing material
- Blinded with known edits and known data
- Unit definition of Cas9 activity

Needs for confidence in genome editing pipeline assessment and comparability

Topic 2- Round 1:

1. What aspects of the genome editing pipeline or comparing genome-editing studies (excluding off-target editing) do you need to be able to make confident statements about, that you are not currently able to?
    - At the cellular level, what is the concentration of the guide RNA, and how that affects editing.
    - General cell and tissue practices. Crossover with cell therapy manufacturing.
    - Bulletin of best practices from NIST would be a valuable resource. Organizing the "pulling together" would help (but may be slow)
    - What does "80% efficiency" mean? How do we measure it reproducibly? Is that good?
- How do you confirm your measurement? Is a pooled measure representative of the true sample?
- Other species, other non-traditional cell lines? (esp. therapeutic)
- Is there a recommended set of optimization experiments to show you have an optimized pipeline?
- Maybe not a "single" set, but to understand what is process-driven v. asset-driven.
- How do I prove my guide RNA is active?

2. What is needed to develop or progress towards developing these tools and competencies and having confidence in assessing their performance?
    - Data standard to share/record data? Do we need standards until there's a little more maturity in the field?
    - Quantifying specific activity of things like Cas9 to compare vendors
    - Identify the materials people are using and characterizing those.
    - In addition to gRNA, cellular concentration/delivery of the donor, of other tools (Cas9 plasmid, for ex.)
    - How do you count efficiency in an aneuploidy genome or a genome of unknown ploidy?

Topic 2- Round 1:

- Side effect of editing on target
- Immunogenicity? (both nuclease and edited protein)
- Sharing info/data, esp. kinetics, ELISA equivalent for Cas9
- Animal models of therapy? How do you replicate human toxicity in them?
- Software/ref stds; on-off target design
- Broad Institute Genome tiling, seeking enhancers, positions activity
- QC and std
- Deep sequencing for off-sites

- Standard reagents/tools that we agree to compare results
- Methods for measuring amount of cas9/ gRNA; validated
- What are the common methods, reagents?
- Simple methods for on-target assessment
- Methods to compare kinetics of what's happening in the cell, competitive events
- DNA repair capabilities?
- Comparing across cells/tissue

- Can we blockoff different sites? Would that allow us to better characterize effects

- Titrating reagents, small molecules to modulate activity
- In vitro vs. in vivo assay. What data can you use?

Data/Informatics standards considerations

Topic 3

1. <u>What informatics, algorithms, data standards or data infrastructure is needed to support genome editing?</u>
   - Minimal information standard
   - Includes metadata on editing expt
   - Includes data on repair outcomes and off-target effects (VCF file, FASTQ files)
   - The "riskome" – a subset of human genome targets to evaluate
   - Arbitrariness of pipelines
     - Too many parameters, subjective
   - Evaluating prediction methods for off-target effects
     - preceded by the problem of characterizing off-target effects
   - Material stds and experimental issues
     - Need for "edited genome in a bottle" DNA standard to use in every experiment Or could be an edited cell preparation

2. <u>What is needed to develop or progress towards developing these tools and competencies and having confidence in assessing their performance?</u>
   - Minimal information standard
     - Intellia version could provide a starting point
     - Need to identify common interest
     - Virtual task force
   - The "riskome" (didn't discuss how to get there)
   - Arbitrariness of pipelines
     - Benefits from a test data set of raw seq reads
   - Evaluating prediction methods for off-target effects

   Material stds and experimental issues
     - (didn't discuss how to get there

VI. Outcomes/Proposed Work Plan

- Participants said the workshop was beneficial and future workshops/meetings were requested to chart a path forward
- Industry competitors agreed measurement assurance for the genome editing pipeline particularly on and off-target editing measurements are PRIORITY needs and are pre-competitive issues and areas for collaboration
- A wide range of need areas were identified by the workshop participants (see breakout group summary above)
- Confidence in detecting and quantifying off-target events and bioinformatics support were identified as high priority areas
- Supporting measurement assurance for this field is aligned with the NIST mission and of value to FDA.
- Leaders in this spaced including: the American Society of Gene and Cell Therapy (ASGCT) and industry such as: Microsoft, Editas, Illumina, IDT and many others are interested in partnering to meet the needs.
- Follow-on in person workshops or discussions held bi-annually coinciding with large meetings including the Keystone Precision Genome Engineering Meeting as well as the ASGCT annual meeting were suggested.