Next-generation Sequencing (NGS) Technology, data formats standardization and promotion of interoperability protocols

September 24 & 25, 2014 8:30 am- 5:00 pm
Porter Conference Center, Building 35, NIH Campus

The goal of the conference is to engage NGS stakeholders in discussions to identify the benefits and limitations of NGS technology as it pertains to FDA-regulated products. FDA would like to involve the broader NGS community, including relevant representatives of other government agencies, academia, industry developers of sequencing and bioinformatics platforms, and end-users of the technology, to discuss these and other relevant questions in an effort to move forward in a more informed manner to better provide guidance to sponsors on these and other issues.

The ultimate goal of this conference to begin the process of developing guidance for a versatile platform of next-generation sequencing technology verification, standardization of data formats and bioinformatics analysis provenance, and promotion of interoperability protocols.

Wednesday September 24, 2014	Thursday September 25,2014
• **8:30– 9:30 am** Introductory remarks • **9:30 – 11:30 am** Next-generation Sequencing Standards *coffee break and poster viewing* lunch on your own • **1:00 – 3:00 pm** Big Data Administration and Computational Infrastructure *coffee break and poster viewing* • **3:30 – 5:00 pm** Database Development	• **8:30 – 10:30 am** Biologics Product Evaluation • **10:30 – 12:00 pm** Clinical Biomarkers and Personalized Medicine lunch on your own • **1:00 – 3:00 pm** Next-generation Sequencing Devices *coffee break* • **3:15 – 4:45 pm** Food Safety and Pathogen Detection • **4:45-5:00 pm** Closing Remarks

SESSION DESCRIPTIONS

Introductory Remarks

Dr.Taha Kass-Hout, Chief Health Informatics Office, FDA

Dr.Vahan Simonyan, Visiting Associate, FDA,CBER

Next-Generation Sequencing Standards (Chaired by Dr. Vahan Simonyan/FDA)

Perspectives of FDA and select government representatives will be highlighted in this section as they pertain to the submission of NGS data to the U.S. FDA and other regulatory organizations and validation of analytical protocols used to acquire this data.

The immediate motivation for the proposed standards is to promote confidence in all data delivered by various technologies to users and regulatory staff, and to ensure the suitability of such data for high-stakes research, including but not limited to studies supporting regulatory submissions to the FDA. To better support regulatory decision-making, there is a need to implement a set of accepted standards that will not only guarantee data of the highest quality, but further ensure the originality of such data and all of the bioinformatics approaches used to generate NGS information. The short term goals and the focus of this conference are centered on the quality of the information and uniform meta-information submission. Future workshops may be needed to address topics such as harmonization of various computational protocols and standardization and validation of the bioinformatics pipelines.

- Dr. Tong Weida, Director, FDA ,Division of Bioinformatics and Biostatistics
- Dr. Jean Thierry-Mieg, NIH, National Center for Biotechnology Information
- Dr. Amnon Shabo, Health Informatics Specialist, European Federation of Medical Informatics
- Dr. Eugene Yaschenko, Chief, NIH, National Center for Biotechnology Information

Big Data Administration and Computational Infrastructure (Chaired by Dr. Eugene Yaschenko/NCBI)

Platforms for NGS data storage and analysis will be the key focus of this session. Stakeholders will represent their perspectives on storage hardware, large data transfer networking, and distributed compute platforms capable of supporting NGS based research and regulatory sciences.

Different stakeholders will have the opportunity to present their perspective of what they envision as the appropriate resources to best enable NGS stakeholders to successfully complete their tasks. The goal is to determine which resources are available, necessary and/or most appropriate.

In addition to introducing diverse platforms, current and developing computing infrastructures will be introduced. This topic will be discussed by some developers whose main focus is NGS analysis. This will help to identify the tools that are widely used within the community and to ensure that these are the optimal tools for producing accurate and reliable results.

- Dr. Warren Kibbe, Director, NCI, Center for Biomedical Informatics and Information Technology
- Dr. Carolyn Wilson, Microbiologist, FDA, CBER and Dr. Vahan Simonyan, Visiting Associate, FDA,CBER
- Dr. Toby Bloom, Deputy Scientific Director, Informatics at New York Genome Center

Database Development (Chaired by Dr. Raja Mazumder /George Washington University)

This session will focus on the need for the development of curated databases that address all of the requirements of NGS data stakeholders. Discussion will focus on validation and integration protocols necessary for maintaining the integrity of the data and databases, keeping in mind the current and future goals of current model organism databases, such as NCBI RefSeq, dbSNP, UniProt and Representative Genomes and clade specific sequence sets. The hope is to establish the key features that must be incorporated when developing NGS resources and the steps needed to produce viable and reliable resources that facilitate collaboration and research.

- Dr. Kim Pruitt, Staff Scientist , National Library of Medicine,
- Dr. Mike Cherry, Professor, Stanford University
- Dr. Rodney Brister, Group Leader, Virus Genome Group, National Center Biotechnology Information

Biologics Product Evaluation (Chaired by Dr. Arifa Khan/FDA)

This section provides an opportunity for industry manufacturers of biologics to communicate their perspectives to FDA.

This session will review different applications of NGS in development, evaluation, and quality control of biological products. The applications include adventitious virus detection in cellular substrates and products manufactured using *ex vivo* ingredients, characterization and identity testing of complex biologicals, and assessment of genetic stability and molecular consistency of live and inactivated vaccines. The talks will describe current efforts for addressing gaps in sample preparations, standards, data analysis pipelines, and viral databases. Additionally, approaches for data collection, format, and transfer and standardization of protocols will be discussed. Speakers will be encouraged to share their vision of and expectations for the NGS-based methods and discuss challenges in making the results comparable and easily transferrable to alternative analysis platforms.

- Dr. Arifa Khan, Supervisory Microbiologist, Food and Drug Administration, CBER
- Dr. Charles Chiu, Associate Professor, UCSF School of Medicine
- Dr. Robert Charlebois, Senior Scientist, Sanofi Pasteur Limited
- Dr. John Thompson, Senior Investigator , Merck

Biomarkers and Personalized Medicine (Chaired by Dr. Eric Donaldson/FDA, Dr. Raja Mazumder/GWU)

Attention will be directed to clinical research using next-generation sequencing technology as a method to validate the efficacy of drugs, detection of biomarkers, personalized medicine and other bio-medicine centric disciplines.

In order to establish the trust of NGS technology in this field, we must first discuss the security and privacy measures involved. Of particular interest will be the needs of FDA and CDER in evaluating NGS data, as well as the needs of the hospitals that are actually implementing HTS pipelines and cancer research.

As new technology is developed and used, it is crucial to establish appropriate protocols for its use. This session should promote the establishment of certain benchmarks and test measurements that can be used to verify the validity of NGS data submissions while maintaining its integrity.

- Dr. Eric Donaldson, Virology Reviewer, FDA CDER Division of Antiviral Products
- Dr. Andrea Ferreira-Gonzalez, Professor of Pathology and Director Molecular Diagnostics Laboratory at the Virginia Commonwealth University Medical Center
- Dr. Andrew Grupe, Senior Director, Pharmacogenomics, Celera
- Dr. Charles Sawyers, Chair, Human Oncology and Pathogenesis Program; Marie-Josee and Henry R. Kravis Chair, Memorial Sloan Kettering Cancer Center
- Dr. Laura J. van 't Veer, Leader, Breast Oncology Program, and Associate Director, Applied Genomics, UCSF Helen Diller Family Comprehensive Cancer Center

Next-Generation Sequencing Devices and Clinical Applications (Chaired by Dr. Zivana Tezak / FDA, Dr. Justin Zook / NIST and Dr. Heike Sichtig)

The key area of this section is the development and validation of NGS hardware, protocols and applications. A description of the current regulatory approval pathway and requirements for NGS platforms will illuminate the challenges posed to the industry from the regulatory perspective. Conversely, an approved NGS platform representative will share his/her perspective and vision for the future. Applications of human microbiological research and clinical use will highlight the efforts of existing platforms in current research efforts, and a brief discussion of the status with input from the major industry platforms will provide an outlook for the future.

- Dr. Ira Lubin, Geneticist, Centers for Disease Control
- Dr. Justin Zook, Researcher, National Institute of Standards and Technology
- Dr. Leanne Kiviharju, Senior Director, Regulatory Affairs Illumina
- Dr. Heike Sichtig, Reviewer, Food Drug Administration, CDRH

Food Safety and Pathogen Detection (Chaired by Errol Strain/FDA, Dr. Heike Sichtig/FDA)

This session will concentrate on food safety and pathogen detection technologies based on next-generation sequencing. The subjects of environmental sample collection and analysis, study of outbreaks, food quality and pathogen control will be highlighted in this section as related to food and veterinary medicine.

The goal is to provide an overview of the use of NGS in public health and molecular epidemiological settings, focusing on outbreak cluster detection and source tracking of infectious bacteria/viruses and foodborne pathogens. Multiple studies have shown the ability of NGS to deliver high-resolution data that allows us to unambiguously link clinical isolates from affected individuals to contaminated food, medical instruments, environmental sources, etc. Standardization of pathogen NGS databases and their associated analysis pipelines is critical to widespread participation by public health organizations.

- Dr. Marc Allard, Research Microbiologist, Food and Drug Administration
- Dr. Bill Klimke, Staff Scientist, NCBI
- Dr. Kristen Holt, USDA/FSIS Liaison, Center for Disease Control and Prevention