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 conference goal is to engage next-generation sequencing (NGS) stakeholders in discussions to identify the benefits and limitations of NGS technology as it pertains to FDA-regulated products. FDA is involving the broader NGS community--including 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 this way, we will advance in a more informed manner to better provide guidance to sponsors on such issues. The conference will begin the process of developing guidance for a versatile platform of NGS 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

- Carolyn Wilson, PhD, Associate Director for Science, CBER, Food and Drug Administration
- Richard G. H. Cotton, PhD, Founding Patron, Scientific Director, Human Variome Project
- Vahan Simonyan, PhD, Lead Scientist, HIVE, CBER, Food and Drug Administration

FDA's activities supporting regulatory application of "Next Generation Sequencing" technologies

Carolyn A. Wilson, PhD, Food and Drug Administration

Recognizing the convergence of NGS technology and the technology's applications that will come under FDA regulatory purview, FDA initiated the FDA Genomics Working Group (GWG) in 2013. The FDA GWG was convened because efforts underway across the agency required better coordination and cross-agency communication. FDA GWG objectives include identifying and implementing the resources and expertise required to support regulatory evaluation of NGS data in a coordinated fashion. FDA GWG is also collaborating with governmental, academia, and industry partners to develop the IT infrastructure, data standards, and analytic approaches to support the use and application of NGS to solve scientific questions supporting FDA's regulatory mission as well as regulatory decision-making. The NGS Data Standards Workshop is one of many internal and external efforts that FDA GWG is spear-heading to achieve these goals.

Standards for medical genetic data storage and sharing

Richard G. H. Cotton, PhD, Human Variome Project

Our knowledge of the human genome, in particular when applied as DNA-based diagnostics, is based on sharing data on genes, variants and phenotypes and is vital for optimal genetic health care. Methods for this activity need to be standardized so data can be reported and shared worldwide. Human Variome Project members including gene variant database (LSDB) curators, lab heads and clinicians have, in a concerted effort, developed a series of recommendations for this process. All recommendations have been reviewed and published and several have been widely accepted including how to describe DNA variants and what content is essential for gene variant database (LSDB) content. Guidelines under development include database disclaimer statements, pathogenicity assignment, database content and quality assessment, and phenotype descriptions.

A National framework is being developed in Australia through a joint project between the Human Genetics society, the Royal College of Pathologists and HVP which appears to be the first such example worldwide.

Development and implementation of novel Next-Generation Sequencing standards

Vahan Simonyan, PhD, Food and Drug Administration

Research and regulatory scientists are developing and implementing novel NGS methodologies for the analysis of vaccines, detection of adventitious agents, identification of viral and bacterial strains, analysis of food contaminants, development of genetic markers, and other aspects of public health science. Researchers at the FDA and other agencies are making efforts to integrate NGS data into their routine regulatory analysis and review of mission-critical scientific experiments However, only select groups are currently capable of using such data due to the multifaceted deluge of NGS data.

To fully achieve the benefits of NGS through experimentation, data production, archival and bio-scientific analytics of NGS content in a single continuous pipeline, we propose development of technology guidelines together with all stakeholders of this revolutionary technology.

The subject of these guidelines is to:

- Create standardized data-typing engine allowing construction of data formats and providing framework for extensibility.
- Create bio-computational protocol validation engine allowing design of validation protocols.
- Define submission and archiving protocols.
- Develop high-performance infrastructure at the FDA to accept, store, archive and compute upon standardized NGS data and to provide framework for bioinformatics provenance.

Next-Generation Sequencing Standards

- Weida Tong, PhD, Director, Division of Bioinformatics and Biostatistics, Food and Drug Administration
- Amnon Shabo, PhD, Chair, Translational Health Informatics, European Federation of Medical Informatics
- Eugene Yaschenko, Chief, Molecular Software Section, NCBI, National Institute of Health

Chairperson's Opening Remarks
Vahan Simonyan, PhD, Food and Drug Administration

Presentation

FDA perspectives and those of 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 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 provenance 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 as well as to address topics such as harmonization of various computational protocols and standardization and validation of the bioinformatics pipelines.

Big Data Administration and Computational Infrastructure

- Vahan Simonyan, PhD, Lead Scientist, HIVE, CBER, Food and Drug Administration
- Warren Kibbe, PhD, Director, CBIIT, National Cancer Institute
- Toby Bloom, PhD, Deputy Scientific Director, Informatics, New York Genome Center

Chairperson's Opening Remarks
Eugene Yaschenko, National Institute of Health

Presentation

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.
Speakers will have the opportunity to present and discuss 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**.**

Database Development

- Kim Pruitt, PhD, RefSeq Project Lead, National Library of Medicine, NCBI, National Institute of Health
- Mike Cherry, PhD, Professor, Department of Genetics, Stanford University
- Rodney Brister, PhD, Staff Scientist, Virus Genome Group, NCBI, National Institute of Health

Chairperson's Opening Remarks
Arifa Khan, PhD, Food and Drug Administration
Philip Krause, MD, Food and Drug Administration

Presentation

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

to FDA. 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.

Biologics Product Evaluation

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

Chairperson's Opening Remarks
Arifa Khan, PhD, Food and Drug Administration
Philip Krause, MD, Food and Drug Administration
Presentation

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

Clinical Biomarkers and Personalized Medicine

Andrea Ferreira-Gonzalez, PhD, Professor, Virginia Commonwealth University Medical Center
Andrew Grupe, PhD, Senior Scientific Director, Pharmacogenomics, Celera, Quest Diagnostics
Charles Sawyers, MD, Chair, HOPP, Memorial Sloan Kettering Cancer Center
Laura J. van't Veer, PhD, Leader, BOP, Associate Director, Applied Genomics, UCSF HDFC Cancer Center

Chairperson's Opening Remarks
Eric Donaldson, PhD, Food and Drug Administration
Raja Mazumder, PhD, George Washington University
Presentation

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.

Next-Generation Sequencing Devices and Clinical Applications

- Ira Lubin, PhD, FACMG, Team Lead, Genetics, OPHSS, Centers for Disease Control and Prevention
- Justin Zook, PhD, Researcher, Genome in a Bottle, National Institute of Standards and Technology
- Mya Thomae, Vice President, Regulatory Affairs, Illumina
- Heike Sichtig, PhD, Principle Investigator, Regulatory Scientist, CDRH, Food and Drug Administration

Chairperson's Opening Remarks
Zivana Tezak, PhD, Food and Drug Administration
Justin Zook, PhD, National Institute of Health
Heike Sichtig, PhD, Food and Drug Administration

Presentation

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.

Food Safety and Pathogen Detection

- Marc Allard, PhD, Research Microbiologist, CFSAN, Food and Drug Administration
- Bill Klimke, PhD, Staff Scientist, NCBI, National Institute of Health
- Kristin G. Holt, DVM, MPH, FSIS Liaison to CDC, Food Safety and Inspection Service, OPHS, USDA

Chairperson's Opening Remarks
Errol Strain, PhD, Food and Drug Administration
Heike Sichtig, PhD, Food and Drug Administration

Presentation

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.