FDA’s Role In Building the ID NGS Diagnostic Toolkit

Heike Sichtig, PhD
Subject Matter Expert
Principal Investigator
CDRH/OIR
Disclaimer

The information in these materials is not a formal dissemination of information by FDA and does not represent agency position or policy.

Opinions are my own
Organizational Chart
Resources For You

Medical Devices

• Emergency Use Authorizations
  – https://www.fda.gov/MedicalDevices/Safety/EmergencySituations/ucm161496.htm

• Device Advice: Comprehensive Regulatory Assistance
  – https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/default.htm

• Recently-Approved Devices (PMA, 510(k), HDE, De novo)
  – https://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/DeviceApprovalsandClearances/Recently-ApprovedDevices/default.htm

• Classify Your Medical Device
  – https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/ClassifyYourDevice/default.htm

• Science and Research (FDA-ARGOS Database)
  – https://www.fda.gov/MedicalDevices/ScienceandResearch/DatabaseforReferenceGradeMicrobialSequences/default.htm
**In Vitro Diagnostic Devices**

**Definition:**

Reagents, instruments, and systems intended for use in the diagnosis of disease or other conditions, including a determination of the state of health, in order to cure, mitigate, treat, or prevent disease or its sequelae. ... for use in the collection, preparation, and examination of specimens from the human body. [21 CFR 809.3]

**US FDA Regulated Uses:**

- Detection and Diagnosis
- Screening
- First Response

- **Not** Environmental Screening
Device Classification

A device should be placed in the lowest class whose level of control will provide reasonable assurance of safety and effectiveness.
Risk Based Regulation of IVDs

Class I - Low likelihood of harm or risk can be mitigated
- General Controls

Class II - Moderate likelihood of harm or risk can be mitigated
- Special Controls

Class III - High or unknown likelihood of harm
- Significant Risk
- Pre-market Approval

Knowledge Mitigates Risk

Class I most 510(k) exempt

Knowledge

Risk

Class III - PMA
Different Use, Same Test

- A CFTR genotyping multiplex assay on the same instrument with the indication
  - for aid in diagnosis $\rightarrow$ 510(k)
  - for fetal screening $\rightarrow$ PMA

- A breast cancer assay to be used
  - for screening, diagnosis $\rightarrow$ PMA
  - for prognosis in already diagnosed patients $\rightarrow$ 510(k)
There is no FDA cleared NGS system for sequencing of microbial genomic DNA for identification of microbial targets or detection of virulence or resistance genes.

**FDA Current Thinking**

**NGS Technologies**

**Targeted** (*amplicon, bioinformatics*)
- Scope limited to defined regions that target a specific organism(s), gene(s) or marker(s).
- Targets are selected *a priori* by any lab or bioinformatics method (e.g., amplicon sequencing or a k-mer signature database) based on the diagnostic devices intended use.

**Hypothesis-Free** (*whole genome, shotgun*)
- No *a priori* knowledge of targets.
- Generally can identify all constituents (e.g., organism(s), gene(s) or marker(s), microbiota, human background, and contaminants) in a sample.

**Sample Applications**
- Single Target (Pathogen, Gene, Marker)
- Pathogen/Marker Panel
- Gene Panel (16S)
- Metagenomics
- Novel and Emerging Pathogens
De Novo Regulatory Pathway

The De Novo process provides a pathway to classify novel medical devices for which general controls alone, or general and special controls, provide reasonable assurance of safety and effectiveness for the intended use, but for which there is no legally marketed predicate device. De Novo classification is a risk-based classification process.

There are two options for when a sponsor can submit a De Novo request for the FDA to make a risk-based evaluation for classification of the device into class I or II.

• Option 1: After receiving a high-level not substantially equivalent (NSE) determination (i.e., new intended use and/or different technological characteristics that raise different questions of safety and effectiveness) in response to a 510(k) submission.

• Option 2: Upon the sponsor’s determination that there is no legally marketed device upon which to base a determination of substantial equivalence (therefore without first submitting a 510(k) and receiving a high-level NSE determination).

Prior to submitting a De Novo request, it is recommended that you consider submitting a pre-submission (pre-sub) to obtain feedback from the appropriate premarket review division.

De Novo Classification Process (Evaluation of Automatic Class III Designation) - Guidance for Industry and Food and Drug Administration Staff (PDF - 139KB)
Risk-Based Evaluation

- Real clinical samples where feasible
  - Prevalence of analyte is low? Consult with FDA about alternative sample types
- Prospective or retrospective evaluation
  - Comparison to a reference method
  - Comparison to a predicate device
  - Comparison to a clinical outcome
- In-Silico evaluation
  - FDA-ARGOS Reference-Grade Genomes (Bioproject 231221)
  - Mixed Microbial Reference Material
ID NGS Diagnostic Toolkit Needs

• ID NGS Diagnostic Assay
• Tools to support regulatory review
  – FDA-ARGOS Reference-Grade Genomes for regulatory use to enable sponsor to perform in-silico validation of claims
  – Mixed Microbial Reference Materials that sufficiently challenge the entire ID NGS Diagnostic Assay workflow
    • Ideally cell-based
    • Performance for assay’s intended use
FDA Tools for ID NGS Dx

**FDA-ARGOS Database**
:microbial reference genomes for **regulatory use**

- More flexible regulatory pathway
  - Enable In-silico analytical and clinical validation
  - Reduce testing burden
- Reference database

**Interagency ID NGS SME Working Group**
: team of NGS agency-wide subject matter experts

- ID NGS Dx Advisory Board
- **Consensus** FDA-ARGOS genome vetting
- Keep current on state of the art
- Tackle open questions (i.e. sens/spec)
- NGS Reference Material
Reference Genomes
For Regulatory Use

Support in-silico analytical and clinical validation

A. Identified by orthogonal reference method
B. Sequenced and de-novo assembled using 2 sequencing methodologies
C. High depth of sequencing coverage
D. Minimum of 20X over 95 percent of the assembled and polished core genome
E. Taxonomy-specific ANI thresholds that are sufficient for identification
F. Placed within a pre-established phylogenetic tree
G. Sample specific metadata, raw reads, assemblies, annotation and details of the bioinformatics pipeline are available
Government-Academic-Clinical Partnership

- In May 2014, the FDA and collaborators established FDA-ARGOS (www.fda.gov/argos)
- With funding support from FDA’s Office of Counterterrorism and Emerging Threats (OCET) and DoD, the FDA-ARGOS team are initially collecting and sequencing 2000 microbes that include biothreat microorganisms, common clinical pathogens and closely related species.
- Currently, FDA-ARGOS microbial genomes are generated in 3 phases.
  - Phase 1 entails collection of a previously identified microbe and nucleic acid extraction from government, academic and clinical collaborators (>30).
  - Phase 2, the microbial nucleic acids are sequenced and de novo assembled using Illumina and Pac Bio sequencing platforms at the Institute for Genome Sciences at the University of Maryland (UMD-IGS).
  - Phase 3, the assembled genomes are vetted by an ID-NGS subject matter expert working group consisting of FDA personnel and collaborators and the data are deposited in NCBI databases.

Principal Investigator: Heike Sichtig, Ph.D.
Center for Devices | U.S. Food and Drug Administration | Heike.Sichtig@fda.hhs.gov
FDA dAtabase for Reference Grade micrObial Sequences (FDA-ARGOS)


To get all associated genbank entries, select the Nucleotide database and enter this search term: ‘231221[BioProject]’

GenBank records (annotations, not RefSeq):

BioSamples:
[https://www.ncbi.nlm.nih.gov/biosample?Db=biosample&DbFrom=bio project&Cmd=Link&LinkName=bio project_biosample&LinkReadableName=BioSample&ordinalpos=1&IdsFromResult=231221](https://www.ncbi.nlm.nih.gov/biosample?Db=biosample&DbFrom=bio project&Cmd=Link&LinkName=bio project_biosample&LinkReadableName=BioSample&ordinalpos=1&IdsFromResult=231221)

Assemblies:

Raw reads:

- American Type Culture Collection/ BEI
- Bernard Nocht Institute for Tropical Medicine, Germany
- Biodefense and Emerging Infections Research Repository
- British Columbia Centre for Disease Control (BCCDC)
- Children’s National Medical Center
- Defense Threat Reduction Agency (DTRA)
- George Washington University
- Joint Program Executive Office for Chemical and Biological Defense (JPEO-CBD)
- Lawrence Livermore National Lab (LLNL)
- Los Alamos National Lab (LANL)
- Mayo Clinic
- National Biodefense Analysis and Countermeasures Center
- National Institute of Allergy and Infectious Diseases (NIH-NIAID)
- New York State Wadsworth Laboratories
- Public Health Agency Canada (PHAC)
- Public Health England (PHE)
- Rockefeller University
- Rutgers University
- Stanford University Medical Center
- University of California, San Francisco (UCSF)
- University of Colorado Denver
- University of Ibadan, Nigeria
- University of Louisville
- University of Michigan
- University of North Carolina at Chapel Hill
- University of Texas Medical Branch (UTMB)
- University of Washington School of Medicine
- U.S. Army Edgewood Chemical Biological Center (ECBC)
- U.S. Army Medical Research Institute for Infectious Diseases (USAMRIID)
- U.S. Food and Drug Administration
- Weill Cornell Medicine
In-Silico Comparator Example

DoD/USAMRIID Collaboration

- Sequencing-based diagnostic device
- Generate FDA-ARGOS Reference Genomes
- Datasets for Regulatory Approval

> Enable In-Silico Data Analysis

### Endemic African Diseases

- Chikungunya virus
- Crimean-Congo hemorrhagic fever virus
- Dengue virus serotype 1
- Dengue virus serotype 2
- Dengue virus serotype 3
- Dengue virus serotype 4
- Ebola virus
- Lassa virus
- Marburg virus (Angola)
- Marburg virus (Ci67)
- *Plasmodium falciparum*
- Rift Valley fever virus
- West Nile virus
- Yellow fever virus
- Zika virus
FDA-ARGOS Pipeline

FDA-ARGOS microbial genomes are generated in 3 phases:
Phase 1- collection of a previously identified microbe and nucleic acid extraction
Phase 2- sequencing and de novo assembly at UMD
Phase 3- Vetting and data deposit in NCBI databases

FDA-ARGOS Reference Genome Characteristics:
A. Identified by orthogonal reference method
B. Sequenced and de-novo assembled using 2 sequencing methodologies
C. High depth of sequencing coverage
D. Minimum of 20X over 95 percent of the assembled and polished core genome
E. Taxonomy-specific ANI thresholds that are sufficient for identification
F. Placed within a pre-established phylogenetic tree.
G. Sample specific metadata, raw reads, assemblies, annotation and details of the bioinformatics pipeline are available.
# Bacteria/Fungi/Eukaryote Pipeline

<table>
<thead>
<tr>
<th>Step</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collect samples /grow samples</td>
</tr>
<tr>
<td>Extract samples</td>
</tr>
<tr>
<td>Q/C extractions at IGS (&gt;10ug cutoff)</td>
</tr>
<tr>
<td>Batch and library prep/sequence on Illumina –Megablast QC</td>
</tr>
<tr>
<td>Library prep/sequence on PacBio –Megablast QC</td>
</tr>
<tr>
<td>Assemble long/short raw reads</td>
</tr>
<tr>
<td>Annotate with in-house pipeline for Q/C</td>
</tr>
<tr>
<td>Data Analytics Q/C Pipeline at FDA</td>
</tr>
<tr>
<td>Register BioSamples and submit raw reads to SRA DB and assemblies to Assembly DB</td>
</tr>
<tr>
<td>NCBI annotates genomes</td>
</tr>
</tbody>
</table>
## Viral Pipeline

1. **Q/C extracted genomic material at IGS (25ng)**
2. **Library Prep/sequence on Illumina**
   - Shotgun
   - Amplicon (may require primer set design –Ebola, Zika)
     - Looking into WNV, Dengue –Broad Institute
     - RACE
3. **Assemble raw reads**
4. **Data Analytics Q/C Pipeline at FDA**
5. **Register BioSamples and submit raw reads to SRA DB and assemblies to Assembly DB**
FDA-ARGOS Sample Progress

• 970 Genomes sequenced
  – 872 bacterial
    • 95 genera
  – 85 viral
    • 9 viral types
      • Ebola, Zika, Dengue, WNV, etc.
    – 13 fungal/parasite
  • 30+ collaborators
Endemic African Diseases

Chikungunya virus
Crimean-Congo
Hemorrhagic Fever virus
Dengue virus serotype 1
Dengue virus serotype 2
Dengue virus serotype 3
Dengue virus serotype 4

Ebola virus
Lassa virus
Marburg virus (Angola)
Marburg virus (Ci67)

Plasmodium falciparum
Rift Valley fever virus
West Nile virus

Yellow fever virus
Zika virus

Non-Curated Database

Misdiagnosis:
- False Positives
- False Negatives

Correct Diagnosis:
- True Positives
- True Negatives

- Minimize Misdiagnosis
- Evolutionary Change
- Rapid Diagnostics
Sustainable Solution
Contributing Genomes to FDA-ARGOS

• Further population and curation of the database will support the success of FDA-ARGOS and promote adoption by the community.

• The FDA-ARGOS team openly invites additional collaborators from the scientific community to assist in filling the gaps in this public resource.

• Specifically searching for unique, hard to source microbes such as biothreat organisms, emerging pathogens, and clinically significant bacterial, viral, fungal, and parasitic genomes.

• The goal is to collect sequence information for a minimum of 5 isolates per species.

• For more information about contributing samples for UMD-IGS sequencing as part of FDA-ARGOS efforts, or to qualify existing genomes by the FDA, please email FDA-ARGOS@fda.hhs.gov.
External Genome Submission

- **Submitter**
  - Sample Metadata
    - NCBI BioSample
  - Sequencing Pipeline
    - Protocols
  - Raw Reads
    - NCBI SRA
  - Bioinformatics Pipeline
    - Protocols or BioCompute Object
  - Assemblies
    - NCBI Assembly
  - Consensus Genome
    - NCBI GenBank

- **FDA-ARGOS**
Reference Materials

• Support analytical validation of entire ID NGS Diagnostic assay workflow
  1. Cell-based in clinical matrix (blood, urine, stool) to test from specimen collection to result
  2. Reference material organism panel should sufficiently capture ID NGS assay’s claimed target characteristics (intended use)
     • Size of the genome, G/C content, DNA/RNA, Near neighbors, Repetitive content, Commensal, Extremes
• Cross-platform comparison
NIST/FDA Reference Material Efforts

• Microbial Genomic DNA Reference Material
  – RM 8375 - Microbial genomic DNA standards for sequencing performance assessment
  – 2 FDA-ARGOS strains/ 2 FDA-CFSAN strains

• Mixed Pathogen DNA Research Material
  – A mixture of genomic DNA from 25 clinically-relevant pathogens plus human genomic DNA.

Build Reference Genomes:

• PacBio/Illumina sequencing of microbial constituents as part of FDA-ARGOS project
Other Reference Material Efforts

• ZymoBIOMICS™ Microbial Community Standards by Zymo Research
  – A mock microbial community consisting of eight bacterial and two fungal strains

• UCSF Control Material

Build Reference Genomes:

• PacBio/Illumina sequencing of microbial constituents as part of FDA-ARGOS project
Which of the following are characteristics of Reference-Grade Genomes for Regulatory Use?

- HMW genomic material from unknown organism
- Sequenced and de-novo assembled using 2 sequencing methodologies
- High depth of sequencing coverage
- Minimum of 1X over 95 percent of the assembled and polished core genome
- Generalized ANI threshold
- Placed within a pre-established phylogenetic tree
- Sample specific metadata, raw reads, assemblies, annotation and details of the bioinformatics pipeline are available.
Which of the following are characteristics of Reference-Grade Genomes for Regulatory Use?

- HMW genomic material from unknown organism
- Sequenced and de-novo assembled using 2 sequencing methodologies
- High depth of sequencing coverage
- Minimum of 1X over 95 percent of the assembled and polished core genome
- Generalized ANI threshold
- Placed within a pre-established phylogenetic tree
- Sample specific metadata, raw reads, assemblies, annotation and details of the bioinformatics pipeline are available.
REVIEW
Characteristics of Reference-Grade Genomes for Regulatory Use

A. Identified by orthogonal reference method
B. Sequenced and de-novo assembled using 2 sequencing methodologies
C. High depth of sequencing coverage
D. Minimum of 20X over 95 percent of the assembled and polished core genome
E. Taxonomy-specific ANI thresholds that are sufficient for microbial organism identification
F. Placed within a pre-established phylogenetic tree
G. Sample specific metadata, raw reads, assemblies, annotation and details of the bioinformatics pipeline are available
FDA-ARGOS team members include representatives from the:

- U.S. Food and Drug Administration
- U.S. Department of Defense
- National Institutes of Health
- Institute for Genome Sciences at University of Maryland

**Funding Agencies**

FDA’s Office of Counterterrorism and Emerging Threats
Joint Program Executive Office for Chemical and Biological Defense (JPEO-CBD)