

FDA's Role and Tools for ID-NGS Diagnostics



Heike Sichtig, MS/PhD

Subject Matter Expert
Principal Investigator

Center for Devices
Division of Microbiology Devices



DHS Sequencing Meeting 2017

Disclaimer

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



Infectious Disease (ID) NGS Dx

There is **no FDA cleared NGS instrument for sequencing of microbial genomic DNA** for identification of microbial targets or detection of virulence or resistance genes.

FEDERAL REGISTER
The Daily Journal of the United States Government

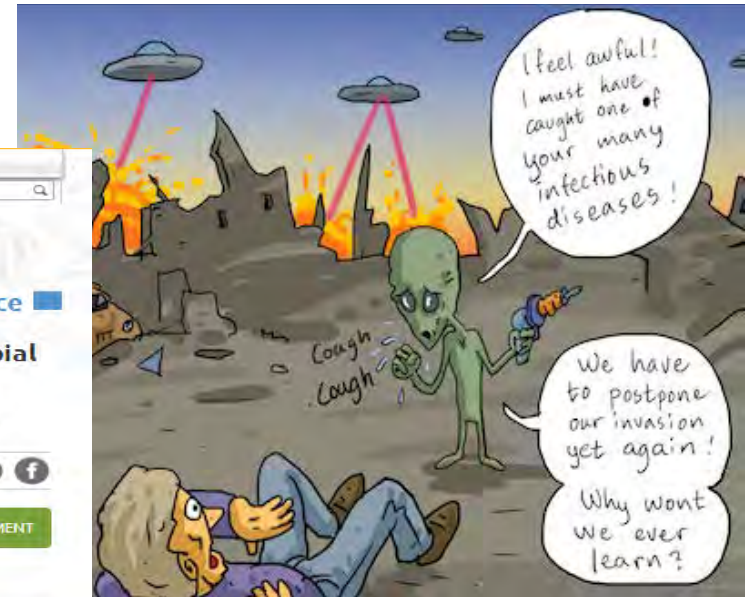
Notice

Infectious Disease Next Generation Sequencing Based Diagnostic Devices: Microbial Identification and Detection of Antimicrobial Resistance and Virulence Markers; Draft Guidance for Industry and Food and Drug Administration Staff; Availability

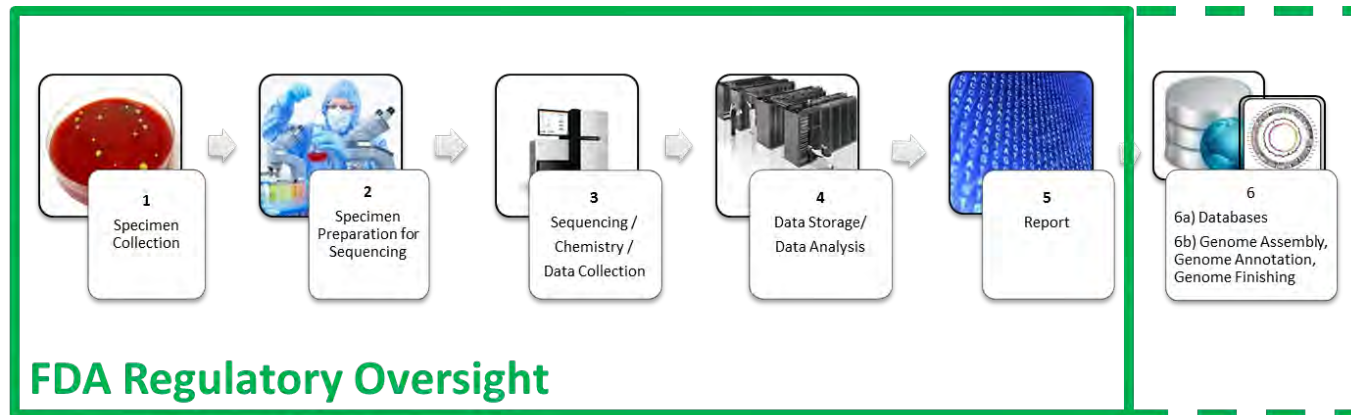
A Notice by the Food and Drug Administration on 05/13/2016

Comments on this document are being accepted at Regulations.gov

SUMMARY The Food and Drug Administration (FDA) is announcing the availability of a draft guidance entitled "Infectious Disease Next Generation Sequencing Based Diagnostic Devices: Microbial Identification and Detection of Antimicrobial Resistance and Virulence Markers." This draft guidance provides recommendations to assist industry in designing studies to establish the analytical and



FDA Current Thinking



NGS Technologies

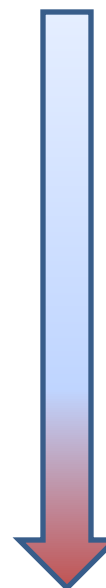
Targeted (*amplicon*)

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

Agnostic (*whole genome, shotgun*)

- No *apriori* 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

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 Team

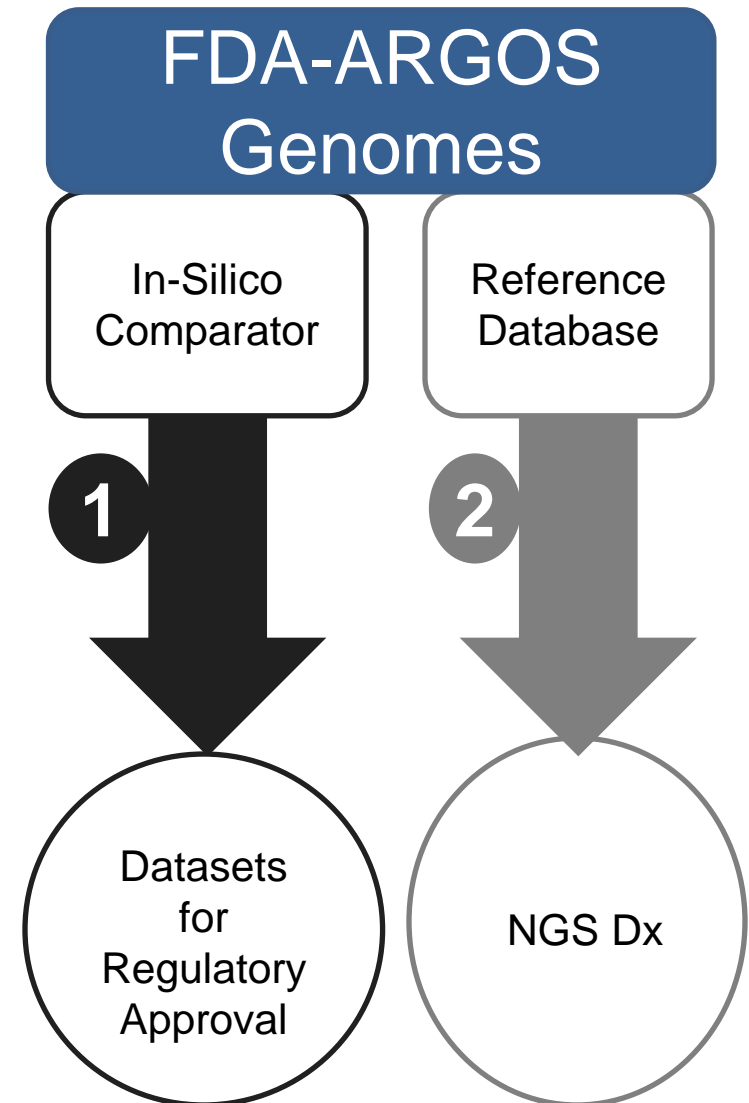
: team of NGS subject matter experts

- ✓ ID NGS Dx Advisory Board
- ✓ Consensus FDA-ARGOS genome vetting
- ✓ Keep current on state of the art
- ✓ Tackle open questions (i.e. sens/spec)

FDA-ARGOS: Goal and Use

- Public Vetted Resource
- Microbial Reference-Grade Genomes for **Regulatory Use**
- US-Initiated
- Medical Countermeasures
- Common clinical
- Near neighbors

- Coverage for US Needs
- Currently not funded to support *Needs for Developing World* and associated *Global Standards*



DoD Collaborations

- USAMRIID
 - PreEUA NGS data prep
 - Enable In Silico validation
- DBPAO
 - West Africa panel
 - Full assay characterization
 - Analytical
 - Clinical
 - Inclusivity/Exclusivity

Pathogen
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)
<i>Plasmodium falciparum</i>
Rift Valley fever virus
West Nile virus
Yellow fever virus
Zika virus

USAMRIID Targeted NGS (environmental)



MTE NGS Pathogen Detection

<i>Borrelia burgdorferi</i> Bundibugyo virus Bunyamwera virus Candiru virus Chikungunya virus Cowpox virus <i>Coxiella burnetii</i> Crimean-Congo hemorrhagic fever virus dengue virus serotype 1 dengue virus serotype 2 dengue virus serotype 3 dengue virus serotype 4 eastern equine encephalitis virus Ebola virus (Zaire) Ebola virus (Kikwit) Getah virus Guanarito virus Hantaan virus Heartland virus Hendra virus Highlands J virus	Human coronavirus 229E Human parainfluenza virus Type 1 Human parainfluenza virus Type 2 Human parainfluenza virus Type 4b Human respiratory syncytial virus (RSV) A-2 Human respiratory syncytial virus (RSV) 18537 influenza B virus (B/Ohio/01/2005) Junín virus La Crosse virus Lassa fever virus Lujo virus Lymphocytic choriomeningitis virus Machupo virus (Malele) Machupo virus (Chicava) Marburg virus (Angola) Marburg virus (Ci67) Marburg virus (Musoke) Mayaro fever virus Monkeypox virus Naples virus Ndumu virus	Nipah virus Onyong-nyong virus Oropouche virus (TR9760) <i>Plasmodium falciparum</i> Powassan virus Ravn virus Reston virus Rift Valley fever virus Ross River virus SARS coronavirus Semliki forest virus Sindbis virus St. Louis encephalitis virus Taï Forest virus Toscana virus Venezuelan equine encephalitis virus Wesselbron virus western equine encephalitis virus yellow fever virus Zika virus
---	---	---

Sequencing Approach

Bacteria

- **Hybrid sequencing** approach using Illumina HiSeq2000 and the BacBio RSII platform to generate industry standard high quality sequences. Use of multiple assemblers. 3 sets of de novo genome assemblies will be produced 1) Illumina only, 2) BioPac only, and 3) Illumina/BioPac hybrid

Virus

- IGS will use existing and well-established laboratory and bioinformatics pipelines within the Genomic Resource Center. A **three-prong** Illumina sequencing approach followed by customized assembly

FDA-ARGOS Genome Status

- There are **827 (bacterial, viral)** samples currently at various stages within the FDA-ARGOS sequencing pipeline.
- **322 (bacterial, viral)** genomes from other efforts (i.e. TTC) to be qualified.
- Goal is to collect and sequence **2000** gap organisms

Overall pipeline

Collaborator -> FDA OSEL -> UMD/IGS -> NCBI/FDA

Collaborator -> USAMRIID -> UMD/IGS -> NCBI/FDA

*Stocks are archived at USAMRIID when available per agreement

NCBI BioProject 231221

:houses FDA-ARGOS genomes generated with the IGS-UMD Sequencing pipeline

NCBI
Resources
How To
hsichig
My NCBI
Sign Out

BioProject
BioProject
Advanced
Search
Help

Display Settings:
Send to:

FDA dAtabase for Regulatory Grade microBial Sequences (FDA-ARGOS):
Supporting development and validation of Infectious Disease Dx tests

Accession: PRJNA231221 ID: 231221

FDA, in collaboration with the National Center for Biotechnology Information (NCBI), the Department of Defense (DoD) and the Institute for Genome Sciences at the University of Maryland, established a publicly available, well-curated reference database (FDA-ARGOS: dAtabase for Regulatory Grade microBial Sequences; BioProject 231221) of regulatory-grade-quality sequences from diverse infectious microorganisms. More...

Accession	PRJNA231221
Data Type	Genome sequencing
Scope	Multispecies
Keyword	GMI
Grants	"Enhancement of Microbial Sequence Quality for Regulatory and Clinical Decision Processes Using High Throughput Sequencing Technologies" (Grant ID TBD, Food and Drug Administration)
Submission	Registration date: 11-Dec-2013 - US Food and Drug Administration - University of Maryland School of Medicine Institute for Genome Sciences (IGS) - sequencing center
NCBI Links	NCBI Pathogen Detection
Relevance	Medical

Related information

- Assembly
- BioSample
- Genome
- Genomic DNA
- Genomic RNA
- Nucleotide
- Protein
- Related Genes
- SRA
- WGS master

Related Resources

NCBI Pathogen Detection

LinkOut to external resources

GOLDCARD: Gp0131387 [Genomes On Line Database]

GOLDCARD: Gp0131384 [Genomes On Line Database]

GOLDCARD: Gp0131375 [Genomes On Line Database]

2871180282: *Listeria monocytogenes* NRRL B-57803 [Integrated Microbial Genomes]

2883763271: *Staphylococcus aureus* FDAARGOS_140 [Integrated Microbial Genomes]

2887528123: *Acinetobacter baumannii* FDAARGOS_131 [Integrated Microbial Genomes]

SILVA SSU Database [SILVA]

SILVA LSU Database [SILVA]

Project Data:

Resource Name	Number of Links
SEQUENCE DATA	
Nucleotide (total)	743
WGS master	125
Genomic DNA	76
Genomic RNA	8
SRA Experiments	545
Protein Sequences	568269
OTHER DATA TYPES	
BioSample	265
Assembly	157

Assembly details:

Assembly level	Number of Assemblies
Complete genome	33

Recent activity

Turn Off Clear

FDA-ARGOS

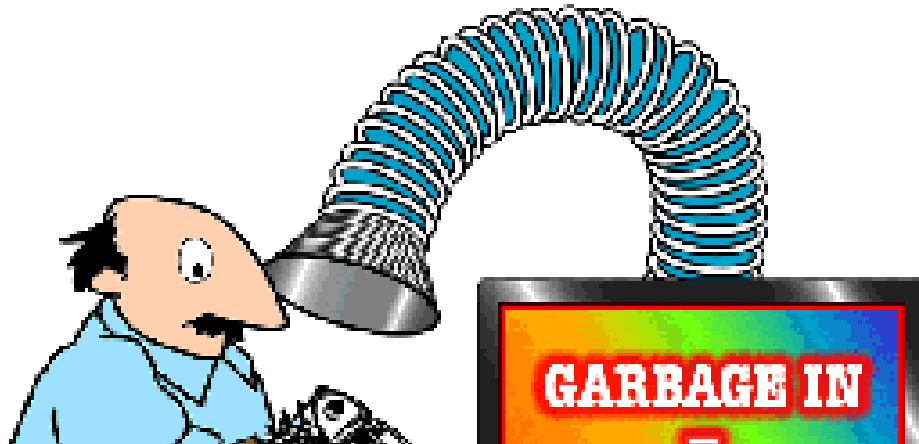


COMPREHENSIVE
ACCURACY
EMERGING THREATS
Precision
Inclusivity
Rigorous
Temporal
VERACITY
Robustness
Specificity
Geography
Sustainability
Spatial
Timeliness

Global-ARGOS



Accuracy

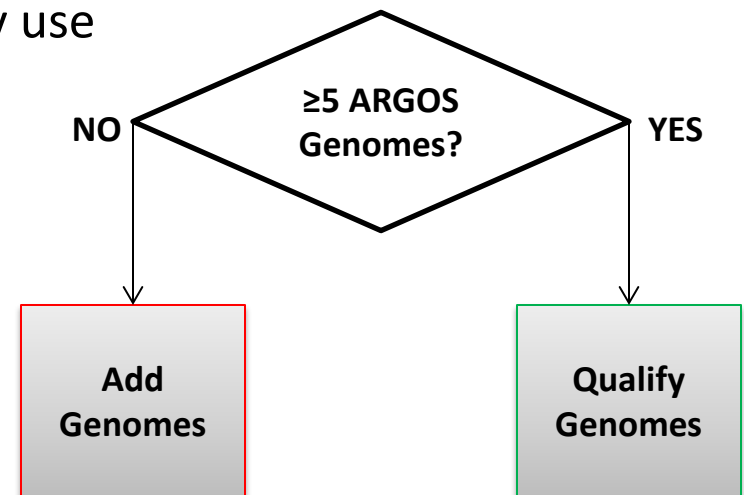


Why Reference Grade Genomes for Regulatory Use?

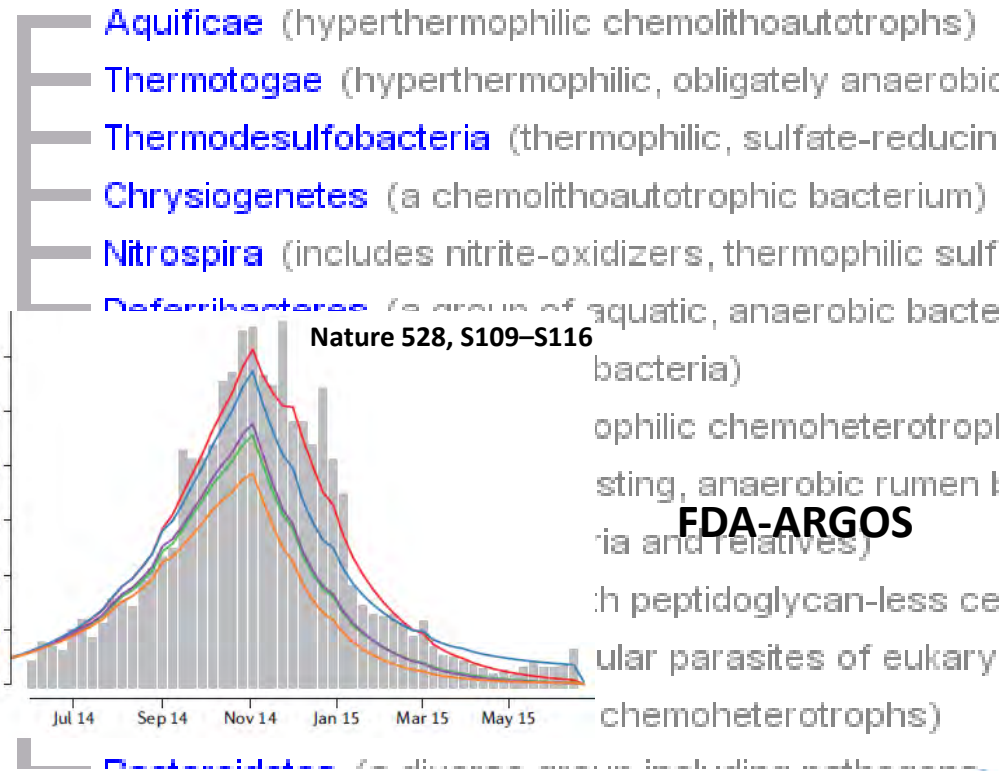
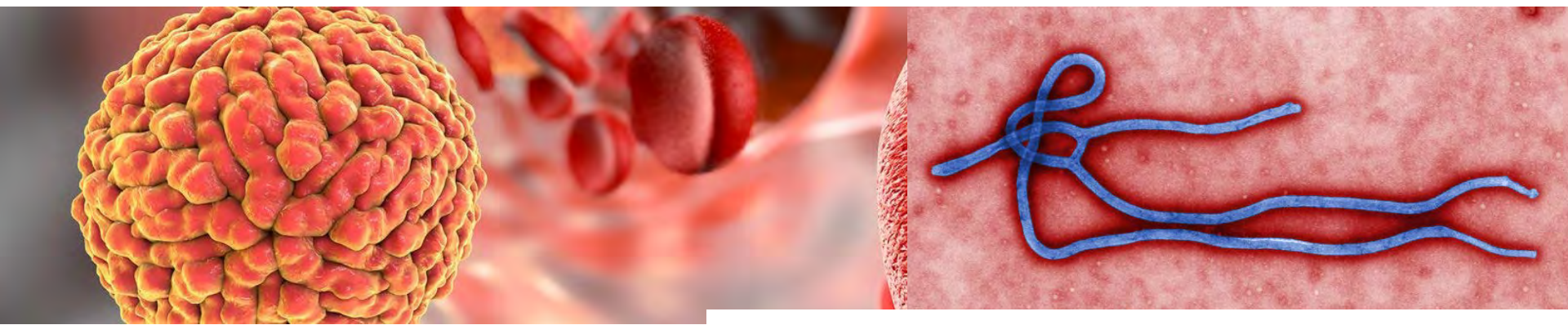


Species-specific Quality Metrics

- Vetted by ID NGS SME Team for regulatory use
- Genome similarity measure and range
 - Based on at least 5 representatives
 - Nucleotide similarity (ANI)
 - Species core genome
 - Minimum 20X over 95%
- Sample-specific Metadata
- Raw reads, assemblies, annotations
- Pipeline
 - Sample Prep, Library , Platform, Assembler



Coverage: Spatial/Temporal



Reference Genome Gap: Ebola

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

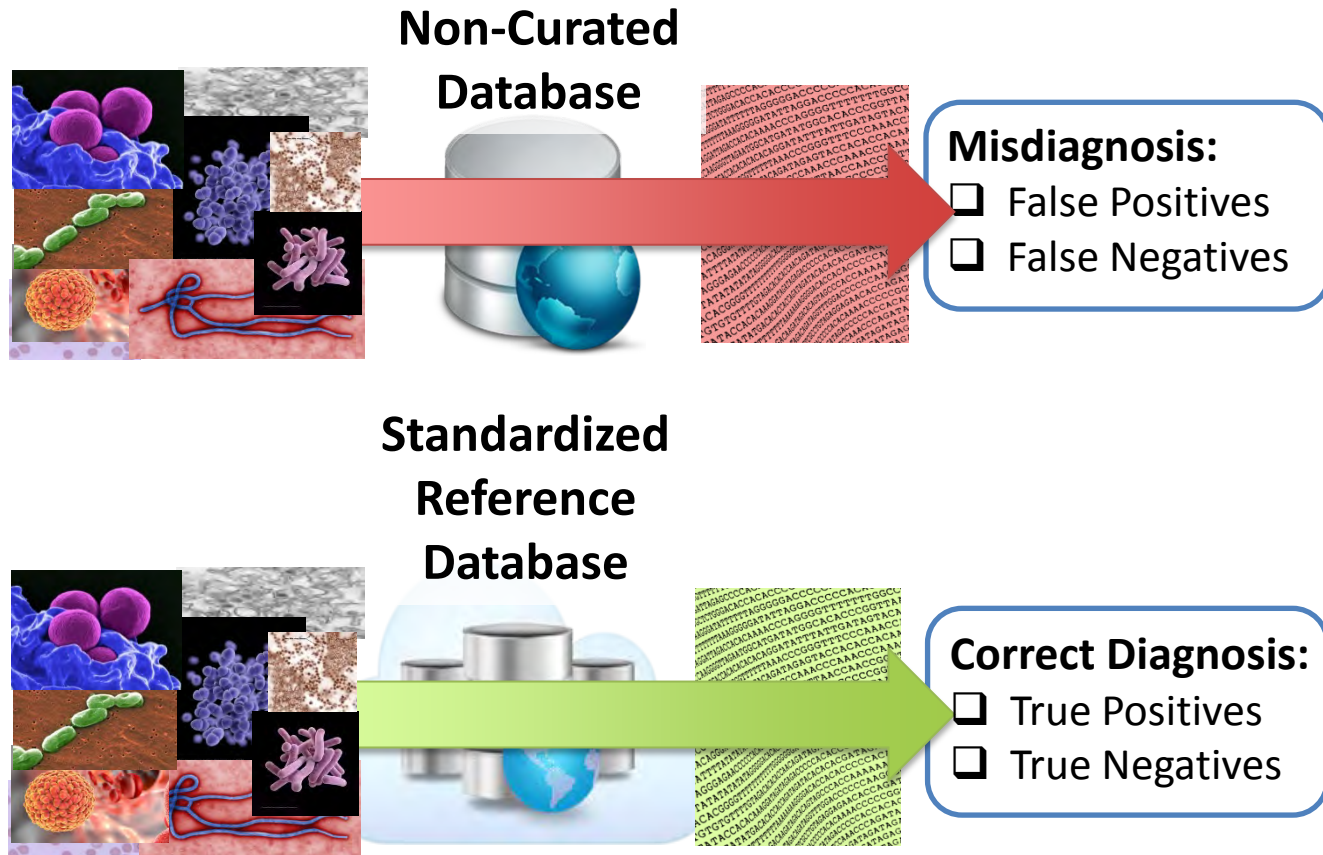
Lassa virus
 Marburg virus (Angola)
 Marburg virus (Ci67)

Plasmodium falciparum

Rift Valley fever virus
 West Nile virus

Yellow fever virus

Zika virus



✓ *Minimize Misdiagnosis*

✓ *Evolutionary Change*

✓ *Rapid Diagnostics*



FDA-----
 ---d**A**tabase for
 ----**R**eference-
 ----**G**rade-----
 micr**O**bial-----
 -----**S**equences



✓ Coverage of circulating strain in reference database first step to catch the bug



Decoding Ebola: Next-Generation Sequencing of the Ebola Genome for the FDA ARGOS Database

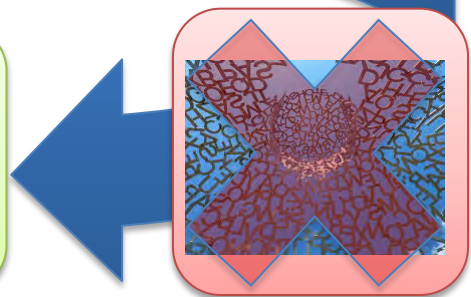
When you need a test to confirm disease in an outbreak, you need it fast. Global

Infections Disease NGS Dx

Platform?

Global ARGOS

Prevent Outbreak



Sustainable Solution

