

Applications of Pathogen Genomics in Public Health: National Level

Duncan MacCannell PhD MBT

DIRECTOR, OFFICE OF ADVANCED MOLECULAR DETECTION

Accelerating the Use of Pathogen Genomics and Metagenomics in Public Health

National Academies of Science, Engineering and Medicine

July 22nd, 2024



Land Acknowledgement: The Division of Infectious Disease Readiness and Innovation acknowledges the Muscogee (Creek) and Cherokee Nations and the Dena'ina people whose Indigenous lands house the CDC Atlanta and AIP Anchorage campuses.

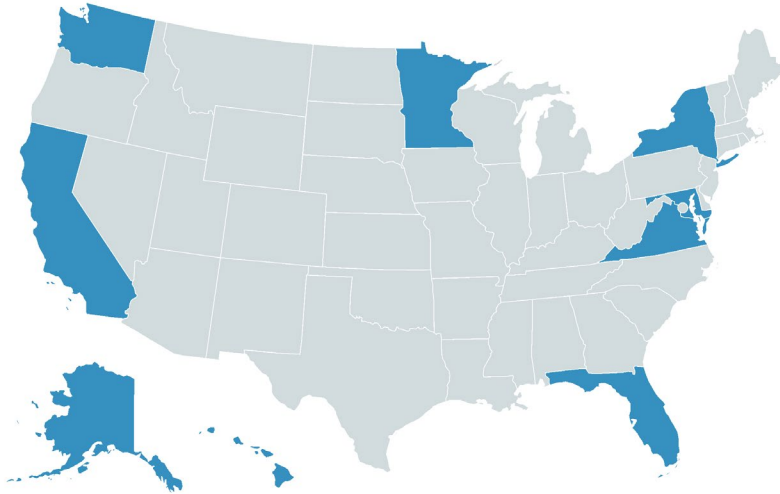
Advanced Molecular Detection (AMD)

- FY2014: **Advanced Molecular Detection and Response to Infectious Disease Outbreaks (AMD)** budget initiative
 - \$40M/year public health innovation program.
 - Focus on emerging laboratory technologies, and the infrastructure and workforce changes needed to sustain them across public health system.
- **Primary Objectives of AMD:**
 - Improve pathogen detection and characterization with emerging technologies.
 - Build integrated information systems and platforms to transform data into action.
 - Strengthen technical workforce, and improve equitable access to technology and skills across the public health system.
 - Promote open data standards, interoperability and reproducible methods to meet a range of public health surveillance and response needs.

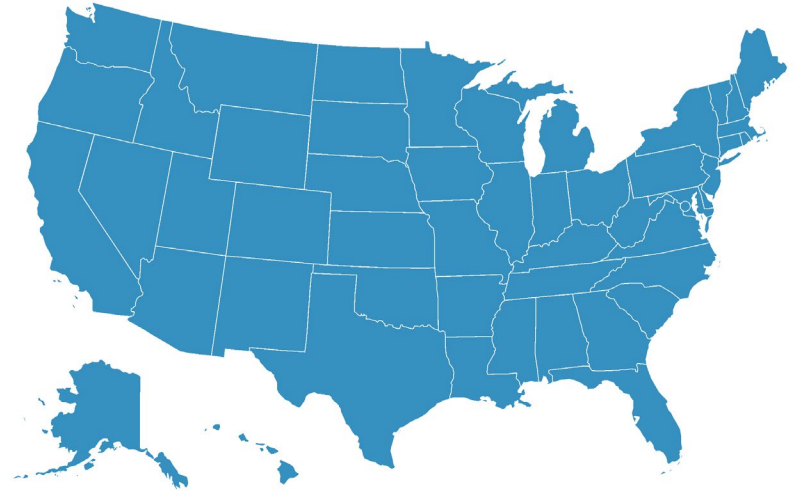


Pathogen Genomics Capacity in US Public Health

2013



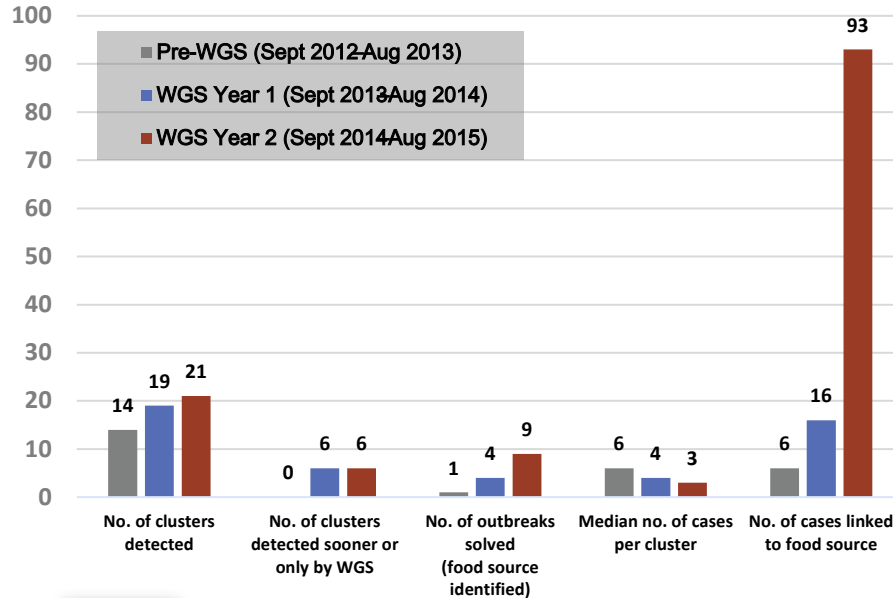
2018



- Foodborne disease surveillance (PulseNet/GenomeTrackr) were early drivers for the adoption of NGS.
- PHLs incorporated sequencing into surveillance and response for national, regional and local priorities.



Listeria monocytogenes: early proof of concept for pathogen genomic surveillance at a national scale



- *Listeria monocytogenes* causes severe foodborne illness, with roughly 1,000 cases in the United States each year.
- September 2013: pilot project to rapidly and comprehensively sequence clinical, environmental and food isolates.
 - Collaboration between federal (CDC, FDA, USDA-FSIS, NCBI), state and international partners.
 - Goal: assess the impact of WGS and feasibility for more common foodborne pathogens.
- Marked improvements in cluster identification & definitive source attribution.

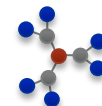
Examples of other CDC pathogen genomics activities

	PATHOGEN	ANALYSIS PLATFORM	SEQUENCING		ANALYSIS		DEFAULT SEQUENCE REPOSITORY	PIPELINE COMPONENTS AND SOURCE CODE
			CDC	STATE	CDC	STATE		
	Foodborne Bacteria	Bionumerics / PulseNet 2.0	■	■	■	■	NCBI SRA	TBD
	Antimicrobial Resistant Bacteria / HAIs	Bionumerics / PHoeNix	■	■	■	■	NCBI SRA	https://github.com/CDCgov/phoenix
	<i>Mycobacterium tuberculosis</i>	Bionumerics / TB GIMS	■	■	■	■		TBD https://github.com/CDCgov/MicrobeTrace
	<i>Candida auris</i>	MycosNP-nf	■	■	■	■	NCBI SRA	https://github.com/CDCgov/mycosnp-nf
	<i>Legionella pneumophila</i>	Bionumerics / Custom	■	■	■	■	NCBI SRA	TBD
	Viral hepatitis	GHOST	■	■	■	■	NCBI	
	Influenza	Influenza Genomic Surveillance	■	■	■	■	NCBI / GISAID	https://hub.docker.com/r/cdcgov/irma https://github.com/cdcgov/mira
	Arboviruses	Custom / Various	■	■	■	■	NCBI	TBD
	Norovirus	Bionumerics / CaliciNet 2.0	■	■	■	■	NCBI	TBD
	mpox	Custom / PolkaPox	■	■	■	■	NCBI	https://github.com/CDCgov/polkapox
	Measles	Custom / VPipe	■	■	■	■	NCBI / MeaNS2	TBD
	Malaria	MaRS	■	■	■	■	NCBI	https://github.com/CDCgov/MaRS
	SARSCoV-2	NS3+Baseline	■	■	■	■	NCBI / GISAID	https://hub.docker.com/r/cdcgov/irma https://github.com/cdcgov/mira

NETWORK MODELS



CENTRALIZED

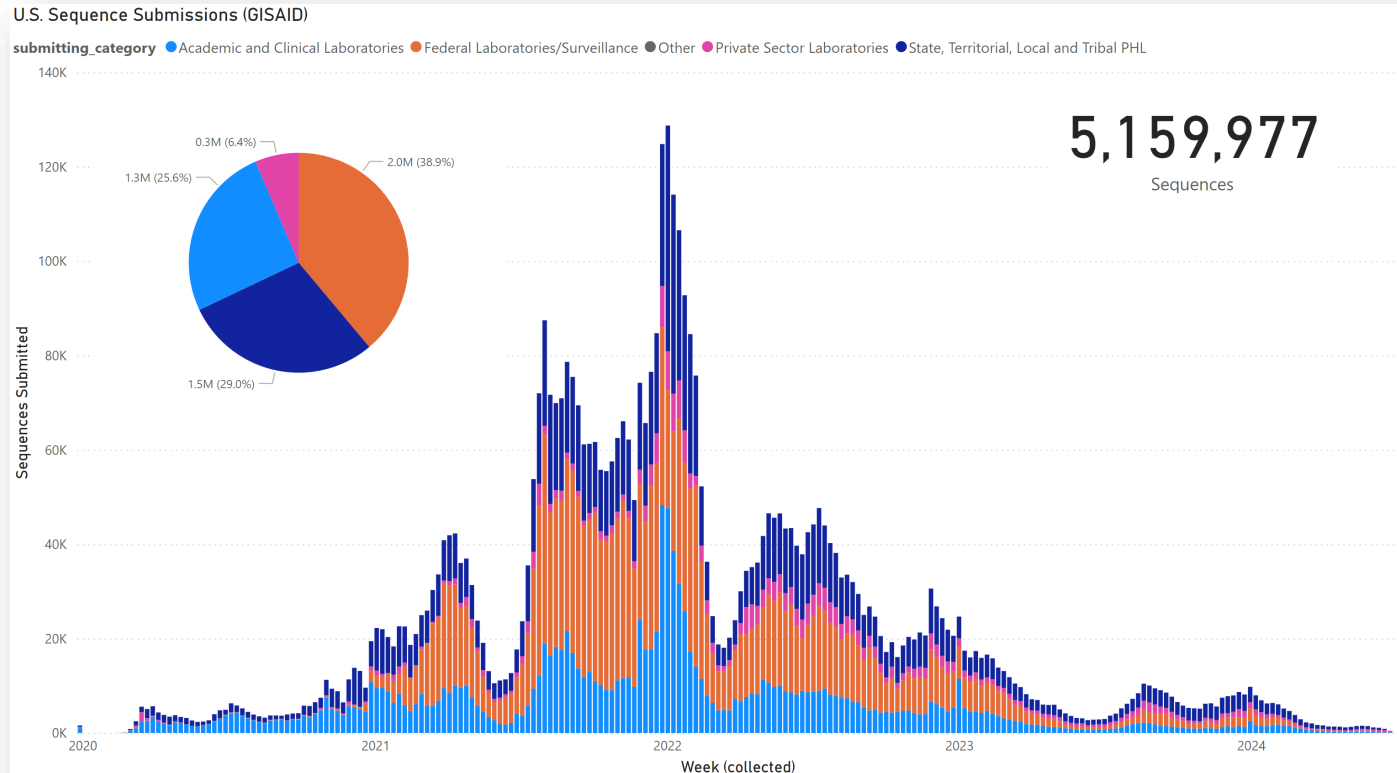


REFERENCE/
COORD. CENTER



DECENTRALIZED

Genomic sequencing of US SARS-CoV-2 (2020-2024).

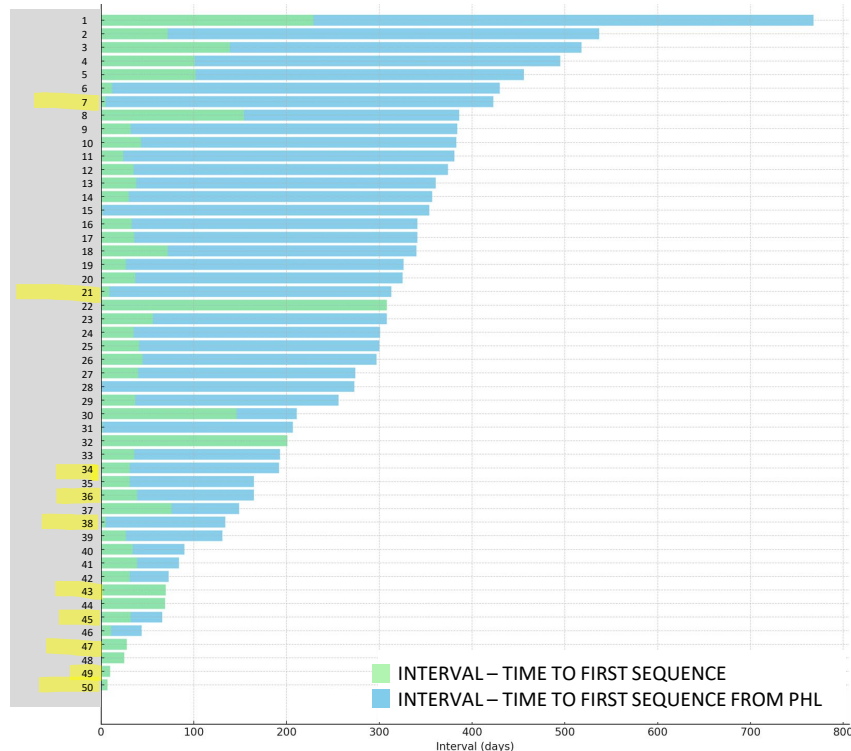


Data source: Line list metadata from US SARS-CoV-2 sequence submissions GISAID EpiCoV™, curated by the Nextstrain team, hosted at www.gisaid.org. Accessed: July 9, 2024.



Observations from SARS-CoV-2 genomic surveillance

Interval Between Earliest Reported COVID-19 Case in Each State and First Sequence Submission



1. Genomic sequencing and bioinformatics capacity that is built around surveillance for **routine pathogens does not automatically translate** into genomic readiness for a novel threat.
2. Strengthening **public health technical workforce** at the regional, state and local level can be a decisive factor.
3. **Partnerships with academia and the private sector** are a highly effective strategy to strengthen and accelerate pathogen genomics in many states.

* Note: many public health institutions did not fully prioritize sequencing until late 2020, when B.1.1.7/Alpha was identified and designated as a “variant of concern”.



The **SPHERES** consortium
brings together over **1,800** scientists
from more than **200** organizations
in a network of federal, state, private,
academic, and non-profit organizations
in the **largest public health
sequencing effort** in US history.



Centers for Disease Control and Prevention



Different objectives and priorities for sequence data



SPHERES

- Technical community of practice, connecting academia, public health and private sector.
- Share capacity and expertise, improve the utility, quality and timeliness of seq data.

NATIONAL PUBLIC HEALTH

- Builds national capacity: incorporates data from state, local, territorial and tribal public health, large commercial laboratories and other surveillance platforms to improve national and global surveillance, policy and decision-making.
- **PRIMARY GOALS:**
 1. National baseline data for variant surveillance
 2. Reference characterization and risk assessment
 3. Diagnostic, vaccine and therapeutic monitoring
 4. Guide resource prioritization and strategy
 5. Cluster identification and outbreak response
 6. Contribute to global picture of circulating viruses.

STATE AND LOCAL PUBLIC HEALTH

- Builds flexible, scalable and resilient genomic surveillance capacity that can be adapted to different scenarios and is responsive to local public health priorities.
- **PRIMARY GOALS:**
 1. Regional/local surveillance and monitoring for unusual variants, phenotypes or clinical outcomes
 2. Outbreak response, infection control, contact tracing
 3. State/local/regional strategy and coordination
 4. Reference characterization

Metagenomics and Pathogen Agnostic Applications

With few exceptions, the use of metagenomics in public health remains driven by specific use-cases: typically limited scale, pilot phase, or secondary to other technologies.

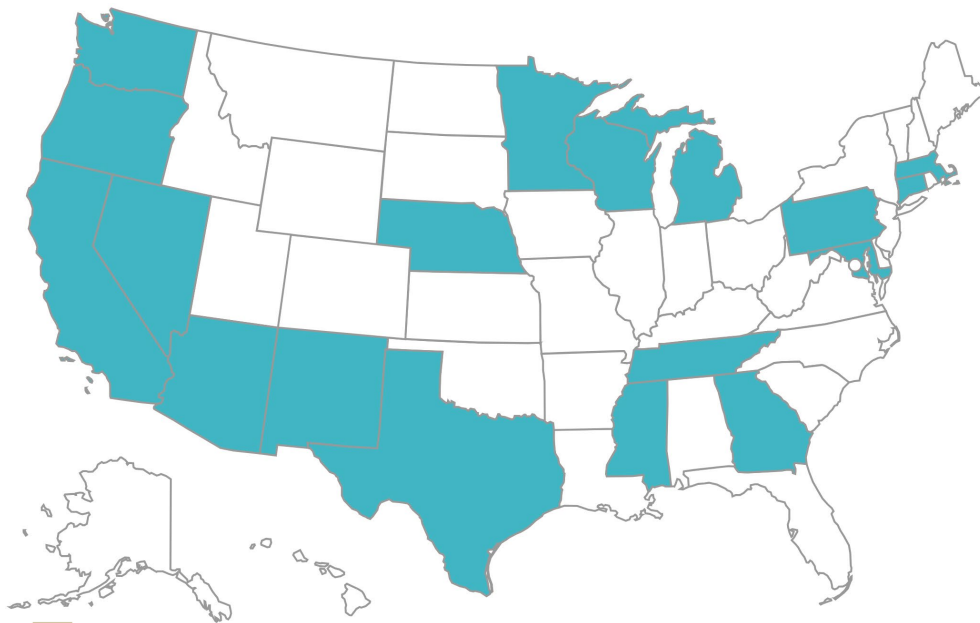
Data complexity, interpretability, timeliness, reproducibility and per-sample cost are all important considerations.

Trend: Pathogen genomics helping to design and validate lower-cost, highly multiplexed assay platforms.

Example Applications from US Public Health

- **Environmental sampling**
 - Wastewater (NWSS)
 - Aircraft wastewater (TGS)
 - Vectorborne monitoring: viral, bacterial
 - Healthcare built environment
 - Environmental mycology
- **Complex clinical samples**
 - Foodborne disease surveillance/CIDT
 - Rapid Genotypic Analysis of MTB
 - Antimicrobial resistance
 - Molecular Pathology
 - Biothreat / Biodefense
 - Unknown pathogens

Accelerating innovation in pathogen (meta)genomics



- 45 BAA Projects launched since 2020.
- Mostly one to three year applied public health research projects.
- Shifting and expanding range of topics and innovation/capability gaps:

— SARS-CoV-2 surveillance, epidemiologic investigations, and transmission dynamics

— SARS-CoV-2 variants and association with different clinical and viral characteristics.

— Bioinformatic software tools for phylogenetic, phylodynamic, genomic epidemiologic, and wastewater analysis.

— Pathogen agnostic and multipathogen detection: platforms, metagenomics, process automation, sample processing, AI/ML use cases, algorithms and data visualization.



Future workforce: CDC/APHL Bioinformatics Fellowship



APHL-CDC Infectious Disease Bioinformatics Fellowship

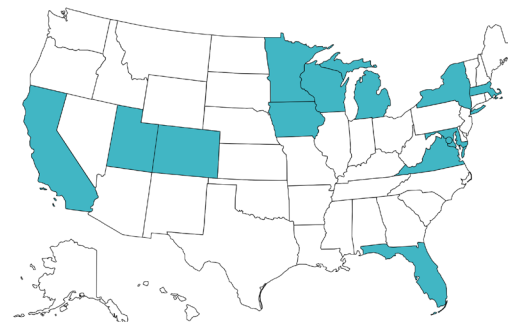
I am a Public Health Bioinformatician

Bioinformatics is revolutionizing the way the world tracks and detects infectious diseases. In public health, the use of Next Generation Sequencing (NGS) technology has reshaped outbreak investigations and pathogen surveillance. Bioinformaticians are crucial for this transition to the use of NGS in public health. They develop pipelines and help interpret the data, identifying and characterizing pathogens; playing a vital role in the public health engine that keeps us all healthy.

The US Centers for Disease Control and Prevention (CDC)'s [Advanced Molecular Detection](#) initiative is spearheading use of NGS technology in public health laboratories. The [Association of Public Health Laboratories \(APHL\)](#) and CDC Office of Advanced Molecular Detection (OAMD) are offering exciting fellowship opportunities for graduates of bioinformatics and related programs to apply their skillset and become part of the public health engine.

"The work is both satisfying and gratifying. I'm getting to use my knowledge and my position to make an impactful and meaningful difference in people's lives by preventing illness and fighting the spread of disease through modern surveillance and computational techniques.

— Logan Fink, 2018 Fellow
Colorado Department of Public Health and Environment



Post-masters/post-doctoral fellowship.

Since 2014: **10** cohorts, **66** fellows.

Placement in **12** state, **3** county/local PHLs.

To date: **71%** have stayed in public health, including **52%** from the first 5 cohorts.

- Retooling fellowship program to **improve recruitment** and incorporate stronger **genomic epidemiology** focus.
- New partnership with the **Dr. James A. Ferguson RISE** graduate fellowship program will help expand opportunities.

Pathogen Genomic Centers of Excellence



Washington PGCoe: Washington State Department of Health

- University of Washington
- Fred Hutchinson Cancer Center
- WA Animal Disease Diagnostic Laboratory
- Public Health – Seattle & King County

Minnesota PGCoe: Minnesota Department of Health

- University of Minnesota
- Mayo Clinic
- University of Pennsylvania College of Veterinary Medicine

New England PGCoe: Massachusetts DOH (Edu Lead)

- Broad Institute of MIT and Harvard
- Harvard Medical School/MassCPR
- Mass General Brigham healthcare system
- Boston University
- Yale University
- Fathom Information Design
- Theiagen Genomics

Virginia PGCoe: Division of Consolidated Laboratory Services

- Virginia Department of Health
- Virginia Commonwealth University
- University of Virginia

Georgia PGCoe: Georgia Department of Public Health

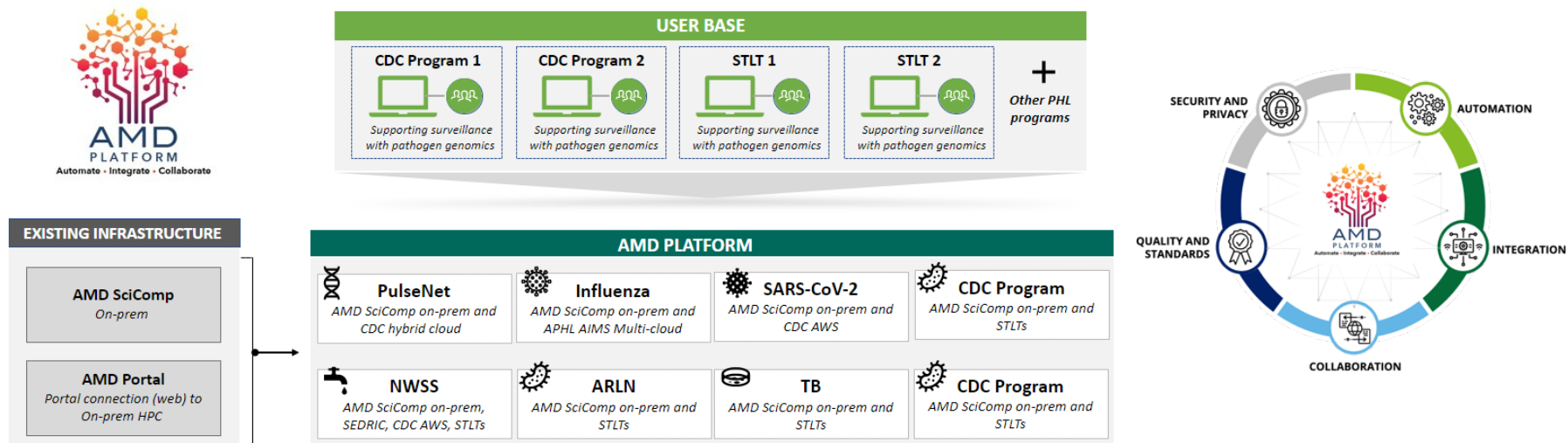
- Augusta University
- Emory University
- Georgia State University
- Georgia Tech Research Institute
- University of Georgia
- University of Texas Health Science Center

PGCoE Core Objectives:

- Applied innovation / R&D
- Establish collaborative networks
- Improve resilience and surge capacity
- Expanded technical workforce
- Engage with other networks and CoEs.



AMD Platform – Modular, cloud-based bioinformatics



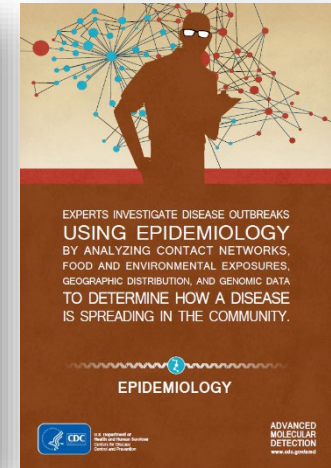
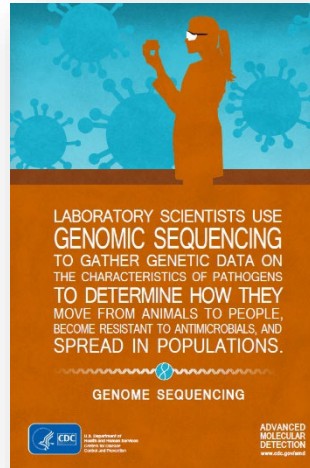
Thank you.

For more information, contact CDC/AMD.

Email: oamd@cdc.gov

Web: <https://www.cdc.gov/amd>

Follow us on X (Twitter) @CDC_AMD & @CDCgov



The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the U. S. Centers for Disease Control and Prevention.