Next Generation Sequencing in Public Health Laboratories

2014 Survey Results

MAY 2015

APHL ASSOCIATION OF PUBLIC HEALTH LABORATORIES™

This project was 100% funded with federal funds from a federal program of $215,972. This publication was supported by Cooperative Agreement # U60HM000803 funded by the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of CDC or the Department of Health and Human Services.
National Center for Immunization and Respiratory Diseases (IP)
Office of Surveillance, Epidemiology and Laboratory Services (OSELS)
National Center for HIV, Viral Hepatitis, STDs and TB Prevention (PS)
National Center for Zoonotic, Vector-borne, and Enteric Diseases (CK)
National Center for Environmental Health (NCEH)
Coordinating Office for Terrorism Preparedness and Emergency Response (CTPER)

Table of Contents

Background and Purpose

Next Generation Sequencing (NGS), also known as high throughput sequencing is a term used to capture a number of modern sequencing technologies that allow DNA and RNA to be sequenced much more quickly and cheaply than the previously used Sanger sequencing methodology. The increasing availability and affordability of NGS technologies is rapidly changing the practice of microbiology and has the potential to revolutionize the practice of public health. These technologies have the potential to more rapidly and accurately provide a high level of detailed information on outbreak causing pathogens, reducing reliance on more time-consuming and costly traditional diagnostic methods.

Several public health laboratories (PHLs) began the process of implementing NGS technologies in their laboratories a few years ago. However, resources, programs and funding made available through CDC's Advanced Molecular Detection (AMD) initiative have pushed implementation forward at a more rapid pace. APHL fielded the NGS in Public Health Laboratories survey in the fall of 2014 to establish a baseline NGS implementation in at the early stages of the AMD roll out to State and Local Public Health Departments. APHL will continue to field this survey over the next several years to capture the uptake and evolution of the technology and impact of the AMD initiative on the way public health is conducted in the US.

Method

The NGS in PHLs Survey was developed by APHL's Emerging Technologies Workgroup, which consists of representatives from CDC and public health laboratories. The final product was reviewed by APHL's Infectious Disease Committee.

The 12 question survey was launched on October 27, 2014 and officially closed on December 24, 2014. The questions were divided into four categories, including: NGS Capacity, Data Storage and Transmission, Data Analysis and Bioinformatics, and Training and Capacity Needs.
The survey was distributed electronically to 97 APHL member State and Local Public Health Laboratories via Qualtrics, a web-based survey platform.

Of the 51 state public health laboratories that received the survey, 50 (98%) responded. Of the 46 local public health laboratories that received the survey, 18 (39%) responded. The overall response rate for state and local public health laboratories was 70%.

None of the 18 responding local public health laboratories (LPHLS) reported having NGS instrumentation at the time of the survey. For this reason, LPHL responses are excluded from much of the data analysis on current PHL practices unless otherwise indicated.

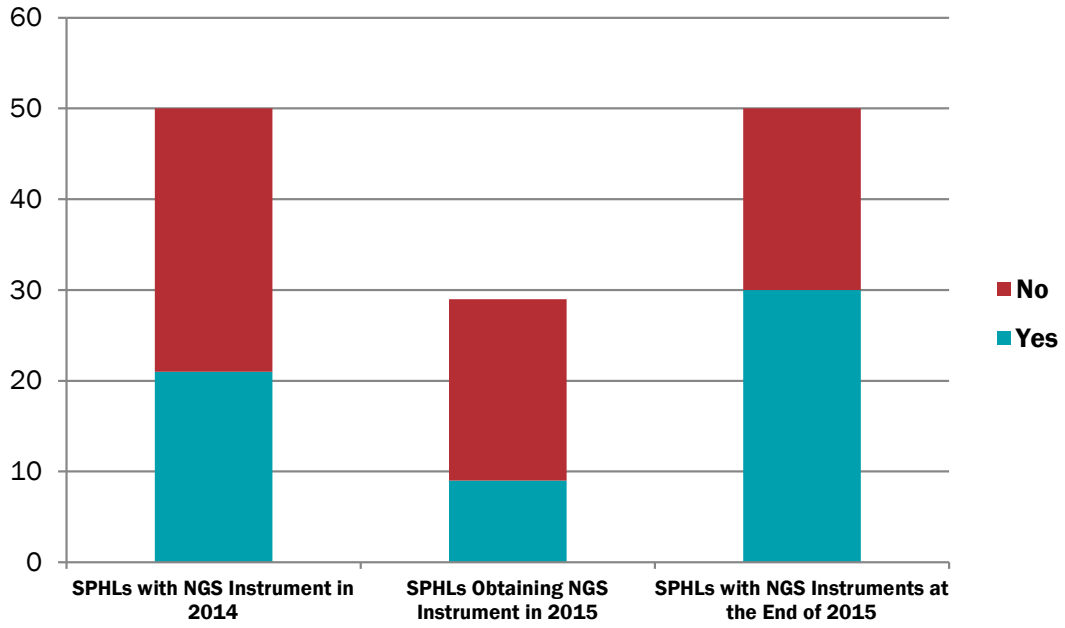
Results

NGS Instrumentation

Twenty one of 50 (42%) of responding state public health laboratories (SPHLs) currently have NGS instrumentation, while none of the responding local public health laboratories (LPHLS) reported having NGS instrumentation at the time of the survey.

Nine of 29 (31%) SPHLS and 3 of 18 (17%) responding LPHLs reported that they do not currently have NGS instrumentation but plan to purchase NGS instruments within the next 12 months. This will lead to 60% of SPHLs and 17% of LPHLs having sequencing instrumentation by the end of 2015. The majority of PHLs that have or are considering NGS instrumentation look at the Illumina MiSeq and the ThermoFisher IonTorrent PGM.

Figure 1: NGS Instrumentation Aquisition in SPHLs 2014-2015 (n=50)



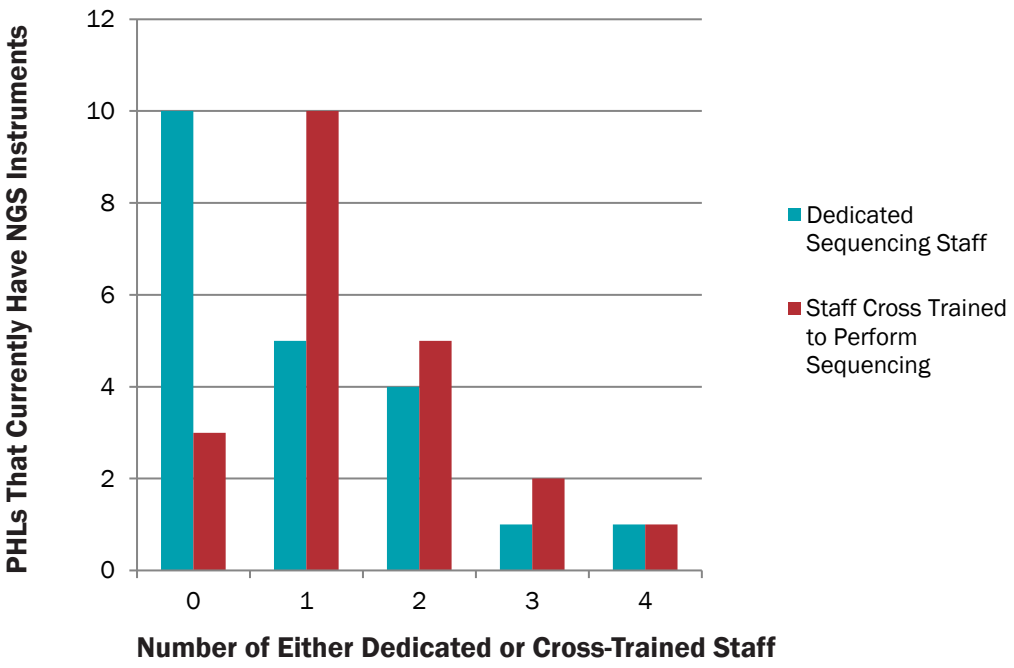
Reasons for Not Purchasing an NGS Instrument

The most common reasons that laboratories cited for waiting to purchase NGS instrumentation were: lack of available funding (19/20) followed by wanting applications to be more fully developed (10/20), the expense of instrumentation (9/20) and insufficient staff to add new methods (8/20). Several PHLs noted in open ended responses that they have access to sequencing technology through partnerships with local universities or other technology centers and are waiting until the needs become clearer before bringing the technology in-house.

Current NGS Practices

Approaches to Staffing

Figure 2: Staffing of NGS in PHLs: Numbers of Dedicated versus Cross Trained Laboratorians (n=21)

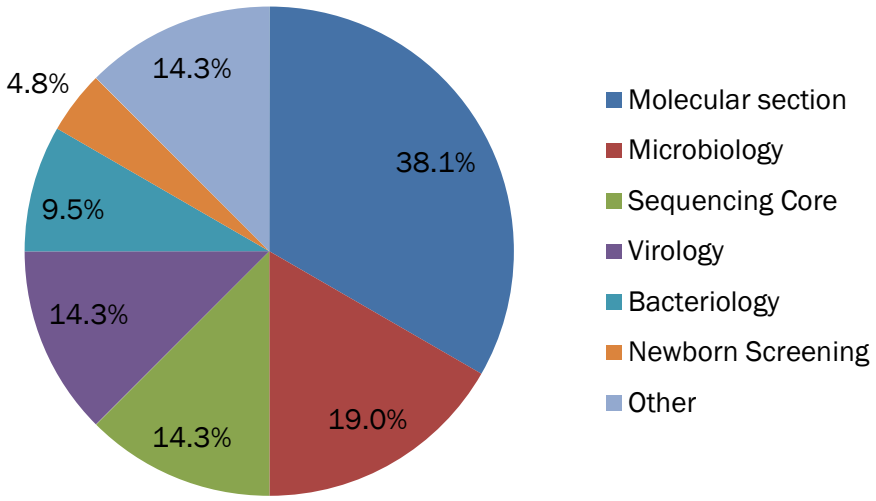


PHLs that currently have NGS Instrumentation have taken different approaches to staffing testing activities. Eight of 21 (38%) laboratories rely exclusively on staff members that have been cross trained from other laboratory sections, while only one PHL relies exclusively on dedicated staff members to perform sequencing. Ten of 21 (48%) PHLs use a combination of dedicated and cross trained staff. Two PHLs reported not having any staff members trained to perform sequencing at the time of the survey. Figure 2 depicts the numbers of dedicated versus cross trained sequencing staff in individual PHLs. For example there are 10 PHLs that do not have any dedicated sequencing staff and three PHLs that do not have any cross trained staff while five PHLs have one dedicated sequencing staff person and 10 PHLs have one staff member cross trained to perform sequencing.

Approaches to NGS Implementation

PHLs are using different approaches to the implementation of NGS in their laboratories. Figure 3 illustrates the diverse approaches to instrument placement within PHLs with most laboratories opting to place their instruments in their molecular (8 or 31.8%) or microbiology (4 or 19%) sections with several (3) laboratories opting to set up a sequencing core. The remaining instruments are in virology (3 or 14.3%), bacteriology (2 or 9.5%) or newborn sequencing (1 or 4.8%). “Other” responses include two laboratories that have opted to place instruments they own in sequencing facilities of a private sector laboratory outside of their own in order to take advantage of their sequencing and bioinformatics capabilities. A third laboratory’s instrument is currently housed in their state’s FERN laboratory.

Figure 3: Section of the Laboratory Where NGS Instrument is Placed (n=21)



NGS Applications

PHLs appear to struggle with identifying applications for NGS beyond characterizing and/or investigating foodborne outbreaks. All PHLs currently implementing or evaluating NGS applications are considering using the technology to characterize and/or identify clusters of foodborne pathogens. Other frequently cited applications include identification and/or characterization of Mycobacteria species (6), characterization of drug resistance for other pathogens including CRE, MRSA or gonorrhea (6) and New Born Screening (3). One or two PHLs cite using or exploring applications for additional pathogens, including respiratory pathogens, *Neisseria meningitidis*, Hepatitis C Virus, HIV, Influenza and *C. difficile*.

The lack of clarity surrounding public health applications of NGS beyond foodborne pathogens points to a need for CDC and APHL to clearly communicate the potential of other applications to both laboratorians and epidemiologists in state and local public health departments.

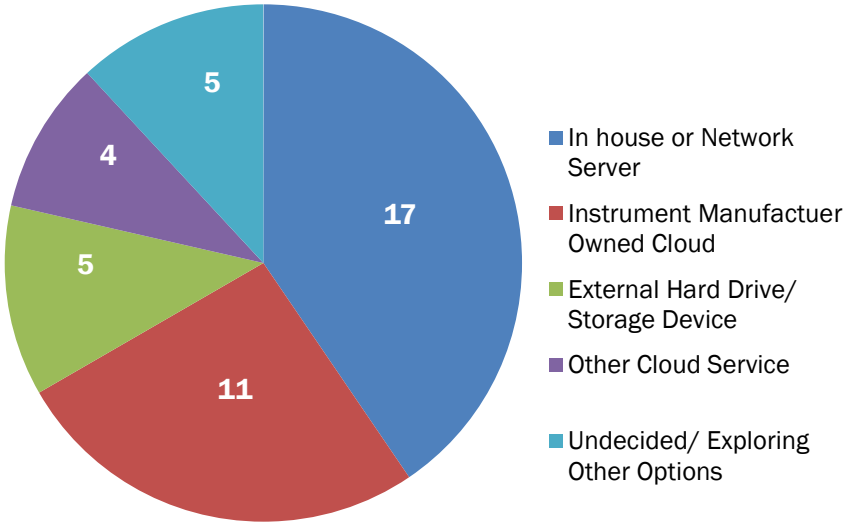
Data Storage and Transmission

One of the major decisions and costs associated with NGS implementation involves the storage the large amounts of data that are generated by the technology. Mechanisms for data transmission including how to transmit data and which metadata to share with network partners are also critical decision points. This survey did not specifically address mechanisms for data sharing although many PHLs are currently relying on cloud based mechanisms or secure FTP sites.

Mechanisms for NGS Data Storage

Current PHL NGS users are exploring several options for data storage with many using a combination of local and cloud based storage (Figure 4).

Figure 4: NGS Data Storage Mechanism (n=21)



Sharing Metadata

Sharing sequence data with network partners and the inclusion of metadata with a sequence is critical to ensuring most effective use of the sequence information to meet epidemiologic and surveillance goals. However, there are many concerns around data sharing agreements and privacy considerations. Survey respondents were asked which data fields they would be willing to share with network partners (Table 1). The majority (80-90%) of responding laboratories were willing or able to share very generic data however when asked about sharing more specific patient (e.g. age) or location (e.g. zip code) data fewer labs responded affirmatively. Two PHLs stated that they would not be willing to share any of the data fields asked about. Three PHLs stated that they would not be willing to share any data with network partners with one laboratory citing limitations with state data sharing allowances and two laboratories citing concerns about the lack of secure data transmission mechanisms. These laboratories did not respond to the question represented in Table 1.

Table 1: Metadata fields PHLs are Willing or Able to Share (n=20)

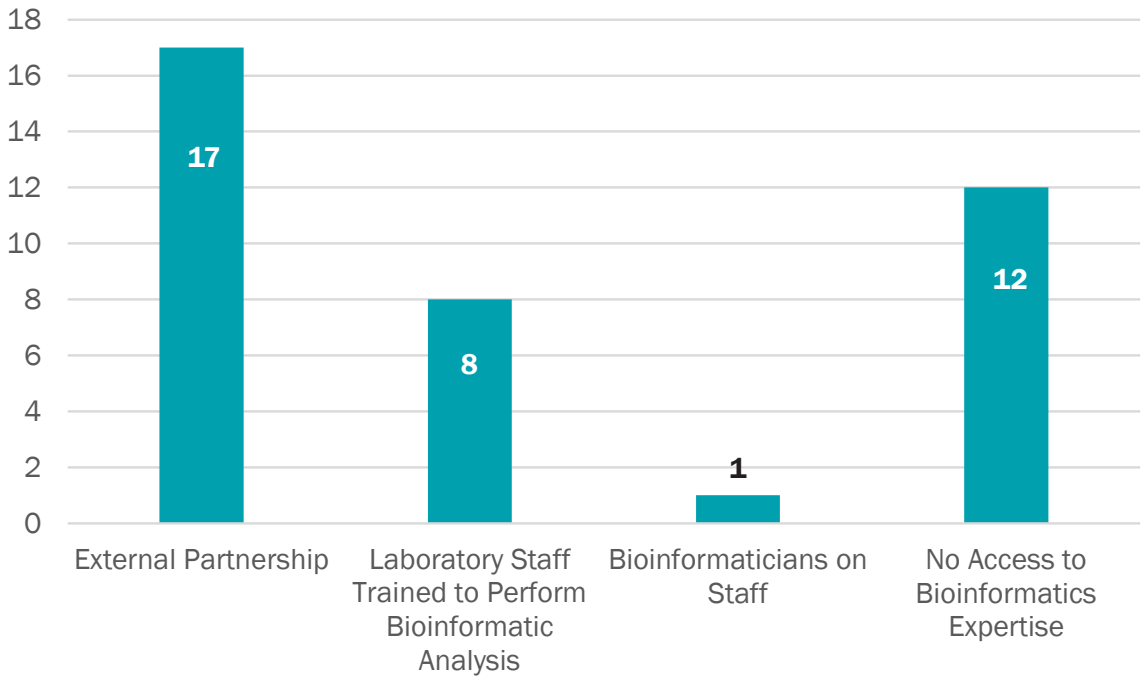
Answer	Response	%
ID for laboratory submitting DNA sequence (e.g., CDC)	18	90.0%
Organism genus and species (e.g., *Listeria monocytogenes*)	18	90.0%
Environmental source (e.g., water, soil, air)	18	90.0%
Collection year	18	90.0%
Commodity type (e.g., meat, produce, seafood)	18	90.0%
Organism serotype (if applicable and available)	17	85.0%
Country of origin (USA)	17	85.0%
Organism source (e.g., blood, cerebral spinal fluid, stool)	17	85.0%
Unique sample ID (WGS_ID)	16	80.0%
Geographic location (state)	16	80.0%
Site of isolation (e.g., blood, cerebral spinal fluid, stool)	14	70.0%

Geographic location (HHS region of patient residence)	13	65.0%
Age category	12	60.0%
Geographic location (zip code)	6	30.0%
None of the above	2	10.0%

Data Analysis and Bioinformatics

Access to bioinformatics expertise is a significant gap among laboratories that have implemented or are near implementation of NGS technologies with 12/33 (36%) reporting no access to bioinformatics expertise (Figure 5). Only 1 responding PHL reports having one bioinformatician on staff. Eight responding PHLs (24%) have cross-trained laboratory staff to perform limited bioinformatics analysis, however across APHL member laboratories there are only 12 laboratory staff trained to perform bioinformatics analysis. Seventeen (52%) of 33 responding PHLs report having access to bioinformatics capacity through an external partnership. CDC (7), FDA (3), NCBI (6) and university partnerships (8) were the most frequent partners cited.

Figure 5: Sources of Bioinformatics Expertise in PHLs (n=33)

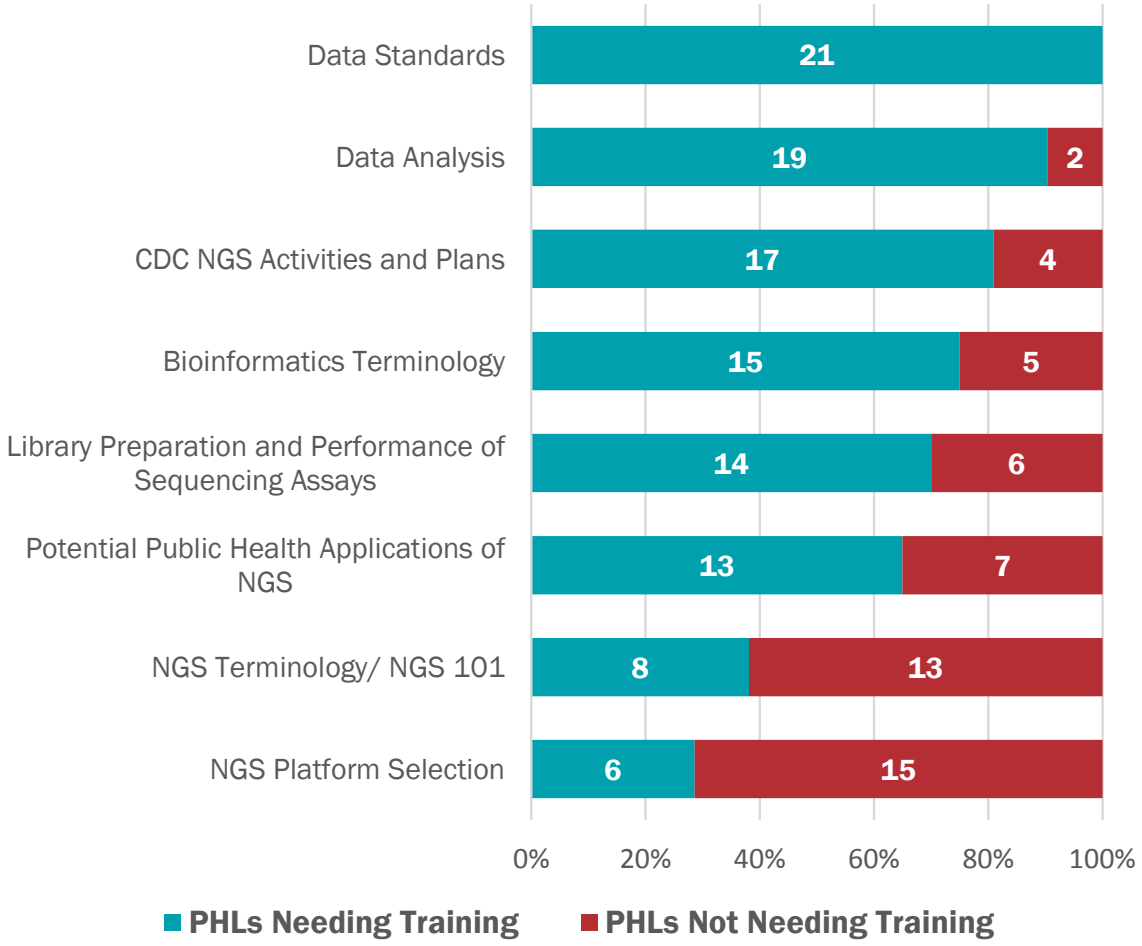


Training and Capacity Needs

Training Needs

Training was cited as a significant need for all responding PHLs regardless of whether or not they had NGS instrumentation. Only 17 of all 68 responding PHLs (25%) have staff members that have received any NGS training. However, the percentage of staff members receiving training shifts when the presence of NGS instrumentation is considered. Sixty-one percent (13/21) of laboratories that have NGS instrumentation have had staff participate in NGS training while only 8% (4/47)) of responding PHLs without NGS instrumentation have attended training. Across all PHLs 31 laboratorians had received NGS training at the time of the survey, with an average of two staff members per PHL training on the methodology. Figure 6 depicts training needs cited by laboratories with NGS instrumentation.

Figure 6: Training Needs in PHLs with Existing Instrumentation (n-21)



Capacity Needs

PHLs were asked to prioritize how they would spend available NGS dollars. The results differed widely between those PHLs that currently own NGS instrumentation and those that do not. Those that currently have instruments would prioritize funding between hiring and training additional personal and purchasing reagents and supplies with results being fairly evenly split. The majority (72%) of PHLs that do not have NGS instruments would place purchasing an instrument as their first or second priority. Table 2 shows the breakdown of how survey respondents prioritized their NGS capacity needs.

Table 2: Priority NGS Needs for PHLs Stratified by Instrumentation Status

	PHLs that have NGS Instrument (n=21)					
	Prioritization Ranking					
	1	2	3	4	5	6
Additional Personnel	7	3	0	3	5	3
Training	6	3	2	3	3	4
Reagents/ Supplies	4	4	3	6	0	4
Equipment/ service contracts	3	2	7	2	4	3
IT infrastructure	1	6	5	1	6	2
Software	0	3	4	6	3	5
	PHLs that do not have NGS Instrument (n=47)					
	Prioritization Ranking					
	1	2	3	4	5	6
Equipment/ service contracts	25	9	2	4	6	1
Additional Personnel	8	4	4	4	7	20
Training	7	10	20	5	4	1
IT infrastructure	3	9	3	14	8	10
Reagents/ Supplies	2	7	9	12	7	10
Software	2	8	9	8	15	5

Conclusion

PHLs are embracing the adoption of NGS technologies with many moving forward with the implementation of surveillance activities that incorporate use of this technology. Over the next few years public health will see increased use of these technologies across the country for many different types of public health investigations. However, in order to fully realize the potential of NGS, several needs and capacity gaps need to be addressed.

- The public health utility of NGS for foodborne pathogens has been well articulated and is well understood. However, more efforts are needed to discuss, describe and communicate applications beyond foodborne pathogens and the state and local role in implementing and executing use of those applications. Engagement of epidemiologists as well as public health laboratorians in these processes is imperative.
- The public health community will also need to explore and navigate various mechanisms for transmitting NGS data as well as establishing data sharing agreements between all users of NGS data. This will include identifying funding and a mechanism to identify an IT infrastructure capable of supporting the storage and exchange of large amounts of data.
- Training of PH laboratorians in performance of NGS testing procedures, bioinformatics and NGS data standards was highlighted as a critical need.

APHL will continue to work with CDC and member laboratories to ensure that these and other needs are addressed. Progress on NGS implementation in PHLs and progress on meeting the needs identified in this survey, will be monitored through fielding of additional surveys in future years.

Association of Public Health Laboratories

The Association of Public Health Laboratories (APHL) is a national nonprofit dedicated to working with members to strengthen laboratories with a public health mandate. By promoting effective programs and public policy, APHL strives to provide public health laboratories with the resources and infrastructure needed to protect the health of US residents and to prevent and control disease globally.

8515 Georgia Avenue, Suite 700
Silver Spring, MD 20910
Phone: 240.485.2745
Fax: 240.485.2700
Web: www.aphl.org