

UC Davis

UC Davis Previously Published Works

Title

Next-generation sequencing capacity and capabilities within the National Animal Health Laboratory Network

Permalink

<https://escholarship.org/uc/item/0cz8c8wf>

Journal

Journal of Veterinary Diagnostic Investigation, 33(2)

ISSN

1040-6387

Authors

Harris, Beth
Hicks, Jessica
Prarat, Melanie
[et al.](#)

Publication Date

2021-03-01

DOI

10.1177/1040638720937015

Peer reviewed

Next-generation sequencing capacity and capabilities within the National Animal Health Laboratory Network

Journal of Veterinary Diagnostic Investigation
2021, Vol. 33(2) 248–252
© 2020 The Author(s)
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1040638720937015
jvdi.sagepub.com

Beth Harris,¹  Jessica Hicks, Melanie Prarat,
Susan Sanchez, Beate Crossley 

Abstract. With the cost of next-generation sequencing (NGS) decreasing, this technology is rapidly being integrated into the workflows of veterinary clinical and diagnostic laboratories nationwide. The mission of the U.S. Department of Agriculture–National Animal Health Laboratory Network (NAHLN) is in part to evaluate new technologies and develop standardized processes for deploying these technologies to network laboratories for improving detection and response to emerging and foreign animal diseases. Thus, in 2018, the NAHLN identified the integration of NGS into the network as a top priority. In order to assess the current state of preparedness across NAHLN laboratories and to identify which have the capability for performing NGS, a questionnaire was developed by the NAHLN Methods Technical Working Group and submitted to all NAHLN laboratories in December 2018. Thirty of 59 laboratories completed the questionnaire, of which 18 (60%) reported having some sequencing capability. Multiple sequencing platforms and reagents were identified, and limited standardized quality control parameters were reported. Our results confirm that NGS capacity is available within the NAHLN, but several gaps remain. Gaps include not having sufficient personnel trained in bioinformatics and data interpretation, lack of standardized methods and equipment, and maintenance of sufficient computing capacity to meet the growing demand for this technology.

Key words: animal diseases; bioinformatics; questionnaire; whole-genome sequencing.

Next-generation sequencing (NGS), or *high-throughput sequencing*, has revolutionized detection, characterization, and analysis of bacterial and viral pathogens. With the cost of NGS decreasing, this technology is rapidly being integrated into the workflows of veterinary clinical and diagnostic laboratories nationwide. NGS has many advantages over existing laboratory approaches including throughput, quality, flexibility, and scalability.³ To date, the era of genome sequencing is divided into 3 generations.

First-generation sequencing techniques were developed in 1977 by Frederic Sanger, Allan Maxam, and Walter Gilbert. Sanger's dideoxy method and the Maxam–Gilbert chain termination method were automated and became the dominant sequencing technology for >20 y. Sanger sequencing was used to sequence the first microbial and human genomes in the 1990s and early 2000s, costing \$50,000 and \$2.7 billion, respectively.^{6,13} First-generation sequencing generates read lengths up to 800 bp in ~3 h, but is labor intensive and costs ~\$2,400 per megabase.¹⁴

Second-generation technologies, including Roche 454 pyrosequencing (Roche Diagnostics), Illumina sequencing by synthesis (SBS; Illumina), and Applied Biosystems Sequencing by Oligo Ligation Detection (SOLiD; Life Technologies), utilize high-throughput, massively parallel

sequencing that generates millions of short reads (50–500 bp) and costs as little as \$0.07 per megabase, but often require >24 h run time.¹²

Third-generation sequencing technologies include Ion Torrent SBS (Thermo Fisher Scientific), PacBio real-time synthesis (Pacific Biosciences), and Oxford Nanopore ion-exchange biopores (Oxford Nanopore Technologies) that generate longer reads (>1,000 bp) and cost \$0.13–\$1.00 per megabase, with sequence results in as little as 30 min.¹⁴

The term *next-generation sequencing* (NGS) refers to massively parallel or deep sequencing using the above technologies, whereas *whole-genome sequencing* (WGS) is

USDA-APHIS-VS-DB National Animal Health Laboratory Network (Harris), USDA-APHIS-VS-DB National Veterinary Services Laboratories (Hicks), Ames, IA; Ohio Department of Agriculture Animal Disease Diagnostic Laboratory, Reynoldsburg, OH (Prarat); Athens Veterinary Diagnostic Laboratory, Department of Infectious Diseases, College of Veterinary Medicine, The University of Georgia, Athens, GA (Sanchez); California Animal Health and Food Safety Laboratory, University of California–Davis, Davis, CA (Crossley).

¹Corresponding author: Beth Harris, USDA-APHIS-VS-DB National Animal Health Laboratory Network, 1920 Dayton Avenue, Ames, IA 50010. beth.n.harris@usda.gov

Table 1. Next-generation sequencing (NGS) capacity within the National Animal Health Laboratory Network as of December 2018.

| Sequencing resource | No. of affirmative responses |
|--|------------------------------|
| In-house or core facility sequencing capability* | 18 of 30 (60%) |
| Access to sequencing equipment† | 24/7: 12 |
| | Intermittent: 10 |
| Sequencing applications† | Detection/client request: 7 |
| | Research: 11 |
| Personnel dedicated to NGS benchwork | 10 of 30 (33%) |
| Personnel trained in NGS analysis | 18 of 30 (60%) |
| Satisfaction with current computer processing capacity | 16 of 30 (53%) |
| Data storage methods used† | Cloud: 8 |
| | Local Server: 14 |
| | External hard drive: 13 |
| Overall satisfaction with analysis tools | 19 of 30 (63.3%) |

* This question did not distinguish between in-house or core facility sequencing capabilities.

† Laboratories could select as many answers as applied.

a specific application of NGS used to determine the complete nucleic acid sequence of an organism's genome at a single time.

The U.S. Department of Agriculture's National Animal Health Laboratory Network (NAHLN) was established in 2002, and is a nationally coordinated network and partnership of federal, state, and university-associated animal health laboratories. NAHLN's mission is to provide animal health laboratory testing to detect emerging and foreign animal diseases (FADs) such as foot-and-mouth disease, African swine fever, and influenza in avian and swine species, among other NAHLN scope diseases. Importantly, the NAHLN also evaluates new technologies and develops standardized processes for deploying these technologies to network laboratories to improve FAD and emerging disease detection and response. The most recent NAHLN Strategic Plan identified the need to incorporate NGS into the NAHLN to support disease identification and surveillance activities.

Laboratories within the NAHLN are classified as levels 1–3 (L1–3), affiliate, or specialty, based on criteria established in the 2012 NAHLN Concept Paper.¹² Level 1–3 laboratories are publicly funded, state and university-associated animal health diagnostic laboratories, whereas affiliate laboratories are either federal or publicly funded laboratories whose mission falls outside the scope of animal agriculture. Specialty laboratories are private laboratories that have a specific testing capability otherwise unavailable within the NAHLN, needed to support APHIS surveillance and disease eradication efforts. Branch laboratories are defined in the NAHLN as multiple laboratories within a state or territory that share one or more of the following: administrative oversight, a quality system, or NAHLN funding mechanism. In December 2018, 59 laboratories and 2 National Veterinary Services Laboratories (NVSL; reference laboratories) constituted the NAHLN: 15 L1 laboratories; 6 L1 branch laboratories; 28 L2 laboratories; 2 L2 branch laboratories; 7 L3 laboratories; and 1 affiliate laboratory.

In 2018, a questionnaire was developed by a 7-member subcommittee of the NAHLN Methods Technical Working Group, with the goals of 1) determining the current state of NGS in the NAHLN laboratory community, 2) identifying the current use and capabilities of NAHLN laboratories performing NGS, and 3) identifying potential equipment platforms, bioinformatics tools, and pipelines that would allow standardization of NGS as a detection or disease response tool in the NAHLN. Thirty-two questions were formulated and divided into 3 general topic areas: reagents and equipment, analysis tools, and cost. Within the reagents and equipment topic, 14 questions targeted NGS technical procedures and addressed equipment, reagents, personnel, sample type(s), and current applications. The analysis tools section included 16 questions on quality control practices, software, data processing, analysis, storage, data sharing, and practices. The final 2 questions addressed fees and specific drivers behind NGS use in the laboratory. The estimated time for completion of the questionnaire was 15 min. Laboratories were given 2 wk from time of receipt to complete and return the questionnaire.

Responses were received from 30 NAHLN laboratories: 13 L1 laboratories, 15 L2 laboratories, and 2 L3 laboratories. No responses were received from any NAHLN branch or affiliate laboratories.

Overall, 18 responding NAHLN laboratories (60%) reported having sequencing capabilities, defined as either in-house or via a core facility (Table 1). This is similar to the 2016 survey of U.S. veterinary diagnostic laboratories,² in which 39% of surveyed laboratories reported the ability to conduct sequencing. Similar to the 2016 study, approximately one-third of the laboratories in the NAHLN still employ Sanger sequencing at some level.

Laboratories were asked for information about which NGS sequencing platforms they used. Thirteen laboratories reported access to and use of multiple sequencing platforms. Overall, Illumina and Oxford Nanopore Technologies were

Table 2. Types of sequencing platforms used by NAHLN laboratories.

| | |
|--|----------|
| No. of laboratories not using any sequencer platform | 9 of 30 |
| No. of laboratories using only one platform | 8 of 30 |
| Applied Biosystems (Sanger; ABI 3700, ABI 3730) | 4 |
| Oxford Nanopore (MinION) | 2 |
| Illumina (MiSeq, iSeq) | 2 |
| No. of laboratories using multiple platforms* | 13 of 30 |
| Applied Biosystems (Sanger; ABI 3700, ABI 3730) | 7 |
| Illumina (HiSeq, MiSeq, iSeq, miniSeq, NextSeq) | 13 |
| Thermo Fisher (Ion Torrent) | 2 |
| Oxford Nanopore (MinION, PromethION) | 10 |
| PacBio | 1 |

* Laboratories were asked to list all platforms that applied.

the 2 NGS manufacturers cited most frequently by all facilities (Table 2), followed by Thermo Fisher. Across the Illumina sequencing platforms, 2 laboratories reported using the HiSeq, 15 had a MiSeq, 6 used an iSeq, and 1 used a NextSeq. For the Oxford Nanopore sequencers, 2 laboratories reported using a MinION, and 2 had access to a PromethION. Two laboratories reported using the Thermo Fisher Ion Torrent sequencing platform, and 1 laboratory had access to a PacBio sequencing platform. These results align with recent surveys of both human public health laboratories¹⁰ and veterinary diagnostic laboratories,² in which Illumina platforms were the most frequently employed.

This questionnaire also asked which commercial reagents were being used by laboratories given that standardization of protocols, including the use of specific reagents or kits, is a cornerstone of the NAHLN. The most common nucleic acid preparations employed were silica membrane spin-column technologies (10 of 15 laboratories responding), with significantly fewer laboratories using magnetic bead-particle technology (5 of 15 laboratories). Library preparation was most often performed using transposon-based fragmentation (14 of 22 responding laboratories), followed by enzymatic shearing (5 of 22 laboratories) and ligation-based fragmentation (3 of 22 laboratories). NAHLN laboratories were also queried about the manufacturers of sequencing kits that they used. Illumina kits were used most commonly (13 of 16 responding laboratories), followed by Oxford Nanopore (3 of 16 laboratories). This is unsurprising, given that the use of reagents from the same manufacturer as the sequencing platform ensures that workflows are optimized.

Information was requested on the volume of NGS testing performed by laboratories, and how NGS was being applied within each institution. On average, laboratories tested between 6 (minimum) and 55 (maximum) samples per run, with a range of 1–192. The number of laboratories performing metagenomics or WGS of individual bacterial or viral isolates was also requested. Eleven of 30 laboratories reported doing either targeted or shotgun metagenomics, and 12 indicated they were using NGS for WGS of bacterial and/or viral isolates.

Additional questions were asked to gain further understanding of applications. Overall, 16 laboratories use NGS technology for pathogen characterization, 11 for pathogen detection, 9 for virus discovery, 7 for outbreak situations or molecular epidemiology, and 4 for other purposes. Comparatively, bacterial identification, antimicrobial resistance profiling, and molecular epidemiology are frequent NGS applications for both human and veterinary clinical laboratories.^{1,2,4,9}

Bioinformatics expertise can be a bottleneck for incorporating NGS into the clinical laboratory setting.^{1,10} Thus, our questionnaire queried personnel resources available to laboratories for conducting NGS. Most laboratories reported having at least some personnel trained in NGS techniques (Table 1). Bioinformatics expertise across all NAHLN laboratories was self-assessed at the beginner or intermediate level for 17 of the 22 laboratories responding to this question, suggesting that additional bioinformatics training may be needed to ensure standardized deployment of sequencing within the NAHLN. Although fewer laboratories reported having personnel dedicated for NGS bench work than bioinformatics or data analysis, our results may not reflect laboratory staff who are trained for routine molecular testing that also perform NGS bench work.

A founding principle of the NAHLN is for laboratories to operate within a quality management system, to provide confidence in test performance. Moreover, laboratories working under a formal quality management system such as ISO 17025 require extensive test validation to demonstrate that employed methods provide high-quality results.¹¹ Thus, our questionnaire also queried quality control parameters employed by laboratories performing NGS work. Although no clear consensus was identified regarding the use of standardized parameters to confirm NGS results, some themes emerged that may be explored for use within the NAHLN. These include evaluating raw data FastQ scores, assessing internal controls, and using third-party bioinformatics tools such as FastQC¹⁵ or QUILT.⁷ Additionally, the use of Clinical Laboratory Improvements Act guidelines may be considered given that these practices are being utilized by some public health laboratories.⁸ The use of proficiency testing to help standardize NGS results was also cited by several reports as a critical aspect of overall quality management of NGS data.^{4,11,16}

Standardized bioinformatics data analysis tools, or “pipelines,” are needed for normalizing NGS methods across NAHLN laboratories. Implementation of such pipelines is often impeded by the level of experience of personnel and types of tools available to them.^{5,10} The questionnaire responses showed differences in skill levels and tools used in NAHLN laboratories, with most using at least one bioinformatics analysis tool. Most laboratories reported having either beginner or intermediate level expertise; advanced expertise was reported by only 5 laboratories. Eight laboratories did not respond to this question. The tools used varied by reported

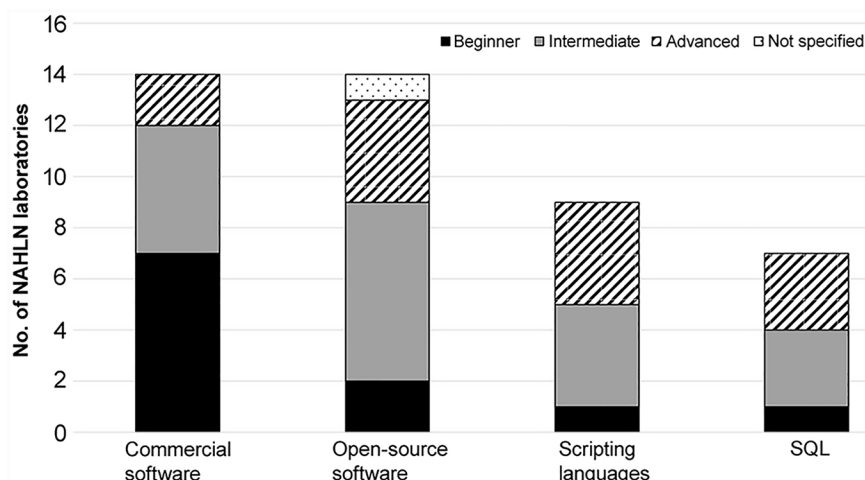


Figure 1. Next-generation sequencing bioinformatics data analysis tools used by NAHLN laboratories. The number of laboratories using each type of bioinformatics tool (commercial software, open-source software, scripting languages, or SQL) are further divided into bioinformatics skill levels, as reported by the laboratories.

skill level (Fig. 1). Laboratories at a beginner level were more likely to use commercial software tools; 7 laboratories using open-source software were at an intermediate level of expertise (Fig. 1). Comparatively, only 9 laboratories reported knowledge of scripting languages necessary for developing pipelines. Of these, 4 of 9 had intermediate, and another 4 of 9 had advanced, experience. Knowledge of additional skills such as Structured Query Language (SQL) were reported by 7 laboratories. Although these results are not unexpected, they present challenges for developing and integrating standardized bioinformatics analysis tools within the NAHLN.

Computational resources can also be a limiting factor for incorporating NGS into clinical laboratories.¹⁰ Overall, 19 of 30 responding NAHLN laboratories indicated that they were satisfied with both their computer processing capacity and data analysis capabilities (Table 1). However, most of these laboratories also reported doing low levels of NGS sequencing, suggesting that, as laboratories find more applications for NGS, current computational resources may become limited.

The information gathered from our questionnaire indicates that many NAHLN laboratories have implemented NGS at some level; however, in order to fulfill the NAHLN's mission of detecting foreign and emerging animal diseases, identifying additional resources for implementing this technology across all laboratories in the network will be needed. The gaps identified through our questionnaire include having sufficient personnel trained in bioinformatics and data interpretation, lack of standardized quality control methods and bioinformatics pipelines, and maintaining sufficient computing capacity to meet the growing demand for this technology. Nonetheless, the opportunities for improved testing that NGS affords to the veterinary community will help drive this forward and continue to make this a priority for implementation.

Acknowledgments

We thank all members of the NAHLN Methods Technical Working Group subcommittee on NGS for their input and critical review of the questionnaire. The findings and conclusions in this publication have not been formally disseminated by the USDA and should not be construed to represent any agency determination or policy.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. Mention of tradenames or commercial products is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Beth Harris  <https://orcid.org/0000-0002-2355-8707>

Beate Crossley  <https://orcid.org/0000-0003-2932-7229>

References

1. Balloux F, et al. From theory to practice: translating whole-genome sequencing (WGS) into the clinic. *Trends Microbiol* 2018;26:1035–1048.
2. Dargatz DA, et al. A survey of methods used for antimicrobial susceptibility testing in veterinary diagnostic laboratories in the United States. *J Vet Diagn Invest* 2017;29:669–675.
3. Didelot X, et al. Transforming clinical microbiology with bacterial genome sequencing. *Nat Rev Genet* 2012;13:601–612.
4. Deurenberg RH, et al. Application of next generation sequencing in clinical microbiology and infection prevention. *J Biotechnol* 2017;243:16–24.

5. European Centre for Disease Prevention and Control. Monitoring the use of whole-genome sequencing in infectious disease surveillance in Europe. ECDC, 2018. Accessed 2020 June 3: <https://www.ecdc.europa.eu/en/publications-data/monitoring-use-whole-genome-sequencing-infectious-disease-surveillance-europe#no-link>
6. Fleischmann RD, et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 1995;269:496–512.
7. Gurevich A, et al. QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 2013;29:1072–1075.
8. Kozyreva VK, et al. Validation and implementation of Clinical Laboratory Improvements Act-compliant whole-genome sequencing in the public health microbiology laboratory. *J Clin Microbiol* 2017;55:2502–2520.
9. Quainoo S, et al. Whole-genome sequencing of bacterial pathogens: the future of nosocomial outbreak analysis. *Clin Microbiol Rev* 2017;30:1015–1063.
10. Revez J, et al. Survey on the use of whole-genome sequencing for infectious diseases surveillance: rapid expansion of European national capacities, 2015–2016. *Front Public Health* 2017;5:347.
11. Rossen JWA, et al. Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology. *Clin Microbiol Infect* 2018;24:355–360.
12. U.S. Department of Agriculture (USDA), Animal and Plant Health Inspection Service. National Animal Health Laboratory Network reorganization concept paper. Docket ID: APHIS-2012-0105-0001. USDA, 2013. Accessed 2020 June 3: <https://www.regulations.gov/docket?D=APHIS-2012-0105>
13. Venter JC, et al. The sequence of the human genome. *Science* 2001;291:1304–1351.
14. Verma M, et al. Genome sequencing. *Methods Mol Biol* 2017;1525:3–33.
15. Wingett SW, Andrews S. FastQ Screen: a tool for multi-genome mapping and quality control. *F1000Res* 2018;7:1338.
16. Yohe S, Thyagarajan B. Review of clinical next-generation sequencing. *Arch Pathol Lab Med* 2017;141:1544–1557.