2018

# Next-generation sequencing bioinformatics: Guidance between the sequencing and sign out

Jeffrey Szymanski
*Washington University School of Medicine in St. Louis*
Eric Duncavage
*Washington University School of Medicine in St. Louis*
John Pfeifer
*Washington University School of Medicine in St. Louis*

# Next-generation Sequencing Bioinformatics: Guidance between the Sequencing and Sign out

Received: 23 March 2018 Accepted: 10 May 2018 Published: 25 June 2018


## Summary

Clinical next-generation sequencing (NGS) has evolved from a novel and promising test modality into a necessary and integral part of diagnosis and prediction of response to therapy. The technology to implement NGS, while complex, is now within the reach of most clinical laboratories. As NGS technology has advanced, a variety of professional and regulatory bodies have developed guidance for clinical laboratories seeking to develop NGS assays [Figure 1].[5,6]


The article by Roy *et al*. in the January issue of JMD is the most recent publication in a series of recommendations for clinical NGS[1] led by the Association for Molecular Pathology. In earlier JMD papers, Li *et al*.[2] and Richards *et al*.[3] established the standards for reporting of somatic and germline sequence variants, respectively. Their work focused on the best practices for selection and use of reference databases and on the evidence-based, clinically focused variant classification. Li *et al*. based their recommendations, in part, on surveys of variant classification from Association for Molecular Pathology (AMP) member laboratories. These surveys highlighted the broad differences not only in how variants were classified but also in the underlying bioinformatics pipelines each laboratory used to identify variants.

Targeted oncology sequencing panels are among the most popular applications of NGS. Jennings *et al*. outlined the recommendations of a second AMP working group for validation of these NGS-based oncology panels.[4] This paper covered panel design, wet laboratory method, and universal validation issues such as reproducibility, limit of detection, and number and the type of samples needed. The authors also noted the diversity of available bioinformatics tools and the importance of different computational approaches based on variant class and assay design. However, detailed guidance for validation of bioinformatics tools was left to a later working group.

The most recent publication by Roy *et al*. provides the much needed guidance on the design and use of bioinformatics pipelines in clinical NGS. The recommendations in this article are informed by a systematic review of 147 publications on validation of clinical NGS bioinformatics. This literature review again highlighted the dissimilarity of computational methods used to align sequence reads and identify variants.

Each of the recent AMP recommendation papers is the result of collaboration with other organizations having expertise in particular aspects of clinical NGS validation. With its focus on bioinformatics, the most recent recommendations were formed in partnership with the College of American Pathologists (CAP) and the American Medical Informatics Association.

## Comments

The recommendations in Roy *et al*. are extensive and address the issues of validation, study design, selection of appropriate and representative variants, information security, data integrity, quality control/quality assurance, and regulatory compliance. Throughout, the focus of these recommendations remains on the specimen and its associated clinicopathological setting. Outside of *de facto* file format standards, there is no effort to

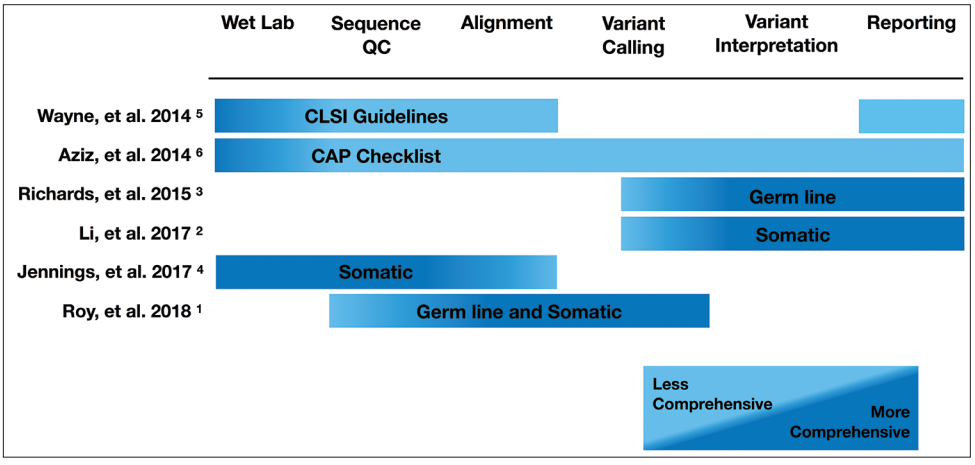


**Figure 1:** Summary of guidelines and recommendations for clinical next-generation sequencing

recommend any specific technology. Unlike most clinical testing approaches, validating all potential findings (variants) on an NGS panel is impractical, and instead, the authors recommend that each potential type of variant (single-nucleotide variant, indel, etc.) is validated, along with its respective specialized variant caller. Because the number of potential variants is too great to validate individually in an analyte-based approach, the choice of an appropriate number of representative variants for use in a method-based validation approach is of considerable importance to laboratory directors. To address this issue, the authors propose a simple but useful formula based on probability and confidence of variant detection.

The 17 recommendations within these guidelines build upon the existing CAP molecular checklist by providing a more general foundation in bioinformatics and more specific guidance for rigorous validation. Directors of established clinical NGS laboratories will find agreement with many of their current validation practices and will likely also identify areas for improvement. For laboratories that are implementing clinical NGS testing for the first time, the "Standards and Guidelines for Validating Next-Generation Sequencing Bioinformatics Pipelines" is a comprehensive guide for all aspects of small sequence variant testing from read alignment through variant annotation. The focus on single nucleotide and short indel variants is an important limitation, as bioinformatics tools to identify structural and copy number variants are already in use in the clinical laboratory. These tools require their own computational approaches and will necessitate additional future recommendations.

**Jeffrey Szymanski[1], Eric Duncavage[1], John Pfeifer[1]**

[1]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA

**Address for correspondence:** Dr. Jeffrey Szymanski, Department of Pathology and Immunology, Division of Laboratory Medicine, Washington University, 660 South Euclid, Campus Box 8118, St. Louis, MO 63110, USA.
E-mail: szymanski@wustl.edu

## References

1. Roy S, Coldren C, Karunamurthy A, Kip NS, Klee EW, Lincoln SE, *et al.* Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: A Joint recommendation of the association for molecular pathology and the College of American Pathologists. J Mol Diagn 2018;20:4-27.
2. Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, *et al.* Standards and guidelines for the interpretation and reporting of sequence variants in cancer: A joint consensus recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. J Mol Diagn 2017;19:4-23.
3. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, *et al.* Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. Genet Med 2015;17:405-24.
4. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, *et al.* Guidelines for validation of next-generation sequencing-based oncology panels: A joint consensus recommendation of the association for molecular pathology and College of American Pathologists. J Mol Diagn 2017;19:341-65.
5. Clinical and Laboratory Standards Institute. CLSI: MM09-A2: Nucleic Acid Sequencing Methods in Diagnostic Laboratory Medicine; Approved Guideline. 2nd ed. Wayne, PA: Clinical and Laboratory Standards Institute; 2014.
6. Aziz N, Zhao Q, Bry L, Driscoll DK, Funke B, Gibson JS, *et al.* College of American Pathologists' laboratory standards for next-generation sequencing clinical tests. Arch Pathol Lab Med 2015;139:481-93.



**Access this article online**

Quick Response Code:

**Website:** www.jpathinformatics.org

**DOI:** 10.4103/jpi.jpi_19_18

**How to cite this article:** Szymanski J, Duncavage E, Pfeifer J. Next-generation sequencing bioinformatics: Guidance between the sequencing and sign out. J Pathol Inform 2018;9:23.

Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2018/9/1/23/235014