INVITAE

Confirmation for clinical genetic testing

Next-generation sequencing (NGS) has largely replaced Sanger sequencing, an older technology, in clinical genetic tests. Compared to Sanger, NGS provides lower costs, higher throughput, and the ability to easily test multiple clinically relevant genes in each patient. In order to minimize the risk of false positives from NGS, a two-step approach is often used, whereby variants uncovered by NGS are confirmed by a separate assay (such as Sanger sequencing). Such confirmatory testing must be "orthogonal" to NGS: it needs to employ different biochemical operating principles and have an uncorrelated chance of error.

Confirmatory testing adds cost, manual labor, and time to the genetic testing process. The ACMG guidelines for NGS state that laboratories should have "extensive experience with NGS... before deciding that result confirmation with orthogonal technology can be eliminated."[1] It has been reported that confirmation of the highest quality NGS variant calls may be unnecessary.[2-5] Moreover, naive use of confirmatory testing can in fact introduce more errors than it actually prevents.[2]

Confirmation is unnecessary and wasteful for high-confidence NGS variant calls

In collaboration with the Partners Laboratory for Molecular Medicine at Harvard and the National Institute of Standards and Technology (NIST), Invitae recently completed the largest study to date on the question of whether and when orthogonal confirmation of NGS results is required.[6] By using both clinical samples (n = 80,000) as well as gold-standard reference samples from NIST, our study considered almost 200,000 variant calls with confirmatory data.

The results reaffirmed other, previous studies in demonstrating that not all variants require confirmation. We showed that high-confidence NGS variant calls can be identified using objective data quality metrics,[6] and that this high-confidence population contains **no** false positives: 100% of the high-confidence variant calls were proven correct by orthogonal data. Many variants meet this "high confidence" criteria and thus do not benefit from confirmation (i.e., confirmation cannot further improve the accuracy of these calls). The remaining, lower confidence calls include a mixture of true and false positives: these cases require, and are resolved by, confirmatory testing.

A significant improvement over others' approaches

The key question is how to consistently identify which NGS calls require confirmation. In this aspect, our study differs from prior publications. Our analysis shows that a battery of quality metrics (based on recommendations in the AMP/CAP NGS bioinformatics guidelines[7]) is required to catch 100% of false positives.[6] Prior studies by other laboratories used only one or two metrics, such as quality score or read depth. We find that these simpler criteria miss some false positives, potentially allowing incorrect pathogenic variants to escape confirmation and be reported as real. We attribute this difference to the size of our study, which was 100 to 1,000 times larger than previous studies, permitting the development of more effective criteria. Our study also employed statistical confidence measures, a critical step that most prior studies did not perform.

Continued on next page

INVITAE

False positive rate and sensitivity in variant calling

There is always a trade-off between sensitivity (the ability to detect variants that are real) and specificity (the ability to avoid false positives). In order to identify clinically important variants with high sensitivity, a wide net must be cast. However, in doing so, a population of lower confidence calls is also identified, some of which are true and some false. Any test that tries to eliminate confirmation by using very strict calling (aiming for high specificity without confirmation) will suffer a sensitivity penalty: true positives will be missed by such a test. Confirmation of some NGS calls continues to be a necessary component of sensitive genetic tests.

Category	Call quality	Action
High-confidence true positives	**Meets highly stringent criteria**	**No benefit from confirmation**
Candidate true positives	**Does not meet stringent criteria**	**Require confirmation**

NGS variants that pass filtering can be placed into high-confidence and intermediate-confidence categories.[6]

Confirmation methods

Variant calls that require confirmation are of many different types, necessitating the use of multiple different confirmation methods. In addition to Sanger sequencing, array CGH, and MLPA, Invitae validated the Pacific Biosciences platform (PacBio) as a confirmation method, showing 100% concordance between PacBio and Sanger.[8] PacBio's technology is highly orthogonal to NGS and can test variants that are difficult for Sanger.[9] Compared to Sanger sequencing, PacBio also provides higher throughput, a higher assay success rate, and improved quality control.[8] By having multiple platforms available, Invitae can use the most appropriate method for each clinical case.

In conclusion

Confirmation significantly increases both cost and turnaround time for patients and clinicians making important healthcare decisions. Our large, interlaboratory study demonstrates that confirmation assays can be focused on a carefully selected subset of variants to deliver high test sensitivity and specificity.

Our team understands that the stakes for clinical genetic testing are high. Results can lead to irreversible action and emotional distress for patients and their families. We are committed to maintaining the highest quality, while continually improving our processes in a responsible and data-driven manner. We hope this study will inform a new standard of data-driven best practices for variant confirmation.

References

1. Rehm HL, Bale SJ, Bayrak-Toydemir P *et al. Genet Med*. 2013;15(9):733-47.
2. Beck TF, Mullikin JC; NISC Comparative Sequencing Program, Biesecker LG. *Clin Chem*. 2016;62(4):647-54.
3. Baudhuin LM, Lagerstedt SA, Klee EW *et al. J Mol Diagn*. 2015;17(4):456-61.
4. Mu W, Lu HM, Chen J *et al. J Mol Diagn*. 2016;18(6):923-32.
5. Strom SP, Lee H, Das K *et al. Genet Med*. 2014;16(7):510-5.
6. Lincoln SE, Truty R, Lin CF *et al. BioRxiv*. 2018, May 18. [Epub ahead of print]
7. Roy S, Coldren C, Karunamurthy A *et al. J Mol Diagn*. 2018;20(1):4-27.
8. McCalmon S, Konvicka K, Reddy NR *et al.* Poster presented at: The American Society of Human Genetics Annual Meeting; 2016, October 18–22;s Vancouver, Canada.
9. Chaisson MJ, Huddleston J, Dennis MY *et al. Nature*. 2015;517:608-611.