GENOOX

TRANSFORMING COMPLEX DATA
INTO CLINICAL ANSWERS

VALIDATING NEXT GENERATION SEQUENCING FOR USE IN CLINICAL LAB PRACTICE

GENE BY GENE

The Genoox platform makes it easy for laboratories hoping to start using NGS based tests to thoroughly validate their process clinically and ensure results are consistent and validated. In addition, the platform provides tools supporting the procedures which are required for obtaining CLIA/CAP certification.

Introduction

Clinical applications of Next Generation Sequencing (NGS) are continually expanding. What was once a "last resort" procedure for exceptional cases, in which a molecular diagnosis could not be achieved by standards means, has become a routine test allowing physicians to translate genomic information into clinically actionable decisions.

However, NGS has its shortcomings, particularly regarding the accuracy of its results. To date, NGS solutions have been associated with higher error rates when compared to Sanger-sequencing, which led to the common practice of confirming NGS results by Sanger. Common problems include unequal coverage throughout the targeted region, sequencing of pseudogenes, and a difficulty in sequencing GC-rich regions.

The Genoox platform makes it easy for laboratories hoping to start using NGS based tests to thoroughly clinically validate their process and ensure that results are consistent and validated. In addition, the platform provides tools supporting the procedures which are required for obtaining CLIA/CAP certification.

Validation method

Genoox was used as the main data analysis platform to technically compare between Sanger sequencing vs. NGS.

Research aims were:

- Validate the Genoox variant calling pipeline for use in clinical practice
- Determine the accuracy of NGS platforms as compared to Sanger-sequencing
- Compare the performance of the common Nextera and SureSelect NGS capture kits
- Compare the accuracy and performance of NGS sequencing results for DNA extracted from blood (peripheralleukocytes) and buccal cells

Sanger-sequencing was completed for 258 genes on DNA extracted from buccal cells of a single individual.

Whole Exome Sequencing (WES) was performed using two commonly used WES enrichment kits: Nextera rapid capture kit and the Agilent capture kit (SureSelectQXT). WES was done in total 10 times in two runs: one including duplicates of both buccal and leukocyte-derived DNA, and one including only leukocyte-derived DNA (yielding a total of 5 NGS assays per each kit).

Each WES sample was processed using the Genoox platform. The platform provides variants calling using multiple variant calling algorithms to allow better variant detection sensitivity and low false detection rates. In addition, the platform provides coverage reports over the alignment (BAM) data and other QC metrics in order to detect regions which were not properly covered. For this research, it was used in order to identify and distinguish between missing variants due to insufficient coverage despite having been targeted by the enrichment kit, and those which are missing since the enrichment kit does not target them.

Lastly, the platform was used in order to compare between the different samples in order to determine and collate the sensitivity levels of each of the methods using each of the kits.

Results

A total of 449 variants were identified in all platforms and samples. Of those, 407/449 (90.6%) variants were detected by all assays. With respect to the 42 variants (9.4%) not identified by all assays, 13 were discordant between Sanger and at least 5/10 NGS assays, and 29 were discordant between NGS assays. Of the 42 variants not detected by all assays, 23 were considered to be true variant calls (including 4 Sanger false negatives and 19 NGS false negatives), and 17 were false positives (3 Sanger, 14 NGS). In summary, 430/449 (95.8%) variants were considered to be true variant calls.

The overall concordance between all platforms and samples was 407/449 (90.6%) with mean "true variants" (Sensitivity) detection rate of ~99% for Sanger-sequencing and SureSelect, and ~c98% for Nextera, with few false calls for each of those methods. The majority of NGS false calls were explained by low coverage, either specific or random. The majority of Sanger false calls were false negatives/positives related to primer design.

The research showed that none of the methods (NGS and Sanger) outperformed the other and no statistically significant differences were noted between the two sequencing methods with respect to the ability to sequence the targeted regions and the observed detection rates.

While no sequencing method outperformed the other, the NGS advantages were clear:

- Direct sequencing requires the lab to design specific primers per region
- The direct sequencing process is time consuming and requires substantial lab work
- NGS allows for analyzing the entire exome or genome at no additional effort and without compromising the variant detection sensitivity.

The Genoox Platform

Genoox removes the final barrier to widespread adoption of clinical NGS, enabling personalized medicine for mainstream patients. Our platform is used in the diagnosis and treatment of genetic disorders and cancer, as well as new drug discovery and family planning.

The Genoox analytical process includes data compression, variant assessment, classification, curation, analysis and verification. Our automated, built-in quality systems ensure accurate fast and simple targeted patient care, and avoid the costly errors associated with missing an important correlated variants, or producing noisy results due to false positives calls. The Genoox analytics dashboard displays evidence and confidence scores with associated details for each variant detected, allowing to distinguish between true and false called variants.

Our proprietary algorithms were clinically validated against both public data sets and in-house data sets, providing an exceptional level of confidence with variant calls as well as a significant reduction in sequencing costs. The Genoox search engine avoids the common problem of relying on limited or fragmented research that miss important results, by scoring hundreds of available data sources, both public and private, for relevant genetics findings.

ESHG

To read full article Click Here>>