

Poor coverage of specific gene sets in exome sequencing gives cause for concern

June 1 2014

Researchers have analysed 44 exome datasets from four different testing kits and shown that they missed a high proportion of clinically relevant regions. At least one gene in each exome method was missing more than 40 percent of disease-causing genetic variants, and the worst-performing method missed more than 90 percent of such variants. This means that there is a substantial possibility of reporting false negative results, they say.

With services based on exome sequencing becoming affordable to patients at a reasonable price, the question of the quality of the results provided has become increasingly important. The exome is the DNA sequence of genes that are translated into protein. These protein-coding regions contain most of the currently-known disease-causing genetic mutations. The American College of Medical Genetics and Genomics (ACMG) has recommended the reporting to patients of clinically actionable incidental genetic findings in the course of clinical exome testing. Specifically, mutations of 56 specific genes with known clinical importance should be reported even when they are incidental to the patient's current medical condition. However, a new study to be reported to the annual conference of the European Society of Human Genetics (ESHG) today (Sunday) shows that exome sequencing, as currently performed, does not always produce high quality results when examining subsets of genes such as the 56 ACMG genes.

Dr Eric Londin, Assistant Professor in the Computational Medicine Centre, Department of Pathology, Anatomy and Cellular Biology,

Thomas Jefferson University, Philadelphia, USA, will tell the conference that analysis of 44 exome datasets from four different testing kits showed that they missed a high proportion of clinically relevant regions in the 56 ACMG genes. "At least one gene in each exome method was missing more than 40 percent of disease-causing genetic variants, and we found that the worst-performing method missed more than 90 percent of such variants in four of the 56 genes," he says.

A central question, the researchers say, is not how often a clinical diagnosis can be made using exome sequencing, but how often it is missed, and the study shows clearly that there is a high false-negative rate using existing sequencing kits. "Our concern is that when a clinical exome analysis does not report a disease-causing genetic variant, it may be rather that the location of that variant has not been analysed rather than the patient's DNA being free of a disease-causing variant," says Dr Londin. "Depending on the method and the laboratory, a significant fraction (more than ten percent) of the exome may be untested and this raises concerns as to how results are being communicated to patients and their families. "

A total of 17,774 disease-causing genetic variants are annotated in the Human Gene Mutation Database (HGMD) for the 56 genes mentioned in the ACMG recommendations. The researchers examined the coverage of the exome datasets for the locations where the 17,774 disease-causing variants can occur. Although the exome datasets are comparable in quality to other published clinical and research exome data sets, the coverage of the disease-causing locations was very heterogeneous and often poor. The researchers believe that clinical laboratories that implement the ACMG reporting guidelines should recognise the substantial possibility of reporting false negative results.

One potential improvement would be to have clinical exome sequencing use methods designed to provide a maximum yield of all clinically

relevant genes. "Many of the currently used exome kits are designed to provide a very broad dataset including genomic features that do not yet have a well-established clinical association. There is a need to develop new kits and methods which provide adequate and reliable coverage of genes with known disease associations. If adequate performance cannot be obtained across the exome, then further use of targeted disease-specific panels of [genes](#) should be explored," Dr. Londin says.

The study also found that exome datasets generated from low amounts of sequence data (fewer than six gigabases) performed much worse than datasets that were generated from higher amounts of sequence data (more than ten gigabases). This finding is consistent with previous studies showing that exome methods do not have a linear relationship between sequence-generated and nucleotide coverage. Instead, a minimum threshold of sequencing data needs to be met before optimum nucleotide coverage is obtained.

"Current consensus and regulatory guidelines do not prescribe a minimum data requirement for clinical exome tests. The result is that when a causative variant cannot be identified it does not necessarily imply that the variant is not present, rather that there may be a technical issue with the exome technology used. In other words, a clinical 'whole exome' study may not be 'wholesome' in coverage. Patients and their families should be made aware of this problem and of the implications of the genomic findings of clinical exome sequencing in its current state," Dr. Londin will conclude.

Provided by European Society of Human Genetics

Citation: Poor coverage of specific gene sets in exome sequencing gives cause for concern (2014, June 1) retrieved 25 March 2023 from <https://medicalxpress.com/news/2014-06-poor-coverage-specific-gene-exome.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.