

Measuring How Many Pathogenic Variants Are Uncovered By Exome Sequencing

November 18, 2013

JM Kocarnik

Whole-exome and whole-genome sequencing are increasingly being incorporated into both clinical and research settings. Whole-exome refers to sequencing only gene-coding portions (the exons) of the whole genome. These broad-coverage approaches raise the possibility of incidental findings--genetic variants that increase risk for diseases that may not be related to the reason the test was performed. Much recent debate has centered on whether, and which of, these so-called incidental findings must be reported. The pervasiveness of the problem, however, had not yet been systematically studied insofar as how many people could be expected to harbor such a finding, or the amount of time necessary to evaluate the importance of potential findings. A pilot study to fill these knowledge gaps was recently completed by a group of researchers and clinicians at the University of Washington, many of whom frequently collaborate with researchers at the Fred Hutchinson Cancer Research Center. Together with Dr. Wylie Burke in the Public Health Sciences Division, the researchers reported their findings in *The American Journal of Human Genetics*.

According to senior author Dr. Gail Jarvik, this report “is particularly timely, since the American College of Medical Genetics and Genomics (ACMG) recently suggested return of incidental findings from genomic tests”. These ACMG guidelines recommend that pathogenic variants in 56 genes should be reported as incidental findings. This raised questions, however, on how many incidental findings might need to be reported, and how much burden this might impose on the medical system. Fortunately, says Jarvik, “we were able to use data from 1000 of the 6500 participants in the NHLBI Exome Sequencing Project (ESP)” to answer these questions. Importantly, this data allowed the researchers to provide estimates for multiple ancestry groups.

To start, the authors expanded on the ACMG list to evaluate 114 genes associated with medically actionable genetic conditions that might remain undiagnosed in adults. Sequencing results for 500 individuals of African descent and 500 individuals of European descent were screened for genetic variants that might be associated with actionable conditions. Any variants found in the predetermined 114 genes of interest were then evaluated by a multistep process to differentiate between variants that should be returned (i.e. those expected to be pathogenic) and those that should not be returned (i.e. variants of uncertain significance).

Initially, 585 instances of 239 unique variants were identified as ‘disease-causing’ in the Human Gene Mutation Database (HGMD). These variants were further evaluated by a team of expert reviewers (medical geneticists, genetic counselors, and certified molecular geneticists). Each of the 239 unique variants required an average of 23 minutes to evaluate (range 1 to 135 minutes). Following literature review and application of stringent evaluation criteria, only 17 of these variants (in 18 participants) were determined to be pathogenic or likely pathogenic. An additional 5 variants not listed as disease-causing in HGMD were also identified. In total, this meant that the exome sequences of 23 out of 1000 participants harbored pathogenic or likely pathogenic variants that could be reported as incidental findings.

Overall, these results “underscored the lack of good clinical classification data for most variants,” says Jarvik, and that “accurate databases that are comprehensive and publically available are needed.” The majority of the variants identified as ‘disease-causing’ by HGMD in this study, for example, did not meet rigorous criteria for classification as high-penetrance pathogenic mutations appropriate to report. Databases that suggest variants are pathogenic but lack sufficient evidence may increase the time required to accurately evaluate and classify variants. Furthermore, says Jarvik, “erroneous classifications could lead to unnecessary worry and treatment”.

Another important finding in this study was that both the number of variants classified as ‘disease-causing’ in HGMD and the number of participants with incidental findings were larger in the European-descent group than the African-descent group (see figure). The reason for this was the “particular dearth of information to classify variants in participants of African descent,” says Jarvik. This suggests that genomic testing is currently more informative for those of European descent, a health disparity that further emphasizes the need for an accurate database informed by studies in multiple ancestry groups.

Given the success of this pilot study, “we are now moving on to classify the variants for the remaining 5500 people in the NHLBI ESP”, says Jarvik. Furthermore, “we are sharing our classifications in a national database.” These future efforts should help provide the information needed for the reliable, efficient, and equitable application of genomic medicine.

[Dorschner MO, Amendola LM, Turner EH, Robertson PD, Shirts BH, Gallego CJ, Bennett RL, Jones KL, Tokita MJ, Bennett JT, Kim JH, Rosenthal EA, Kim DS, National Heart, Lung, and Blood Institute Grand Opportunity Exome Sequencing Project, Tabor HK, Bamshad MJ, Motulsky AG, Scott CR, Pritchard CC, Walsh T, Burke W, Raskind WH, Byers P, Hisama FM, Nickerson DA, Jarvik GP.](#) 2013. Actionable, pathogenic incidental findings in 1,000 participants’ exomes. *Am J Hum Genet.* 93(4):631-40.

See also: [Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, McGuire AL, Nussbaum RL, O'Daniel JM, Ormond KE, Rehm HL, Watson MS, Williams MS, Biesecker LG. 2013. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* 15\(7\):565-74.](#)

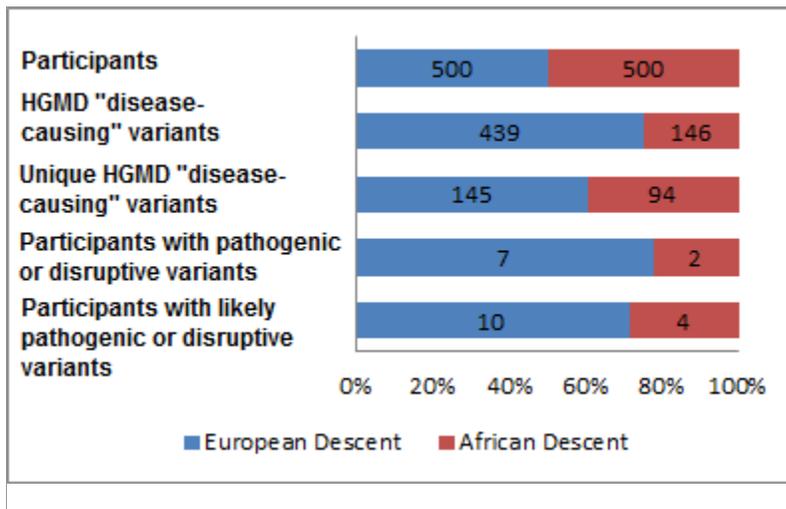


Image provided by Jonathan Kocarnik

Number and percentage of variants and identified from performing exome sequencing in 1000 individuals, and number and percentage of participants with reportable incidental findings, by ancestry. HGMD = Human Gene Mutation Database.