Novel gene-searching software improves accuracy in disease studies

January 22 2013

A novel software tool, developed at The Children's Hospital of Philadelphia, streamlines the detection of disease-causing genetic changes through more sensitive detection methods and by automatically correcting for variations that reduce the accuracy of results in conventional software. The software, called ParseCNV, is freely available to the scientific-academic community, and significantly advances the identification of gene variants associated with genetic diseases.

"The algorithm we developed detects copy number variation associations with a higher level of accuracy than that available in existing software," said the lead inventor of ParseCNV, Joseph T. Glessner, of the Center for Applied Genomics at The Children's Hospital of Philadelphia. "By automatically correcting for variations in the length of deleted or duplicated DNA sequences from one individual to another, ParseCNV produces high-quality, highly replicable results for researchers studying genetic contributions to disease."

Glessner is the lead author of a study describing ParseCNV, published Jan. 4 in Nucleic Acids Research.

Copy number variations (CNVs) are particular sequences of DNA, ranging in length from 1000 to millions of nucleotide bases, which may be deleted or duplicated. While in any given region of a person's DNA, CNVs are very rare, everyone's genome has CNVs, many of which play important roles in causing or influencing disease.

In searching for associations between CNVs and diseases, researchers typically perform case-control studies, comparing [DNA samples](#) from patients to DNA from healthy individuals, looking for telltale differences in how CNVs are overrepresented or underrepresented.

CNVs, however, occur in multiple types among individuals, said senior author Hakon Hakonarson, M.D., Ph.D., director of the Center for Applied Genomics at The Children's Hospital of Philadelphia. "One person may have a 60-kilobase deletion, while another may have a 100-kilobase deletion; that may determine the difference between a healthy state versus disease. Many CNV detection softwares may misread the boundary of a CNV region, which could lead to a misclassification and result in false-positive or false-negative associations."

ParseCNV is designed with built-in corrections to adjust for these size variations and other red flags that confound results. Using polymerase chain reaction testing to validate the initial findings, the study team determined that the software had called 90 percent of the CNVs accurately—a better rate than conventional CNV association softwares, which typically produce validation rates that are notably lower.

The authors say the program's comprehensive design, statistical capabilities, and quality-control features lend it versatility, applicable not just to case-control studies, but also to family studies, and quantitative analyses of continuous traits, such as obesity or height.

Glessner says the Center for Applied Genomics team will continue to refine ParseCNV's features as CNV research progresses. Hakonarson adds that the ParseCNV algorithm will advance genomic diagnostics: "It is likely to play a future key role as a research tool in improving detection of CNV association in individual patients enrolled in disease studies—perhaps through an initial diagnostic screen, to be followed up

with a CLIA-certified laboratory test."

   More information: "ParseCNV integrative copy number variation association software with quality tracking," *Nucleic Acids Research*, published online Jan. 4, 2013. dx.doi.org/10.1093/nar/gks1346

Provided by Children's Hospital of Philadelphia

Citation: Novel gene-searching software improves accuracy in disease studies (2013, January 22) retrieved 13 March 2024 from https://medicalxpress.com/news/2013-01-gene-searching-software-accuracy-disease.html