AgileROH: Rapid identification of autozygous regions using Illumina short read sequencing data

Ian Carr[1], Christopher Watson[2], Morag Raynor[1], Marilena Elpidorou[1], Ummey Hany[1], Laura Crinnon[2], Eamonn Sheridan[1], Alexander Markham[1], James Poulter[1], and David Bonthron[1]

[1]University of Leeds
[2]Leeds Teaching Hospitals NHS Trust

August 1, 2022

Abstract

Rare autosomal recessive diseases are a major cause of mortality and morbidity. They occur more frequently in individuals with consanguineous parents, in which case the pathogenic variants are often located within regions of genetic identity by descent. A well-established and effective way of identifying these "autozygous" genomic regions has been to search for runs of homozygous genotypes in microarray SNP data. However, with the widespread use of whole genome and exome sequencing in both diagnostic and research settings, it has become desirable to be able to both map autozygous regions and to identify the deleterious variants using a single data set. We have developed and optimized an algorithm, implemented as a set of three applications, that identifies autozygous regions in consanguineous individuals using standard whole exome variant data. These applications have been successfully used in both research and diagnostic settings. Availability and implementation: User guides, compiled programs and sample VCF files are freely available at http://www.dna-leeds.co.uk/agile/AgileROH/ and http://www.dna-leeds.co.uk/agile/AgileMultiIdeogram/.

Hosted file

`AgileROH_manuscript.docx` available at https://authorea.com/users/498522/articles/579279-agileroh-rapid-identification-of-autozygous-regions-using-illumina-short-read-sequencing-data