nature.com   Publications A-Z index   Browse by subject

Basic statistical analysis in genetic case-control studies

Geraldine M Clarke, Carl A Anderson, Fredrik H Pettersson, Lon R Cardon, Andrew P Morris & Krina T Zondervan

Abstract

Abstract   **author Information**   **supplementary Information**

This protocol describes how to perform basic statistical analysis in a population-based genetic association case-control study. The steps described involve the (i) appropriate selection of measures of association and relevance of disease models; (ii) appropriate selection of tests of association; (iii) visualization and interpretation of results; (iv) consideration of appropriate methods to control for multiple testing; and (v) replication strategies. Assuming no previous experience with software such as PLINK, R or Haploview, we describe how to use these popular tools for handling single-nucleotide polymorphism data in order to carry out tests of association and visualize and interpret results. This protocol assumes that data quality assessment and control has been performed, as described in a previous protocol, so that samples and markers deemed to have the potential to introduce bias to the study have been identified and removed. Study design, marker selection and quality control of case-control studies have also been discussed in earlier protocols. The protocol should take ~1 h to complete.

Subject terms: Computational biology   Genetic analysis   Genomics   Model organisms

ARTICLE PREVIEW
view full access options

Author information

Abstract   author Information   **supplementary Information**

Affiliations

Genetic and Genomic Epidemiology Unit, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.

Geraldine M Clarke,Fredrik H Pettersson,Andrew P Morris &Krina T Zondervan

Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK.

Carl A Anderson

GlaxoSmithKline, King of Prussia, Pennsylvania, USA.

Lon R Cardon

Contributions

G.M.C. wrote the first draft of the manuscript, wrote scripts and performed analyses. G.M.C., C.A.A., A.P.M. and K.T.Z. revised the manuscript and designed the protocol. L.R.C. conceived the protocol.

Competing financial interests

The authors declare no competing financial interests.

Corresponding author

Correspondence to: Geraldine M Clarke

Supplementary information

Abstract   **author Information**   supplementary Information

Zip files

1. Supplementary Data 1 (197MB)
   Example genome wide association (GTA) data.

2. Supplementary Data 2 (84KB)
   Example candidate gene 9 (CG) data.