NEALE LAB

RAPID GWAS OF THOUSANDS OF PHENOTYPES FOR 337,000 SAMPLES IN THE UK BIOBANK

September 20, 2017

The UK Biobank recently released genome-wide association data on ~500,000 individuals. The genotype data for these samples have been cleaned, imputed and released to the scientific community. This public release of data represents an extraordinary advance for genetics, pushing the envelope for data sharing and rapid uptake by the research community. These data will be used for novel discovery of disease-associated genes, in the development of new methods, and to serve as an example for how future efforts in genetics and biology ought to proceed.

To further enhance the value of this resource, we have performed a basic association test on ~337,000 unrelated individuals of British ancestry for over 2,000 of the available phenotypes. We're making these results available for browsing through several portals, including the Global Biobank Engine where they will appear soon. They are also available for download here.

We have decided not to write a scientific article for publication based on these analyses. Rather, we have described the data processing in a detailed blog post linked to the underlying code repositories. The decision to eschew scientific publication for the basic association analysis is rooted in our view that we will continue to work on and analyze these data and, as a result, writing a paper would not reflect the current state of the scientific work we are performing. Our goal here is to make these results available as quickly as possible, for any geneticist, biologist or curious citizen to explore. This is not to suggest that we will not write any papers on these data, but rather only write papers for those activities that involve novel method development or more

NEALE LAB

is now a relatively well-established activity, and while the scale of this is a bit grander than before, that in and of itself is a relatively perfunctory activity. Simply put, let the data be free.

We do view these results as likely to change as we continue to refine the quality control analyses and as we continue to dig into the results themselves. Nevertheless, we've started to use them in a variety of downstream analyses and for other scientific projects and hope that others find them useful too.

If you would like more information, please contact us at: nealelab.ukb@gmail.com.

FAQ

Where can I find the GWAS results?

All summary statistics will be browsable through the Global Biobank Engine soon, and are already available for download here. This Global Biobank Engine is a public browser that allows you to explore the results for different loci and phenotypes.

What is Hail and how can I use it?

Hail is an open-source, scalable framework for exploring and analyzing genomic data. For this analysis, we ran Hail on Google Cloud. Have a go at using Hail yourself here.

What phenotypes did you analyze?

We analyzed 2,419 phenotypes which were processed using a customized version of PHESANT. A list of the phenotypes can be found in the manifest where our results are available for download.

What UK Biobank applications cover this activity?

This work was conducted under UK Biobank applications 18597 (V. Anttila) and 11898 (J. Hirschhorn).

In sharing these results before publication aren't you worried about getting "scooped"?

single cohort is insufficient for a (good) publication. Replication in an

independent cohort, careful and thorough examination of putative associated

loci, and secondary analyses to interpret and give insight into the relevance

and meaning of GWAS hits are what we think important for advancing

biological understanding. In sharing these results as soon as they are

produced, we hope to dramatically cut down the time researchers spend in

obtaining GWAS results so that they may focus on all the other components

that are less standardized across studies.

The team:

Liam Abbott

Verneri Anttila

Krishna Aragam

Alex Baumann

Jon Bloom

Sam Bryant

Claire Churchhouse

Joanne Cole

Mark J. Daly

Rob Damian

Andrea Ganna

Jackie Goldstein

Mary Haas

Joel Hirschhorn

Daniel Howrigan

Eric Jones

Dan King

Ruchi Munshi

Benjamin Neale

Duncan Palmer

Tim Poterba

Manuel Rivas

Cotton Seed

Sailaja Vedantam

Raymond Walters

NEALE LAB

Authored by Claire Churchhouse and Ben Neale

f 🐦 ⓟ ♥

PREVIOUS

Heritability of >2,000 traits and disorders in the UK Biobank

TWITTER   INTRANET

Powered by Squarespace