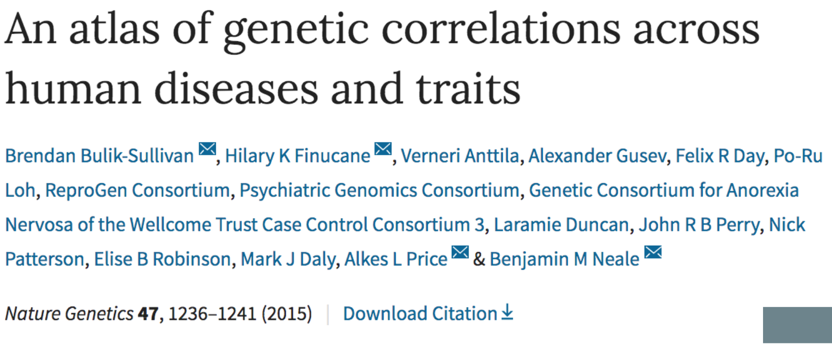
Step 2: Estimating ldsc

LDSC produces genetic "heat maps" of genetic correlations across traits

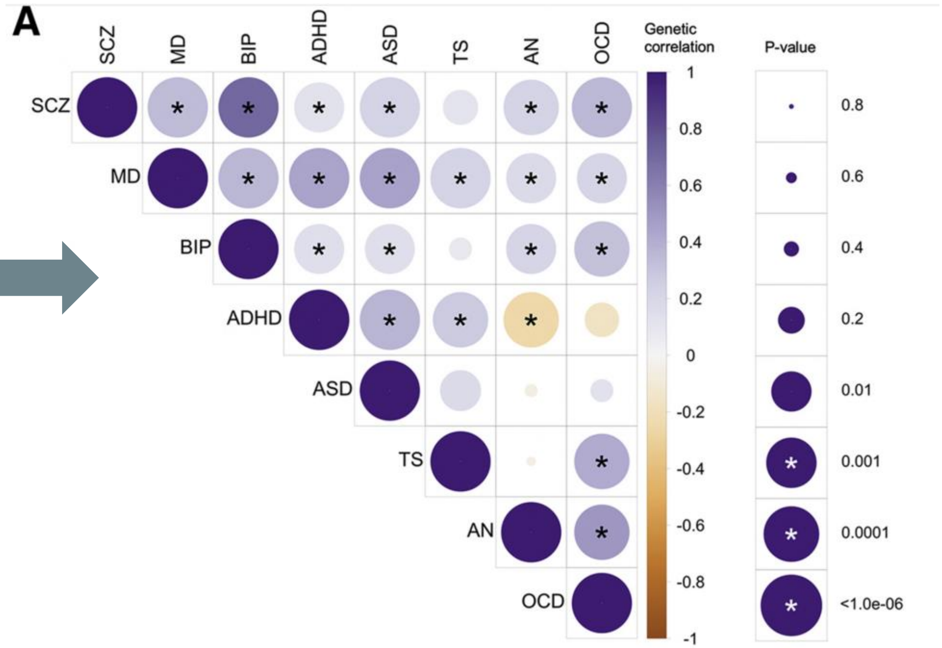


An atlas of genetic correlations across human diseases and traits

Brendan Bulik-Sullivan, Hilary K Finucane, Verneri Anttila, Alexander Gusev, Felix R Day, Po-Ru Loh, ReproGen Consortium, Psychiatric Genomics Consortium, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3, Laramie Duncan, John R B Perry, Nick Patterson, Elise B Robinson, Mark J Daly, Alkes L Price & Benjamin M Neale

Nature Genetics **47**, 1236–1241 (2015) | Download Citation ↓

LDSC estimates genetic correlations between samples with varying degrees of sample overlap using publicly available data



Bulik-Sullivan et al., 2015; Lee et al., 2019

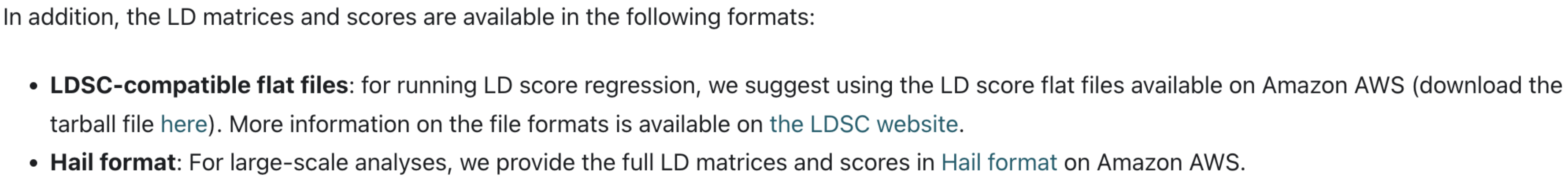
Where to get ancestry-specific LD-scores

- European and East Asian available in Box link on GenomicSEM wiki:

https://utexas.app.box.com/s/vkd36n197m8klbaio3yzoxsee6sxo11v

- African, East Asian, Central/South Asian and Middle Eastern LD-scores are available on the panUKB website:

https://pan.ukbb.broadinstitute.org/downloads



In addition, the LD matrices and scores are available in the following formats:

- **LDSC-compatible flat files**: for running LD score regression, we suggest using the LD score flat files available on Amazon AWS (download the tarball file here). More information on the file formats is available on the LDSC website.
- **Hail format**: For large-scale analyses, we provide the full LD matrices and scores in Hail format on Amazon AWS.

The *ldsc* function takes 6 arguments:

1. **traits**: a vector of file names/paths to files which point to the munged GWAS data
2. **sample.prev**: A vector of sample prevalences of length equal to the number of traits. Enter 0.5 when inputting sum of effective N for binary traits.
3. **population.prev**: A vector of population prevalences. If the trait is continuous the values should equal NA.
4. **ld**: A folder of LD scores used as the independent variable in LDSC
5. **wld**: A folder of LDSC weights (Typically same folder as specified for the ld argument)
6. **trait.names**: The trait names.

Key output from LDSC

LDSC_INT$S : the genetic covariance matrix

Functions like cov2cor would give you the genetic correlation matrix

LDSC_INT$V: the sampling covariance matrix