loneliness.readme.pdf

Jianjun Gao, et al Neuropsychopharmacology 2016
http://www.ncbi.nlm.nih.gov/pubmed/27629369

INTRODUCTION

These are the summary statistics of the Genome-Wide Association Study (GWAS) of loneliness using a combination of genotyped and imputed SNPs and limited to individual of European descent. Citation for all studies that use any of these data: DOI:10.1038/npp.2016.197. These results are shown in Figure 1 of that paper.

All genotype and phenotype data were collected as part of the Health and Retirement Study (http://hrsonline.isr.umich.edu/). The Health and Retirement Study genetic data is sponsored by the National Institute on Aging (grant numbers U01AG009740, RC2AG036495, and RC4AG039029) and was conducted by the University of Michigan." Summary statistics are shared here with permission.

DISCLAIMER

These data are provided "as is", and without warranty, for scientific and educational use only. It is your responsibility to use the data correctly. If you download these data, you acknowledge that these data will be used only for non-commercial research purposes; that the investigator is in compliance with all applicable state, local, and federal laws or regulations and institutional policies regarding human subjects and genetics research; that secondary distribution of the data without registration by secondary parties is prohibited.

To prevent identifiability of individual participants, we only distribute summary statistics data.

METHODS. See paper for full details. Briefly:

We derived three phenotypes from the loneliness scale for GWAS study: 1) “linear” - a continuous phenotype obtained by summing the scores from all three questions, thus yielding a score between 3 (least lonely) and 9 (most lonely); 2) “multivariate” – a single score for each question ranging from 1 (least lonely) to 3 (most lonely); and 3) “case: control” – a dichotomous score in which participants who answered 1 on all three items were considered controls (totally loneliness score = 3), and individuals with a loneliness score of ≥ 6 were considered cases (participants with scores of 4 or 5 were treated as missing).

We used Linear Mixed Model (LMM) or Multivariate Linear Mixed Models (MLMM) implemented in the Genome-wide Efficient Mixed Model Association (GEMMA) software package to further correct for residual population structure due to ancestry or cryptic relatedness (Zhou and Stephens, 2012) in our GWAS. We examined the linear, multivariate and case:control phenotype models using a combination of genotyped and imputed SNP data, adjusting for sex, age and marital status. Marital status data available from HRS included multiple unmarried categories (e.g. single, widowed, etc) which we collapsed to create a binary variable. We excluded SNPs with MAF<0.01. For the case:control studies, controls were coded as 0 and

cases were coded as 1, as suggested in the GEMMA documentation (Zhou et al, 2012). The association analyses were performed using the 7,556 European American subjects.

DOWNLOAD FILES

The following files are available for download:

Linear4PGC.txt	Summary statistics for all SNPs; loneliness as a continuous variable (5,768,558 SNPs)
Multi4PGC.txt	Summary statistics for all SNPs; loneliness as a single score for each question (multivariate analysis, 5,768,558 SNPs)
CaCr4PGC.txt	Summary statistics for all SNPs; loneliness as a dichotomous score (3 = not lonely, 6-9 = lonely, 4 and 5 excluded. 5,768,486 SNPs actually analyzed rather than 5,768,507 in Table S1 of our paper, which we forgot to remove head-lines for CHR2-22)

The files contain the following common information (columns):

SNPID CHR BP A1 A2 BETA SE PVALUE remle EUR_MAF

SNPID SNP ID

CHR chromosome number on which the SNP is located (build 37, hg19)

BP base pair position of the SNP (build 37, hg19)

A1 effect allele

A2 alternate allele

BETA unstandardized regression coefficient in LMM

SE standard error of the beta

PVALUE p-value associated with the effect size from Wald test (performed using GEMMA)

remle remle estimates for lambda

EUR_MAF: minor allele frequency from European super population of 1000 Genomes Project - Phase 3

REFERENCES:

Gao J, Davis LK, Hart AB, Sanchez-Roige S, Han L, Cacioppo JT, Palmer AA. Genome-Wide Association Study of Loneliness Demonstrates a Role for Common Variation. Neuropsychopharmacology. 2016 Sep 15. doi: 10.1038/npp.2016.197. PMID: 27629369

Zhou X, Stephens. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* M (2012) **44**(7): 821-824.

Health and Retirement Study. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant numbers U01AG009740, RC2AG036495, and RC4AG039029), Ann Arbor, MI.