**Gene-based test for rare variants**

**GCTA-ACAT: a fast gene- or set-based association test based on Cauchy distribution**

This method, Aggregated Cauchy Association Test (ACAT), was originally proposed by [Liu et al, 2019](https://www.sciencedirect.com/science/article/pii/S0002929719300023). It is a general, powerful, robust, and computationally efficient p-value combination test method for rare variants. Only summary statistics are required for the test. We have implemented ACAT by efficient C code and made it compatible with the output from [fastGWA-GLMM](https://cnsgenomics.com/software/gcta/#fastGWA-GLMM). The test is very efficient. For example, to test through 26,292 genes (~11 million SNPs), the average runtime is around 30 seconds per trait. Credits: [Longda Jiang](mailto:longda.jiang@uq.edu.au), [Hailing Fang](mailto:fanghailing@westlake.edu.cn) and [Jian Yang](http://researchers.uq.edu.au/researcher/2713).

**References**

Liu Y, Chen S, Li Z, Morrison AC, Boerwinkle E, Lin X (2019) ACAT: A fast and powerful p value combination method for rare-variant analysis in sequencing studies. The American Journal of Human Genetics, 104(3), 410-421.

Jiang L, Zheng Z, Yang J (2021) FastGWA-GLMM: a generalized linear mixed model association tool for biobank-scale data, 12 February 2021, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-128758/v1](https://europepmc.org/article/PPR/PPR283012)

--gene-list gene\_list.txt  
Input gene list with gene start and end positions.

Input file format  
gene\_list.txt (columns are gene ID, chromosome, left and right boundary of the gene region)

1 19774 19899 Gene1

1 34627 35558 Gene2

......

Please click the link below to download the gene list file (provided by [Plink1.9](https://www.cog-genomics.org/plink/1.9/resources#genelist)).

Gene list (hg19): [glist-hg19.txt](file:///Users/uqljian5/Documents/gcta2_doc_github/build/res/glist-hg19.txt)

--snp-list gwas.fastGWA  
The GWAS summary statistics produced by [fastGWA-GLMM](https://cnsgenomics.com/software/gcta/" \l "fastGWA-GLMM).

--max-maf 0.01  
The maximum minor allele frequency (MAF) allowed for a variant to be included in the ACAT test. Any variant with MAF larger than this value will be excluded.

--min-mac 10  
The minimum minor allele count (MAC) allowed for a variant to be included in the ACAT test. Any variant with MAC smaller than this value will be excluded. Please note that the [original ACAT test](https://www.sciencedirect.com/science/article/pii/S0002929719300023) aggregates variants with MAC <= 10 and performs a burden test based on raw genotypes of the variants. However, this has not been implemented in GCTA-ACAT, because in a standard GWAS, variants with such small MAC are usually removed during the QC process.

--wind 0  
A value to define a flanking gene region (unit: kilobase). By default, the original region (+-0kb) defined in the gene\_list.txt file is used to map and extract variants.

Examples

# Gene-based ACAT test for rare variants

gcta64 --acat --maf 0.01 --snp-list assoc.fastGWA --gene-list gene\_list.txt --max-maf 0.01 --min-mac 10 --wind 0 --out test.acat.res

Output file format (columns are chromosome, gene ID, left and right boundary of the gene region, number of qualified variants in the gene region, and ACAT test p-value)

chr gene\_name start stop snp\_num p\_acat

1 FAM87B 752750 755214 2 0.772495

1 LINC01128 762970 794826 13 0.250037

1 LOC100130417 852952 854817 4 0.416386

......