

---

# **Pathway-PRS User's Guide**

**Version 1.0**

**January 2023**

---

Developers: Feng Zhang

URL: <https://github.com/LCE-XJTU/Pathway-PRS.git>

Email: [fzhxjtu@xjtu.edu.cn](mailto:fzhxjtu@xjtu.edu.cn)

## 1. Running pathway-PRS

Pathway-PRS is developed by R for efficient data analysis. Please make sure that R (<http://www.r-project.org/>) has been installed on your system. Pathway-PRS is the command lines. Unzip the downloaded "Pathway-PRS" package and run "code.R" program at your terminal window to start analysis.

## 2. Parameter Setting

Pathway-PRS will ask the user to prepare SNP lists in advance before starting the analysis:

**[1] SNP list:** Please input the storage path and name of the SNP list. We have provided an example file (SNPlist.xlsx) in the Pathway-PRS package. Users can also use their own SNP list. SNP list should contain basic information about each SNP, such as rsid, chromosome, position, effect allele, other allele, beta, P-value, and gene name, in which rsid, effect allele, beta, and gene name are required.

**Note:** Each line of this file records the details of one SNP.

**[2] Gene ontology gene sets and the canonical pathways:** The Gene ontology (GO) gene sets and the canonical pathways from the Gene Set Enrichment Analysis (GSEA) (<https://www.gsea-msigdb.org/>) have been provided in the Pathway-PRS package to extract the SNP set of each path. Users can also download the gene sets from the GSEA by themselves.

## 3. Output files Pathway-PRS will output three result files:

**[1] score file:** Each score file (such as score1, score2, ... score1000) represents one of the pathways. Each line of the file records the basic information of one SNP (including rsid, effect allele, and beta). This file will eventually be used to calculate the PRS for each pathway.

*Example:*

```
rs1625895 T -0.0290993078252748
rs6847780 T -0.0176956485356892
rs17246544 T -0.0182963622831166
rs10954668 A -0.0208049294331883
rs8007110 A -0.02960390325858
rs196010 A 0.0172994970780611
rs11641592 A -0.0183982139464423
rs8046118 A -0.0306035477981776
rs12972559 T 0.0263010744193707
```

**[2] name\_score.txt:** The name\_score.txt file is used to record the gene and pathway name/term represented by each score.

*Examples:*

```
score1 PRKAR1B BIOCARTA_AGPCR_PATHWAY
score2 GIT2 BIOCARTA_AGR_PATHWAY
score3 PRKCE BIOCARTA_AKAPCENTROSOME_PATHWAY
score4 MYL2 BIOCARTA_ALK_PATHWAY
score5 PRKCB BIOCARTA_ARENRF2_PATHWAY
score6 TP53 BIOCARTA_ARF_PATHWAY
score7 PRKCB BIOCARTA_AT1R_PATHWAY
score8 TP53 BIOCARTA_ATM_PATHWAY
score9 TP53 BIOCARTA_ATRBRCA_PATHWAY
score10 IGF1R BIOCARTA_BAD_PATHWAY
```

**[3] SNPlist.txt:** The SNPlist.txt file is used to record the “rsid” of all the SNPs involved in the analysis.

*Examples:*

```
rs16958751
rs7169763
rs73055696
rs12369015
rs11207855
rs2782549
rs437021
rs12088813
rs11590808
rs3782274
```