

Guidance to reproduce SUMMIT's results

Simulation

Figure 2(a) and 2(b)

```
Rscript /gpfs/research/chongwu/zichenzhang/SUMMIT-test/code/Simulation.R \
--h2_e A \
--h2_p 0.2 \
--p_causal B \
--n 31684 \
--sumstats TRUE \
--gene_ENSG ENSG00000258289 \
--UKB TRUE \
--folder_output SIM \
--t1e FALSE \
--seed SEED \
```

| A | B | SEED |
|-------|------|------|
| 0.005 | 0.01 | 1 |
| 0.005 | 0.05 | 2 |
| 0.005 | 0.10 | 3 |
| 0.005 | 0.20 | 4 |
| 0.01 | 0.01 | 5 |
| 0.01 | 0.05 | 6 |
| 0.01 | 0.10 | 7 |
| 0.01 | 0.20 | 8 |
| 0.1 | 0.01 | 9 |
| 0.1 | 0.05 | 10 |
| 0.1 | 0.10 | 11 |
| 0.1 | 0.20 | 12 |

Figure 2(c)

```
Rscript /gpfs/research/chongwu/zichenzhang/SUMMIT-test/code/Simulation.R \  
--h2_e A \  
--h2_p 0.2 \  
--p_causal 0.05 \  
--n B \  
--sumstats TRUE \  
--gene_ENSG ENSG00000258289 \  
--UKB TRUE \  
--folder_output SIM \  
--t1e FALSE \  
--seed SEED \  

```

| A | B | SEED |
|-------|-------|------|
| 0.005 | 300 | 13 |
| 0.005 | 600 | 14 |
| 0.005 | 3000 | 15 |
| 0.005 | 10000 | 16 |
| 0.005 | 31684 | 17 |
| 0.01 | 300 | 18 |
| 0.01 | 600 | 19 |
| 0.01 | 3000 | 20 |
| 0.01 | 10000 | 21 |
| 0.01 | 31684 | 22 |
| 0.1 | 300 | 23 |
| 0.1 | 600 | 24 |
| 0.1 | 3000 | 25 |
| 0.1 | 10000 | 26 |
| 0.1 | 31684 | 27 |

Supplementary Figure 1

```
Rscript /gpfs/research/chongwu/zichenzhang/SUMMIT-test/code/Simulation.R \  
--h2_e A \  
--h2_p 0.1 \  
--p_causal B \  
--n 31684 \  
--sumstats TRUE \  
--gene_ENSG ENSG00000258289 \  
--UKB TRUE \  
--folder_output SIM \  
--t1e FALSE \  
--seed SEED \
```

To reproduce this figure, use exact same options for Figure 2(a) and 2(b) but change the flag `p_causal` to 0.1, 0.5, and 0.8.

Supplementary Figure 2

```
Rscript /gpfs/research/chongwu/zichenzhang/SUMMIT-test/code/Simulation.R \  
--h2_e 0.05 \  
--h2_p 0.2 \  
--p_causal 0.2 \  
--n A \  
--sumstats B \  
--gene_ENSG ENSG00000258289 \  
--UKB TRUE \  
--folder_output SIM \  
--t1e FALSE \  
--seed SEED \  

```

| A | B | SEED |
|-------|-------|------|
| 300 | FALSE | 28 |
| 600 | FALSE | 29 |
| 3000 | FALSE | 30 |
| 10000 | FALSE | 31 |
| 31684 | FALSE | 32 |
| 31684 | TRUE | 33 |

Supplementary Figure 3

```
Rscript /gpfs/research/chongwu/zichenzhang/SUMMIT-test/code/Simulation.R \  
--h2_e A \  
--h2_p 0.2 \  
--p_causal B \  
--n 31684 \  
--sumstats TRUE \  
--gene_ENSG ENSG00000003147 \  
--UKB TRUE \  
--folder_output SIM \  
--t1e FALSE \  
--seed SEED \  

```

| A | B | SEED |
|-------|------|------|
| 0.005 | 0.01 | 34 |
| 0.005 | 0.05 | 35 |
| 0.005 | 0.10 | 36 |
| 0.005 | 0.20 | 37 |
| 0.01 | 0.01 | 38 |
| 0.01 | 0.05 | 39 |
| 0.01 | 0.10 | 40 |
| 0.01 | 0.20 | 41 |
| 0.1 | 0.01 | 42 |
| 0.1 | 0.05 | 43 |
| 0.1 | 0.10 | 44 |
| 0.1 | 0.20 | 45 |

Supplementary Figure 4

```
Rscript /gpfs/research/chongwu/zichenzhang/SUMMIT-test/code/Simulation.R \  
--h2_e A \  
--h2_p 0.2 \  
--p_causal B \  
--n 31684 \  
--sumstats TRUE \  
--gene_ENSG ENSG0000001631 \  
--UKB TRUE \  
--folder_output SIM \  
--t1e FALSE \  
--seed SEED \  

```

| A | B | SEED |
|-------|------|------|
| 0.005 | 0.01 | 46 |
| 0.005 | 0.05 | 47 |
| 0.005 | 0.10 | 48 |
| 0.005 | 0.20 | 49 |
| 0.01 | 0.01 | 50 |
| 0.01 | 0.05 | 51 |
| 0.01 | 0.10 | 52 |
| 0.01 | 0.20 | 53 |
| 0.1 | 0.01 | 54 |
| 0.1 | 0.05 | 55 |
| 0.1 | 0.10 | 56 |
| 0.1 | 0.20 | 57 |

Supplementary Figure 5

```
Rscript /gpfs/research/chongwu/zichenzhang/SUMMIT-test/code/Simulation.R \  
--h2_e 0.05 \  
--h2_p 0.2 \  
--p_causal 0.2 \  
--n A \  
--sumstats B \  
--gene_ENSG ENSG0000003147 \  
--UKB TRUE \  
--folder_output SIM \  
--t1e FALSE \  
--seed SEED \  

```

| A | B | SEED |
|-------|-------|------|
| 300 | FALSE | 58 |
| 600 | FALSE | 59 |
| 3000 | FALSE | 60 |
| 10000 | FALSE | 61 |
| 31684 | FALSE | 62 |
| 31684 | TRUE | 63 |

Supplementary Figure 6

```
Rscript /gpfs/research/chongwu/zichenzhang/SUMMIT-test/code/Simulation.R \
--h2_e 0.05 \
--h2_p 0.2 \
--p_causal 0.2 \
--n A \
--sumstats B \
--gene_ENSG ENSG0000001631 \
--UKB TRUE \
--folder_output SIM \
--t1e FALSE \
--seed SEED \
```

| A | B | SEED |
|-------|-------|------|
| 300 | FALSE | 64 |
| 600 | FALSE | 65 |
| 3000 | FALSE | 66 |
| 10000 | FALSE | 67 |
| 31684 | FALSE | 68 |
| 31684 | TRUE | 69 |

Supplementary Figure 7

```
Rscript /gpfs/research/chongwu/zichenzhang/SUMMIT-test/code/Simulation.R \  
--h2_e A \  
--h2_p 0.2 \  
--p_causal 0.01 \  
--n B \  
--sumstats FALSE \  
--gene_ENSG ENSG00000258289 \  
--UKB TRUE \  
--folder_output SIM \  
--t1e FALSE \  
--seed SEED \  

```

| A | B | SEED |
|-------|-------|------|
| 0.005 | 300 | 70 |
| 0.005 | 600 | 71 |
| 0.005 | 3000 | 72 |
| 0.005 | 10000 | 73 |
| 0.005 | 31684 | 74 |
| 0.01 | 300 | 75 |
| 0.01 | 600 | 76 |
| 0.01 | 3000 | 77 |
| 0.01 | 10000 | 78 |
| 0.01 | 31684 | 79 |
| 0.05 | 300 | 80 |
| 0.05 | 600 | 81 |
| 0.05 | 3000 | 82 |
| 0.05 | 10000 | 83 |
| 0.05 | 31684 | 84 |
| 0.1 | 300 | 85 |
| 0.1 | 600 | 86 |
| 0.1 | 3000 | 87 |

| A | B | SEED |
|-----|-------|------|
| 0.1 | 10000 | 88 |
| 0.1 | 31684 | 89 |

Supplementary Figure 8

```
Rscript /gpfs/research/chongwu/zichenzhang/SUMMIT-test/code/Simulation.R \  
--h2_e 0.05 \  
--h2_p 0.2 \  
--p_causal 0.01 \  
--n 31684 \  
--sumstats TRUE \  
--gene_ENSG ENSG00000258289 \  
--UKB TRUE \  
--folder_output SIM \  
--t1e TRUE \  
--seed 90 \
```

Real-data results

To reproduce SUMMIT's real-data results, please follow our steps:

1. Download GWAS summary data through the links provided in Supplementary Table 1
2. Preprocess the downloaded GWAS summary data using APSS.R
3. Get the real-data result using TestAssociation.R
4. Use our plotting code to draw the plots

Tips

Since we were working on a SLURM-managed environment, in line 55 of `Simulation.R`, we used

```
Sys.getenv("SLURM_ARRAY_TASK_ID")
```

to get the sub-job's index from the global environment. Depending on your computing environment, you may need to manually modify this line.

In line 663 of `Simulation.R`, you can change `runs` to a smaller number to save time and still get results that are very similar to ours.

Notes

We made every effort to help users replicate our results exactly. All results were tested on CentOS Linux 8. For each figure and table, you may need to change the parameters (denoted by A, B, and SEED) as we suggested. The codes can also serve as a launching point for your own studies.

All the above codes are for replication purposes and the users may need to change the directory and install the necessary packages to run it smoothly. If you have any questions, feel free to contact us (Zichen Zhang, zz17@fsu.edu)

Disclaimer

The codes are provided "as is" and the author disclaims all warranties with regard to these codes including all implied warranties of merchantability and fitness. In no event shall the author be liable for any special, direct, indirect, or consequential damages or any damages whatsoever resulting from loss of use, data or profits, whether in an action of contract, negligence or other tortious action, arising out of or in connection with the use or performance of these codes.

Author

Zichen Zhang, and Chong Wu