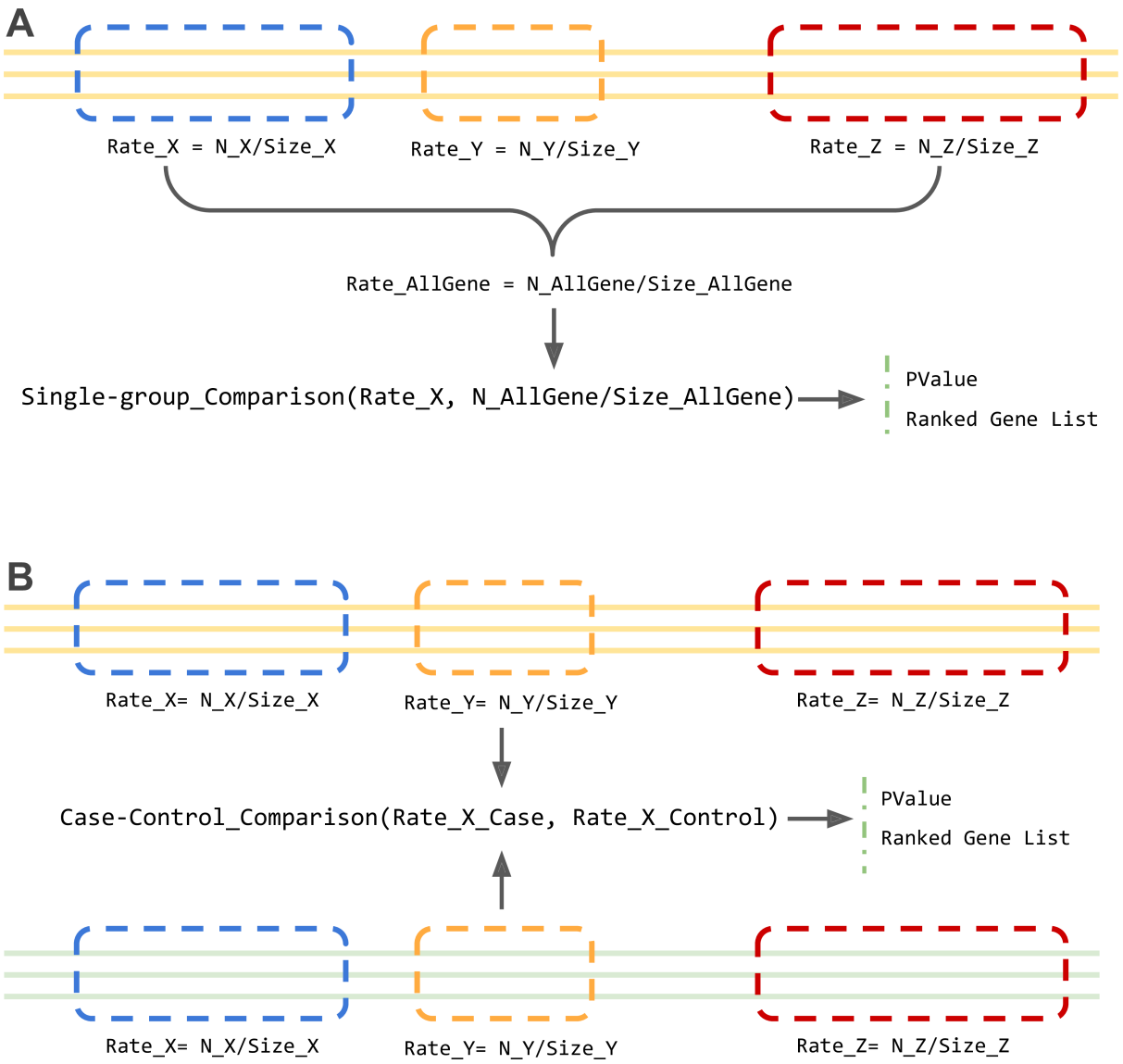


Figure S1: **Illustration figures for the "single-group" analysis (A) and "case-control" analysis (B).** $N\_X$ is the number of families with qualified rare variants in gene $X$. $Size\_X$ is the exome sequencing covered size for gene $X$. The binomial test was used to evaluate the enrichment of families with qualified rare variants in each gene.

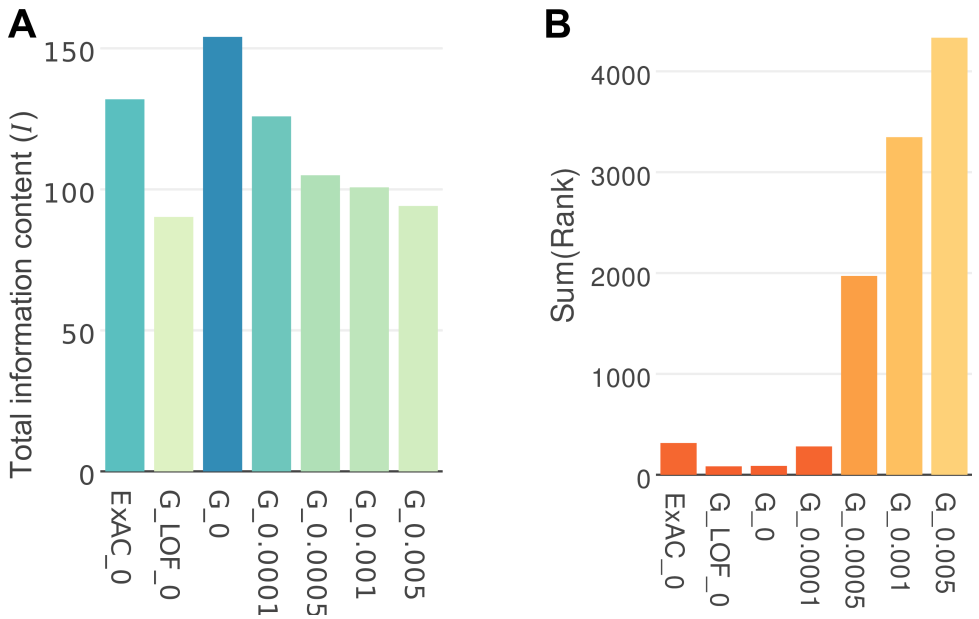


Figure S2: **"Single-group" burden test for extreme rare variants of dominant model, showing the statistics for known FSGS genes WT1, COL4A5, ACTN4, INF2, TRPC6, and PAX2.** A) The total information content (bits) of positive control genes for different parameter combinations. The total information content was measured by $I = -\sum log_2(P_i)$, $P_i$ is the $Pvalue$ of gene $i$. B) The overall rank sum of positive control genes for different parameter combinations.

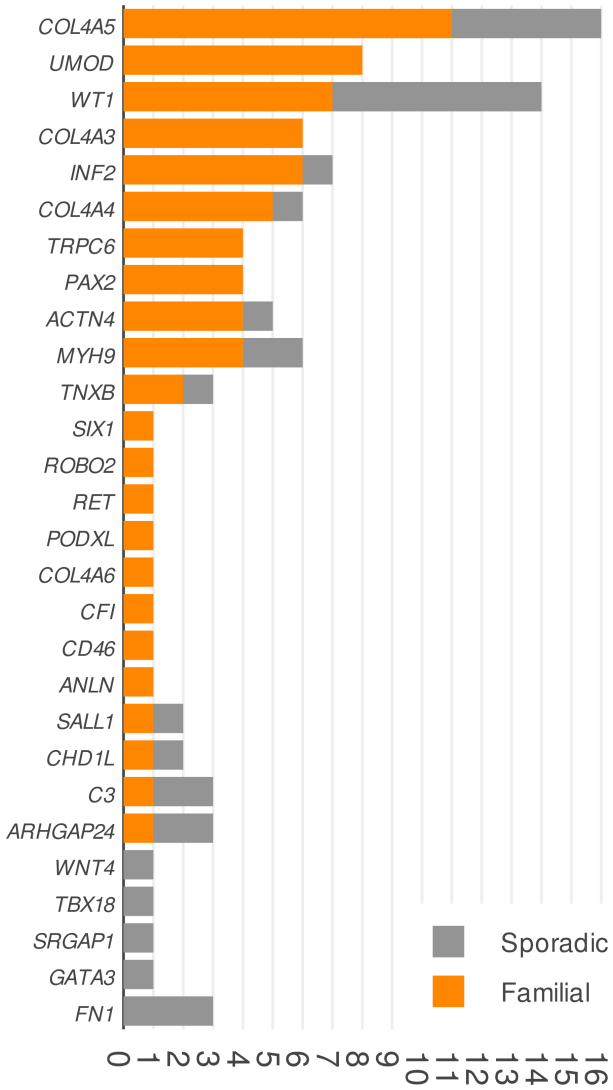


Figure S3: **The rare variant distribution in familial and sporadic families for each dominant disease-causing gene.**

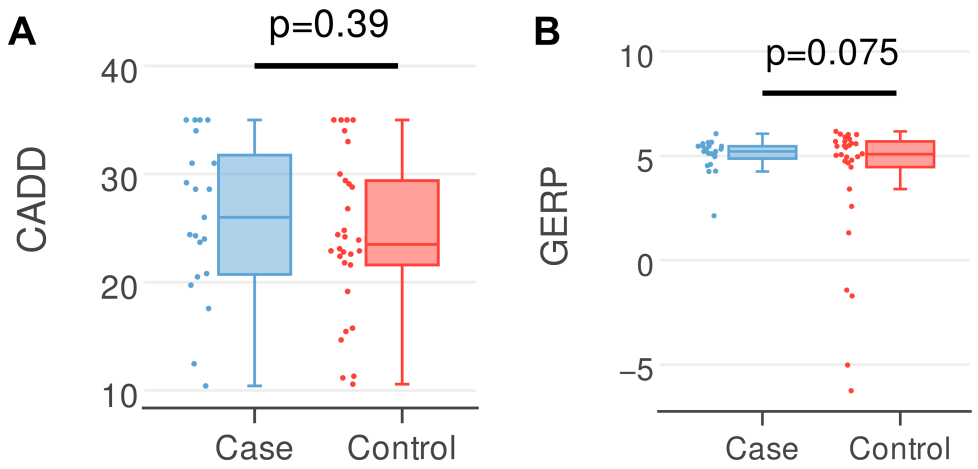


Figure S4: **The comparison of rare variant CADD and GERP score between the case and control samples for those dominant genes with high burden in control samples.** The CADD and GERP score of rare variants in MYH9, FN1, SALL1, PODXL, TBX18, ROBO2, COL4A6 and SRGAP1 gene were compared in cases and controls; no significant differences were detected, $Pvalue > 0.05$, $T$ test.

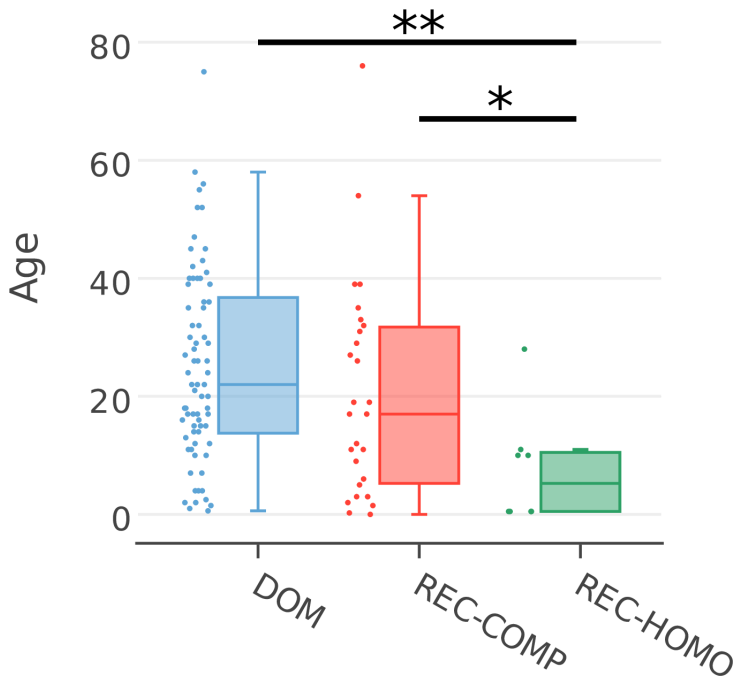


Figure S5: **Diagnosis age distribution for affected individuals with rare variants detected in known disease-causing genes (without genes with high level of burden in control samples), partitioned by inheritance model.**
* represents $Pvalue < 0.05, Utest$. ** represents $Pvalue < 0.005, Utest$.

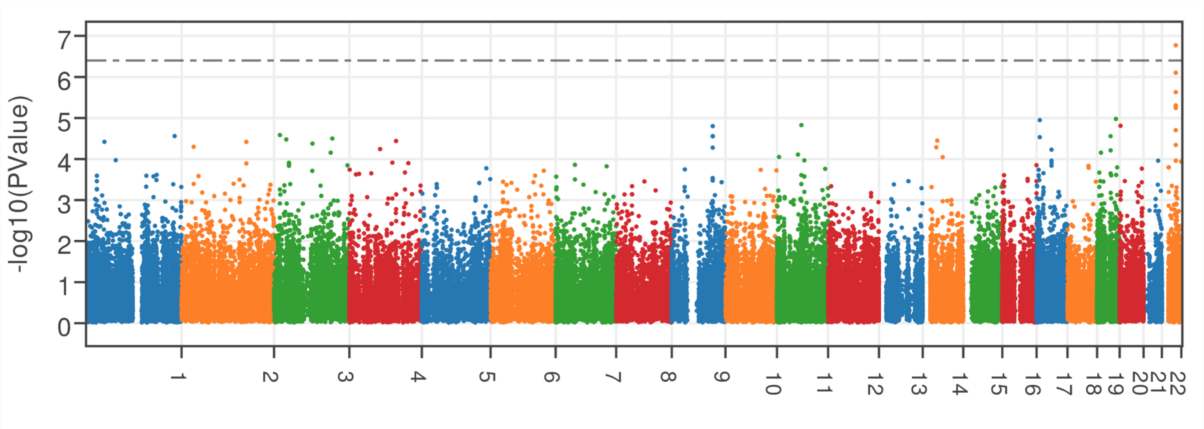


Figure S6: **Manhattan plot for the results of the association test of the exonic common variants.** Dashed line shows the genome-wide significance cut-off ($0.05/123424 = 4.0510^{-7}$, 123,424 is the total number of exonic common variants tested). The APOL1 G1 variant reached genome-wide significance.

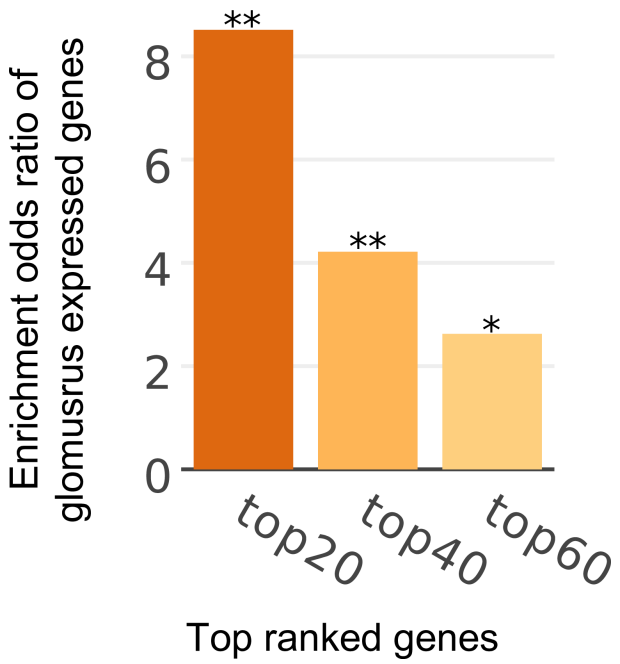


Figure S7: **The enrichment of renal glomerulus expression enriched genes in top ranked genes.**

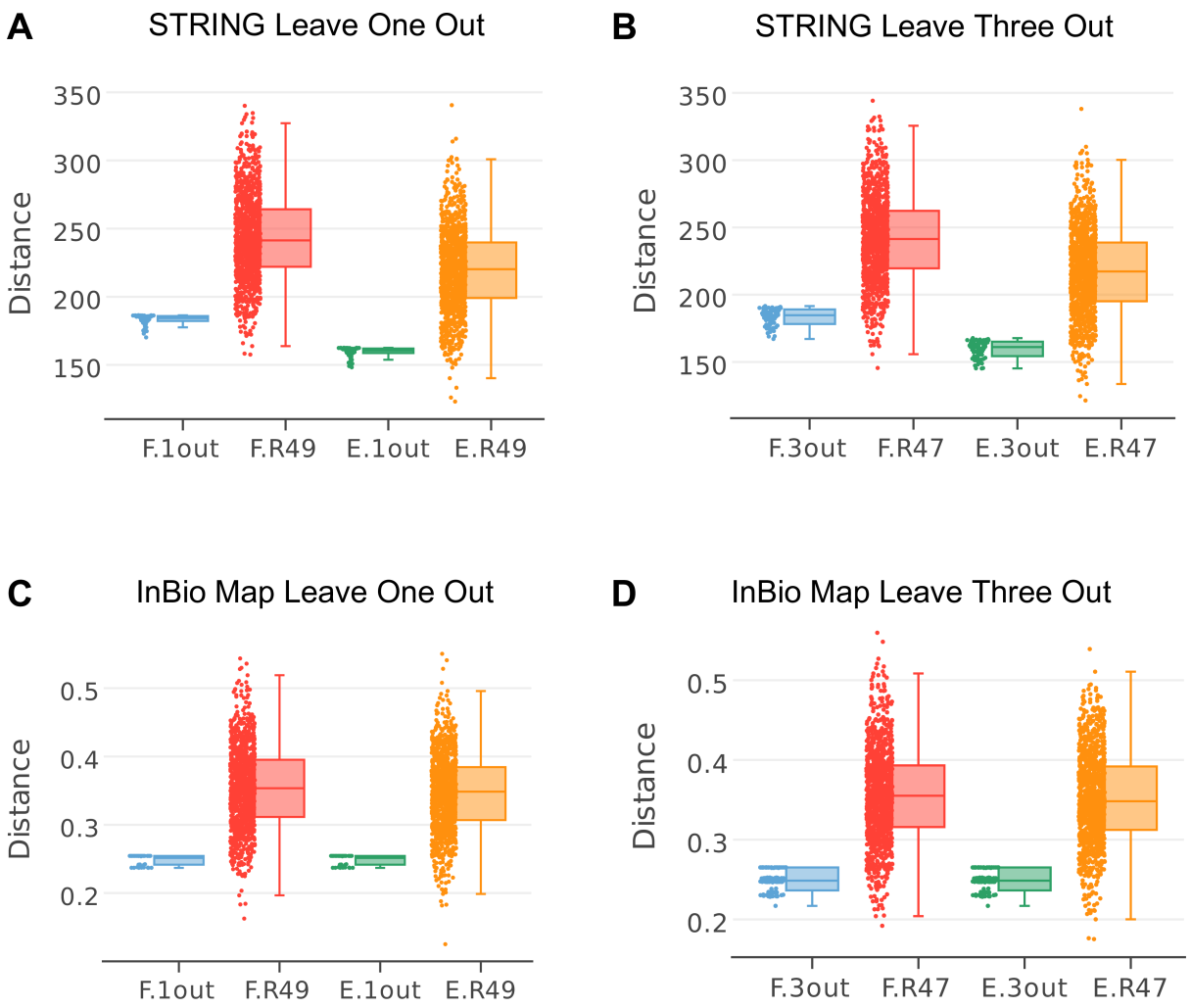


Figure S8: **Robustness analysis of gene set distance comparison.** Each dot is a distance between two gene sets. The F.1out and E.1out are the gene set distances between the top ranked 50 new genes (leaving one out) with the FSGS gene panel and the expanded gene panel, respectively. This process was repeated 50 times, each time dropping one gene in the iteration. F.R49 and E.R49 show the gene set distance between the 49 random picked genes with the FSGS gene panel and expanded gene panel, respectively. This process was repeated 1000 times. The F.3out and E.3out are the same analyses as "leave one out", except now each time dropping 3 randomly chosen genes, labeled as "leave three out". The random picking for the "leave three out" analysis was repeated 100 times. F.R47 and E.R47 are the sample analyses done such that each set of 47 genes were picked in random to perform a "leave three out" analysis.
A) and C) Leave one out analysis for STRING network and in Bio Map network, respectively.
B) and D) Leave three out analysis for STRING network and in Bio Map network, respectively.