

Appendix S1:

$$\begin{aligned}
 & \begin{bmatrix} q_1 q_1 & q_1 q_2 & \dots & q_1 q_N \\ q_2 q_1 & q_2 q_2 & \dots & q_2 q_N \\ \dots & \dots & \dots & \dots \\ q_N q_1 & q_N q_2 & \dots & q_N q_N \end{bmatrix} \\
 &= \begin{bmatrix} Q & Q^\perp \end{bmatrix} \Lambda \begin{bmatrix} Q & Q^\perp \end{bmatrix}^T \\
 &= \begin{bmatrix} q_1 & \dots & \dots & \dots \\ q_2 & \dots & \dots & \dots \\ \dots & \dots & Q^\perp & \dots \\ q_N & \dots & \dots & \dots \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 0 & & \\ & & \dots & \\ & & & 0 \end{bmatrix} \begin{bmatrix} q_1 & q_2 & \dots & q_N \\ & & \dots & \\ \dots & \dots & (Q^\perp)^T & \dots \\ & & \dots & \end{bmatrix},
 \end{aligned}$$

where Q^\perp is matrix with column vectors orthogonal to vector $Q = (q_1, q_2, \dots, q_N)^T$, and Λ is diagonal matrix with all element equals 0 except that the first element is 1.

Table S1: simulation setting for models without confounding effect and a fixed sample size

Effect	Pct*	Setting*							
		Binary		Normal		Student's t		Cauchy	
		μ_β	σ_β	μ_β	σ_β	μ_β	σ_β	μ_β	σ_β
Null	0	0	0	0	0	0	0	0	0
	5								
	10								
	30	0	1	0	0.3	0	0.5	0	0.5
A1	50								
	5								
	10								
	30	0.25	0.25	0.1	0.1	0.25	0.25	0.25	0.25
A2	50								

* Percentage of functional rare variants

Table S2: simulation setting for models without confounding effect and varying sample sizes

Effect	Sample size	Setting							
		Binary		Normal		Student's t		Cauchy	
		μ_β	σ_β	μ_β	σ_β	μ_β	σ_β	μ_β	σ_β
Null	50								
	100								
	200	0	0	0	0	0	0	0	0
	500								
A1*	50								
	100								
	200	0	1	0	0.3	0	0.5	0	0.5
	500								

* 50% of the rare variants are set as functional.

Table S3: simulation setting for models with cofounding effect and varying sample sizes

Effect	Sample size	Setting							
		Binary		Normal		Student's t		Cauchy	
		μ_β	σ_β	μ_β	σ_β	μ_β	σ_β	μ_β	σ_β
Null	50								
	100	0	0	0	0	0	0	0	0
	200								
	500								
A1*	50								
	100	0	1	0	0.3	0	1	0	0.5
	200								
	500								

* 50% of the rare variants are set as functional.

Table S4: Type I error rates for WU-SEQ based on 100000 simulated replicates

Level	Method	Type I error*			
		Binary	Normal	Student's t	Cauchy
0.05	WU-SEQ _{RK} **	0.0295	0.0379	0.0379	0.0382
	WU-SEQ _{QT} ***	0.0295	0.0495	0.0495	0.0499
5×10^{-4}	WU-SEQ _{RK} **	1×10^{-05}	6×10^{-05}	7×10^{-05}	1×10^{-04}
	WU-SEQ _{QT} ***	1×10^{-05}	0.00028	0.00039	0.00049

*The sample size is 200 and other simulation settings are the same as in Table S2.

**In WU-SEQ_{RK}, we use cross product kernel based on rank of the phenotype values without quantile transformation.

***In WU-SEQ_{QT}, we use cross product kernel with quantile transformation.

Table S5: The association of 4 candidate genes with 3 binary phenotypes* in Dallas Heart Study

Gene	BMI		Cholesterol		VLDL	
	SKAT	WUSEQ	SKAT	WUSEQ	SKAT	WUSEQ
<i>ANGPTL3</i>	0.805**	0.791	0.734	0.655	0.024	0.051
<i>ANGPTL4</i>	0.340	0.365	0.242	0.263	0.012	0.020
<i>ANGPTL5</i>	0.750	0.585	0.752	0.648	0.310	0.413
<i>ANGPTL6</i>	0.698	0.49	0.006	0.006	0.286	0.375
<i>All 4 genes</i>	0.765	0.632	0.047	0.059	0.010	0.022

*For each phenotype, we use the highest quartiles and lowest quartiles to form the binary phenotype.

** P-value from the association analysis by adjusting for age, gender, and race.

Figure S1: Missing data distribution for Dallas Heart Study

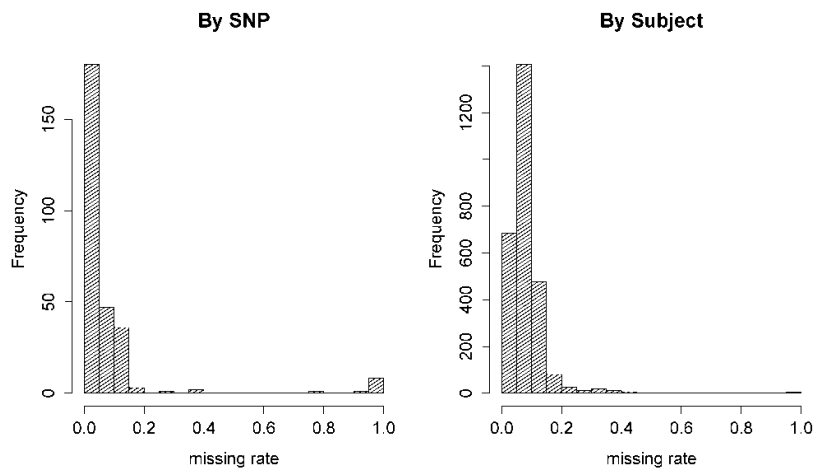


Figure S2: MAF for all SNPs

