

Supplemental Data

Meta-analysis of Gene-Level Associations

for Rare Variants Based on Single-Variant Statistics

Yi-Juan Hu, Sonja I. Berndt, Stefan Gustafsson, Andrea Ganna, Genetic Investigation of ANthropometric Traits (GIANT) Consortium, Joel Hirschhorn, Kari E. North, Erik Ingelsson, and Dan-Yu Lin

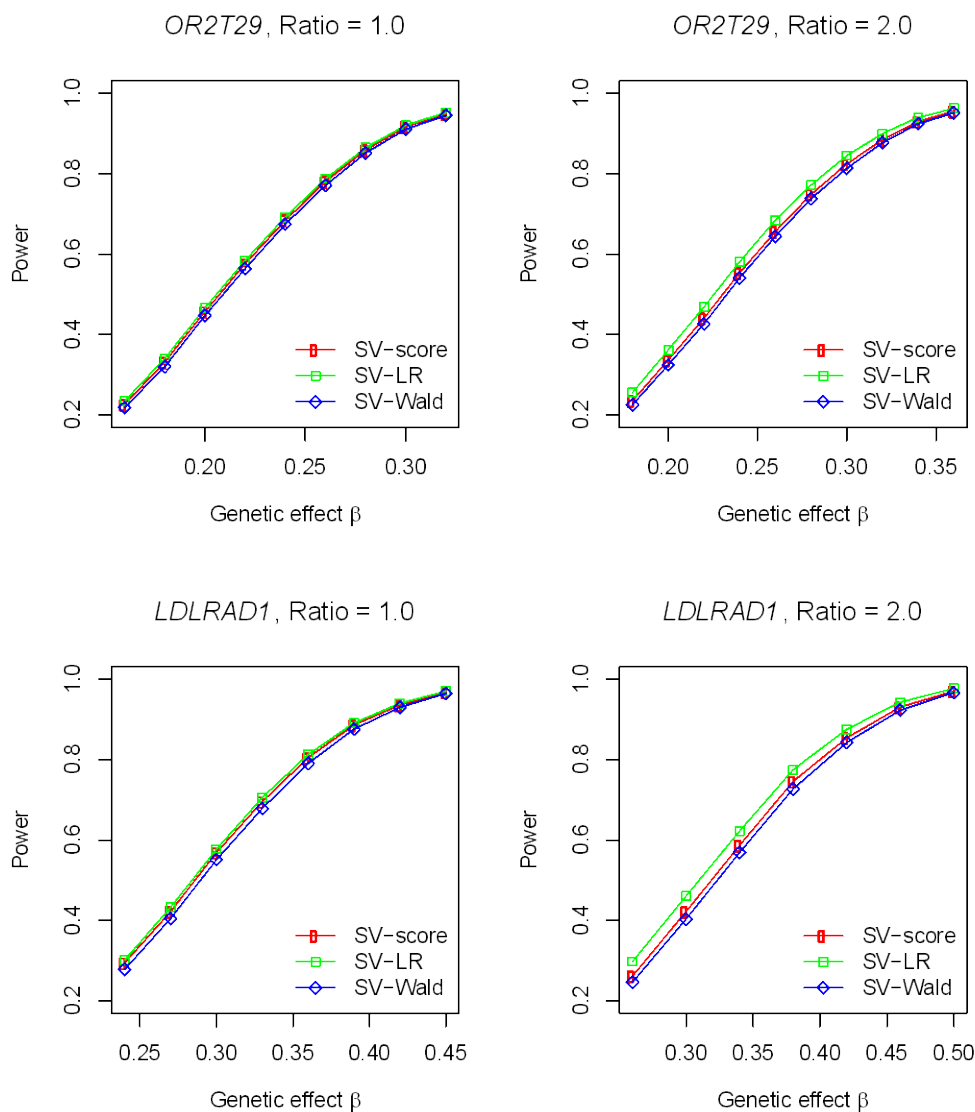


Figure S1. Power of T5 at the nominal significance level α of 0.001 for binary traits with an internal reference. Ratio is the case-control ratio for Study 2 and Study 3. SV-score, SV-LR and SV-Wald mean that single-variant p -values are based on the score, LR and Wald tests, respectively. Each power estimate is based on 10,000 replicates.

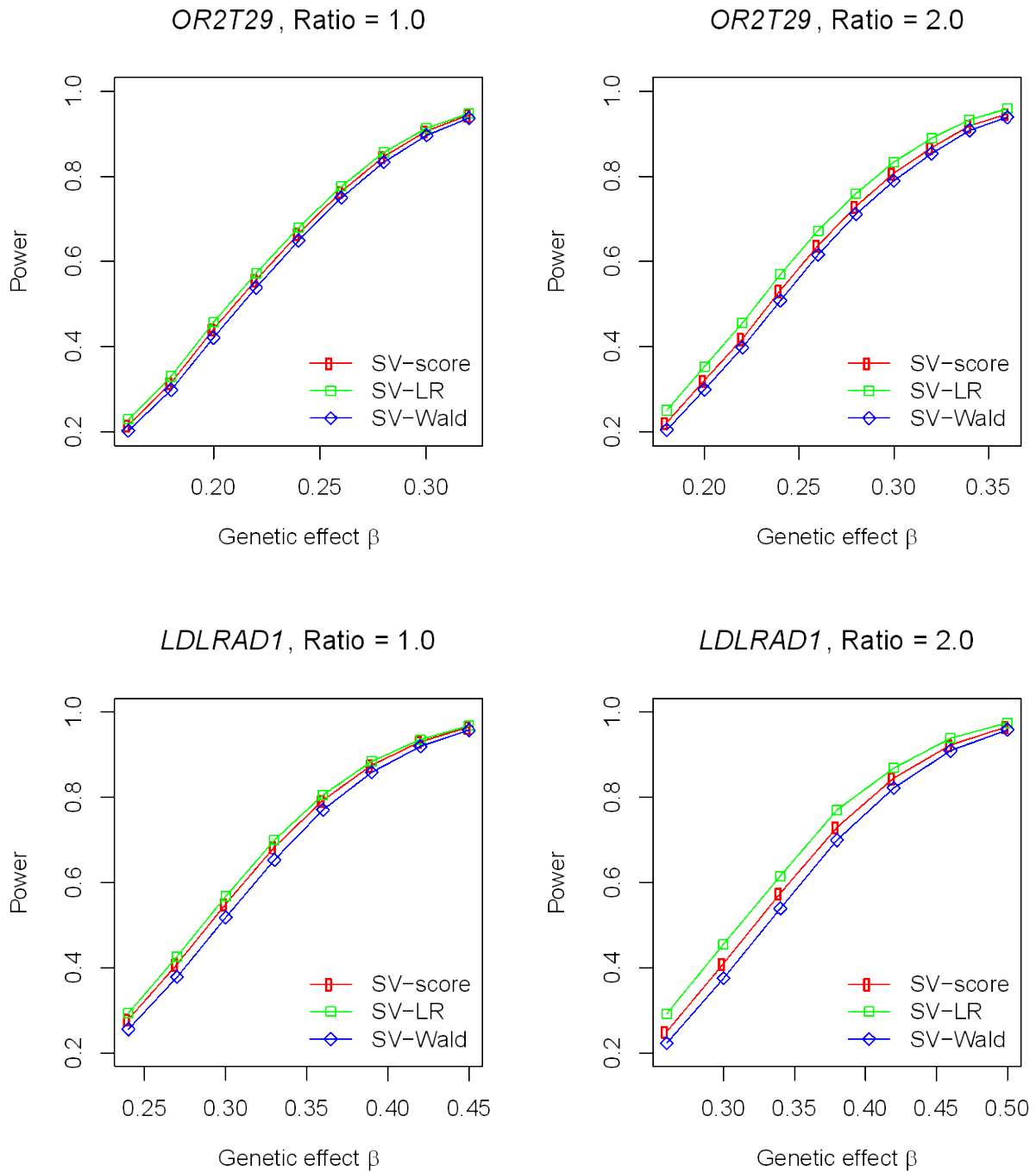


Figure S2. Power of T5 at the nominal significance level α of 0.001 for binary traits with an external reference. Ratio is the case-control ratio for Study 2 and Study 3. SV-score, SV-LR and SV-Wald mean that single-variant p -values are based on the score, LR and Wald tests, respectively. Each power estimate is based on 10,000 replicates.

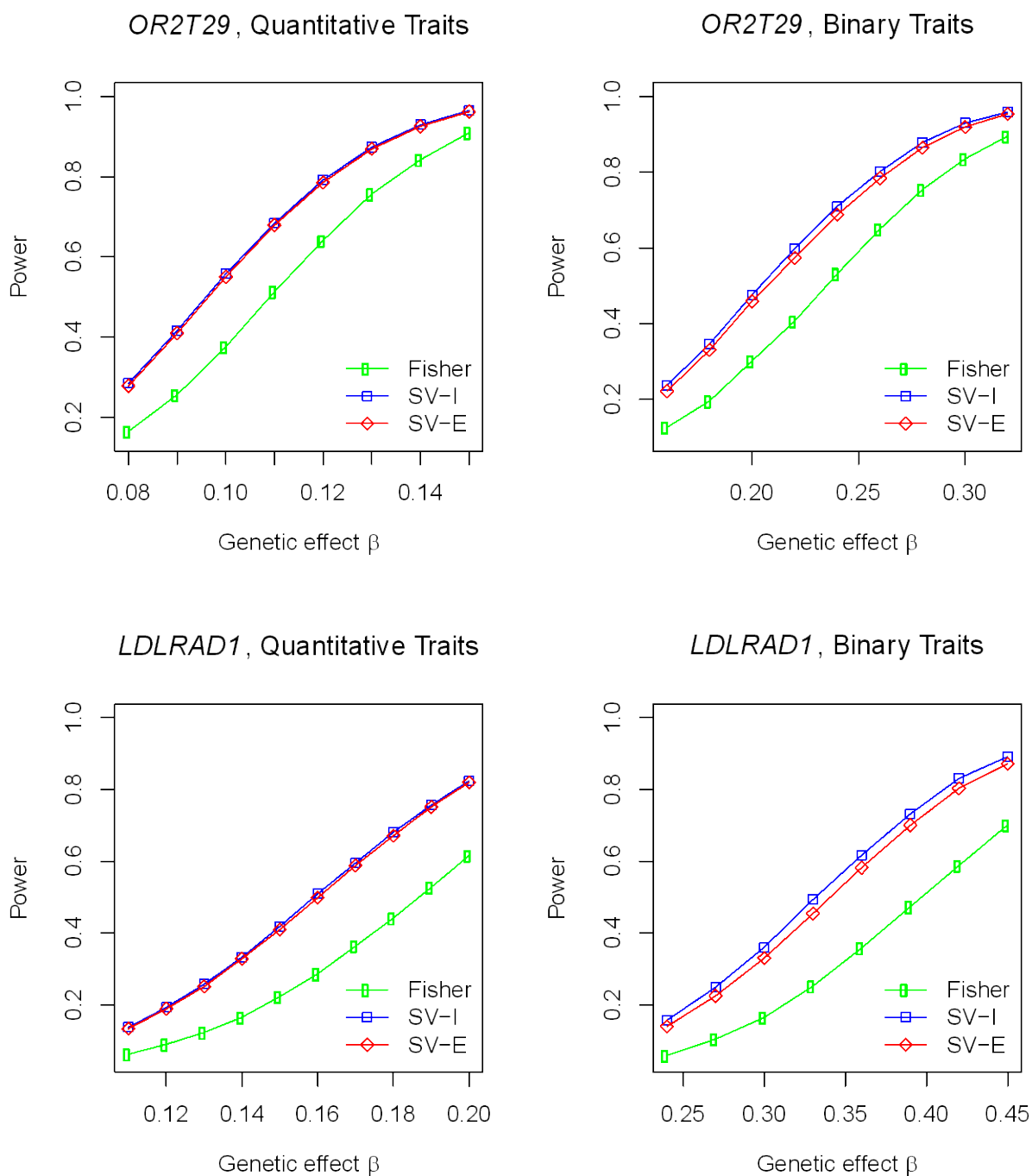


Figure S3. Power of SKAT for the proposed and Fisher's methods at the nominal significance level α of 0.001. For quantitative traits, the error variances in Study 2 and Study 3 are 1. For binary traits, the case-control ratios in Study 2 and Study 3 are 1. For SV-I and SV-E, single-variant p -values are based on the score test. Each power estimate is based on 10,000 replicates.

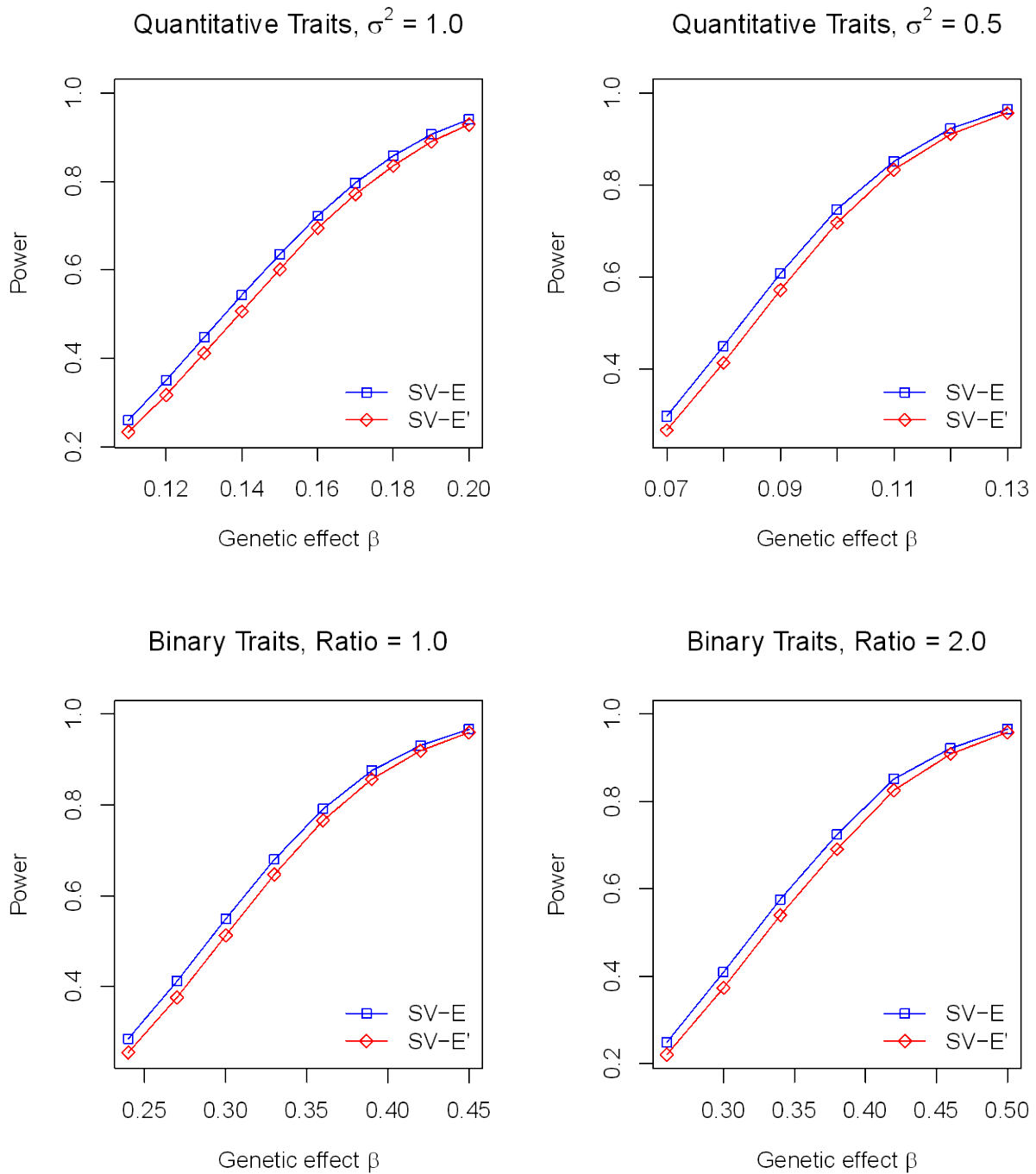


Figure S4. Power of T5 for *LDLRAD1* at the nominal significance level α of 0.001: SV-E and SV-E' pertain to correct and incorrect specifications of the external panel. For quantitative traits, σ^2 is the error variance of Study 2 and Study 3. For binary traits, "Ratio" is the case-control ratio of Study 2 and Study 3. Single-variant p -values are based on the score test. Each power estimate is based on 10,000 replicates.

Table S1. Type I error (divided by the nominal significance level) for quantitative traits

Gene	Test	σ^2	SV-I		SV-E		naive	
			score	LR	score	LR	score	LR
<i>OR2T29</i>	T5	0.5	0.98	0.99	0.96	0.97	75.67	75.90
		1.0	1.00	1.01	0.97	0.98	75.54	75.77
	VT	0.5	1.00	1.01	0.97	0.98	53.53	53.75
		1.0	1.00	1.00	1.01	1.02	53.34	53.53
	SKAT	0.5	0.95	0.97	0.94	0.96	19.19	19.33
		1.0	1.01	1.01	0.98	1.00	19.13	19.24
<i>LDLRAD1</i>	T5	0.5	0.98	0.99	0.97	0.98	2.87	2.90
		1.0	0.98	0.98	0.97	0.98	2.81	2.85
	VT	0.5	1.04	1.04	1.02	1.03	2.92	2.95
		1.0	1.05	1.06	1.01	1.02	2.98	3.00
	SKAT	0.5	1.02	1.03	1.01	1.03	2.39	2.44
		1.0	1.00	1.01	1.03	1.05	2.38	2.41

Note: The summary statistics do not include the standard error estimates. The nominal significance level $\alpha = 0.001$. σ^2 pertains to the error variance in Study 2 and Study 3. Each entry is based on 1,000,000 replicates.

Table S2. Type I error (divided by the nominal significance level) for binary traits

Gene	Test	Ratio	SV-I		SV-E		naive	
			score	LR	score	LR	score	LR
<i>OR2T29</i>	T5	1.0	0.88	0.94	0.86	0.92	73.13	74.75
		2.0	0.99	1.00	0.93	0.97	73.60	73.93
	VT	1.0	0.89	0.95	0.83	0.92	51.10	52.54
		2.0	1.02	1.01	0.98	0.98	51.91	52.14
	SKAT	1.0	0.90	0.96	0.84	0.92	17.65	18.64
		2.0	0.98	0.98	0.92	0.95	18.16	18.45
<i>LDLRAD1</i>	T5	1.0	0.83	0.89	0.76	0.86	2.18	2.46
		2.0	1.06	1.00	0.99	0.96	3.12	2.81
	VT	1.0	0.76	0.85	0.64	0.77	2.04	2.46
		2.0	1.07	0.98	0.88	0.83	3.20	2.82
	SKAT	1.0	0.83	0.96	0.70	0.85	1.58	1.92
		2.0	0.93	0.94	0.75	0.82	1.95	1.98

Note: The summary statistics do not include the standard error estimates. The nominal significance level $\alpha = 0.001$. Ratio is the case-control ratio in Study 2 and Study 3. Each entry is based on 1,000,000 replicates.

Table S3. Correlation matrix of the *LDLRADI* genotypes in ESP HeartGO data

	rs143619888	rs150468103	rs141759859	rs149768061	rs147345740	rs145889899	rs142900519	rs149114405
rs143619888	1	0	0.003	0.014	0	0.011	0	0
rs150468103		1	0	0	0	0	0	0
rs141759859			1	0.004	0	0.003	0.001	0
rs149768061				1	0	0.780	0.004	0
rs147345740					1	0	0	0
rs145889899						1	0.003	0
rs142900519							1	0
rs149114405								1

Table S4. Type I error (divided by the nominal significance level) of SV-E' for testing *LDLRADI*

Test	Quantitative traits				Binary traits			
	σ^2	score	LR	Wald	Ratio	score	LR	Wald
T5	1.0	0.74	0.74	0.75	1.0	0.52	0.60	0.39
	0.5	0.72	0.72	0.74	2.0	0.64	0.63	0.50
VT	1.0	0.90	0.91	0.92	1.0	0.48	0.57	0.32
	0.5	0.89	0.90	0.90	2.0	0.71	0.67	0.48
SKAT	1.0	0.90	0.91	0.92	1.0	0.64	0.74	0.46
	0.5	0.89	0.90	0.91	2.0	0.60	0.68	0.45

Note: The nominal significance level $\alpha = 0.001$. For quantitative traits, σ^2 is the error variance of Study 2 and Study 3. For binary traits, "Ratio" is the case-control ratio of Study 2 and Study 3. Each entry is based on 1,000,000 replicates.