Supplementary Materials
Powerful Multi-marker Association Tests: Unifying Genomic Distance-Based Regression and Logistic Regression

Fang Han, Wei Pan

Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN 55455

June 8, 2010

Correspondence author: Wei Pan

Telephone: (612) 626-2705

Fax: (612) 626-0660

Email: weip@biostat.umn.edu

Address: Division of Biostatistics, MMC 303,

School of Public Health, University of Minnesota,

Minneapolis, Minnesota 55455–0392, U.S.A.

Logistic regression models and tests

We aim to incorporate genotype scores, HWD parameters, LD measurements, and dissimilarity-derived scores, and possibly their various combinations, as covariates into logistic models, then characterize the Type I error and power properties. In particular, we would like to assess whether such an expanded model and its associated tests (e.g. SSU) can maintain high power by combining multiple types of information in genotypic distributional differences between the case and control groups.

Suppose that X is an $n \times k$ genotype matrix with the dosage coding; that is, $X_{il} = 0$, 1 or 2 represents the copy number of an allele in locus l for subject i. We denote XX as the cross-product matrix with the ith row as $(XX)_{i.} = (X_{i1}^2, X_{i1}X_{i2}, ..., X_{i1}X_{ik}, X_{i2}^2, X_{i2}X_{i3}, ..., X_{ik}^2)$. Suppose Z_G, Z_{H_1} and Z_{H_2} are matrices derived from similarity matrices S^G, S^{H_1} and S^{H_2} respectively. We will consider the following logistic regression models:

L1: Logit $\Pr(Y = 1) = \beta_0 + X\beta_1$,

L2: Logit $\Pr(Y = 1) = \beta_0 + X\beta_1 + XX\beta_2$,

L3: Logit $\Pr(Y = 1) = \beta_0 + X\beta_1 + XX\beta_2 + Z_G\beta_3 + Z_{H_1}\beta_4 + Z_{H_2}\beta_5$,

L4: Logit $\Pr(Y = 1) = \beta_0 + X\beta_1 + Z_G\beta_3 + Z_{H_1}\beta_4 + Z_{H_2}\beta_5$,

L5: Logit $\Pr(Y = 1) = \beta_0 + X\beta_1 + Z_{H_2}\beta_5$,

corresponding to five null hypotheses:

$H_{0,1}$: $\beta_1 = 0$,

$H_{0,2}$: $\beta_1 = 0$ and $\beta_2 = 0$,

$H_{0,3}$: $\beta_1 = 0$, $\beta_2 = 0$, $\beta_3 = 0$, $\beta_4 = 0$, and $\beta_5 = 0$,

$H_{0,4}$: $\beta_1 = 0$, $\beta_3 = 0$, $\beta_4 = 0$, and $\beta_5 = 0$,

$H_{0,5}$: $\beta_1 = 0$ and $\beta_5 = 0$.

Model L1, perhaps the most popular one in use, aims to detect the mean difference between the genotypes scores of the case and control groups. In addition to the mean difference in genotype scores, model L2 incorporates the possible differences in

HWD parameters (through the quadratic terms of the genotype scores) and in LD patterns (through the pairwise cross-products or interactions of the SNPs) (Kim et al 2009). Model L3 aims to capture all four types of information: mean genotype scores, HWD parameters, pairwise or two-way genotype score interactions, and high-order interactions. Note that a dissimilarity matrix is regarded as representing some complex high-order interactions among the SNPs. To reduce the number of parameters, we also consider a more parsimoneous model L4 by eliminating the high-dimensional XX from model L3. Since there may be some overlapping information in the use of the three dissimilarity matrices while H2 performed best as shown by Lin and Schaid (2009), a further simplified model is L5.

There are two ways to test each composite null hypothesis. The first is *joint testing*: we test all the parameters simultaneously by applying a joint test, such as the score, SSU or UminP test. The second is *combining p-values* of a test being applied to multiple components of the entire parameter vector. For example, for $H_{0,2}$, the joint testing is to test its two components $\beta_1 = 0$ and $\beta_2 = 0$ at the same time by applying a score test, a SSU test or a UminP test to model L2. As a comparison, the second way is a two-step procedure: we first obtain a p-value for each of the two components, β_1 and β_2, using any of the above tests, then combine the two p-values. An advantage of the joint testing is the availability of asymptotic null distributions to facilitate p-value calculations, in contrast to the use of permutations or simulations to obtain p-values in combining p-values. A weakness of joint testing is possible loss of power due to large DF, which may (or may not) be overcome by combining p-values.

Since the SSU will be shown to perform best, we only consider combining p-values from the SSU tests. Given L p-values, p_1,...,p_L, obtained from L SSU tests on individual components of a null hypothesis, we consider three popular methods of combining the p-values:

- The MinP method: $T_{\text{MinP}} = \min(p_1, ..., p_L)$.

- Fisher's (1932) method: $T_{\text{Fisher}} = \prod_{j=1}^{L} p_j$.
- The truncated product method (TPM): $T_{\text{TPM}} = \prod_{j=1}^{L} p_j I(p_j < \tau)$, where τ is some cut-off; as in Zaykin et al (2002), we used $\tau = \alpha = 0.05$ throughout.

To obtain a p-value for each combining function, say $C(p_1, ..., p_L)$, we can use permutations by shuffling Y, which however is computationally demanding for its requirement of fitting models many times. Here we propose using a simulation based approach. First, we note that each individual test is based on a component of or the whole score vector U. Second, because of the asymptotic null distribution of U is known as $U \sim N(0, V)$, we can simulate B iid copies of U^b's from $N(0, V)$ with $b = 1, 2..., B$. Based on each U^b, we can calculate individual p-values as $p_1^b, ..., p_L^b$, and thus $C(p_1^b, ..., p_L^b)$. Third, the p-value for $C(p_1, ..., p_L)$ is simply $\sum_{b=1}^{B} I[C(p_1, ..., p_L) < C(p_1^b, ..., p_L^b)]/B$. We used $B = 1000$ for simulated data and $B = 1E6$ for the ALS data.

Simulation results

Figure 1 shows the Type I error rates of various methods.

The left panel of Fig 2 shows that the permutation-based F-test in GDBR with any similarity matrix (G or H1 or H2) had almost the same overall power as the SSU test in the corresponding logistic regression model (i.e. with the corresponding decomposed G or H1 or H2 matrix as predictors).

The middle panel of Fig 2 shows the overall power of the joint testing with the SSU, score and UminP tests applied to models L1-L5. It can be seen that the SSU test in model L4 had the highest power, though the SSU tests in models L1 and L5 has almost equally high power, closely followed by the UminP test applied to model L1. Overall, for any given model, the SSU test yielded the highest power, followed by UminP, and the score test performed badly. Again, as shown in Table 1, for a candidate region, it is possible that the UminP test was more powerful than the SSU test.

The right panel of Fig 3 shows the performance of the three methods of combining the p-values of the SSU test in models L2-L4. For the overall power, it is clear that Fisher's method performed best, followed by TPM and SSU. The minP method did not work well, possibly due to the fact that each type of information (contained in each component of a logistic regression model) contributed to disease-marker association, which also explains why the SSU outperformed the UminP in the middle panel. For Fisher's method, model L3 gave the highest power with all types of information being combined; in contrast, for the other two combining methods, the simpler model L5 yielded highest power.

In summary, in terms of the overall power, Fisher's method in model L3 was most powerful among all the tests, showing the power gain by combining multiple types of information.

We also conducted stratified power analysis, and similar conclusions can be drawn (Figs 3-5).

Results for ALS data

The results for all the nine SNPs are shown in Table 1.

Table 1: P-values of the single-marker and multi-marker tests for the ALS data.

							Logistic								
	Single-marker		LD blk	GDBR			G			H1			H2		
SNP	1-DF	2-DF	#SNPs	G	$H1$	$H2$	Score	SSU	UminP	Score	SSU	UminP	Score	SSU	UminP
rs4363506	3.64E-6	1.52E-6	3	< .001	< .001	< .001	1.64E-5	3.21E-6	2.44E-5	7.50E-5	3.82E-6	3.72E-5	0.0196	1.30E-5	0.0001
rs16984239	1.04E-5	1.36E-6	16	0.01	0.006	< .001	0.0003	0.0140	0.0015	7.16E-5	0.0110	5.20E-5	2.14E-5	1.00E-5	0.0002
rs12680546	0.0054	1.15E-5	11	< .001	< .001	< .001	0.0022	0.0039	0.0375	0.0019	0.0032	0.0356	0.0034	0.0025	0.0172
rs6013382	0.0096	6.31E-6	5	0.007	0.009	0.015	0.0151	0.0121	0.0430	0.0181	0.0106	0.0369	0.0195	0.0124	0.0270
rs2782931	0.3100	6.04E-6	19	0.513	0.534	0.708	0.0390	0.5464	0.0387	0.0479	0.5373	0.0116	0.6271	0.7420	0.2464
rs7976059	0.0008	6.56E-5	4	0.018	0.022	0.007	0.0037	0.0161	0.0061	0.0051	0.0192	0.0079	0.0057	0.0060	0.0531
rs10773543	0.0003	3.55E-5	2	< .001	< .001	< .001	0.0004	2.46E-5	7.62E-5	0.0002	1.53E-5	5.14E-5	0.0001	3.13E-5	0.0001
rs332389	0.8101	5.46E-6	4	< .001	< .001	< .001	4.92E-5	2.12E-6	8.25E-6	5.46E-5	2.56E-6	9.46E-6	0.0002	0.0002	0.0019
rs2767584	0.2083	0.0001	7	0.406	0.371	0.715	0.0233	0.3917	0.0049	0.4479	0.5967	0.3587	0.8152	0.8644	0.6478

	Logistic: Joint testing														
	L1			L2			L3			L4			L5		
SNP	Score	SSU	UminP	Score	SSU	UminP	Score	SSU	UminP	Score	SSU	UminP	Score	SSU	UminP
rs4363506	3.90E-5	8.95E-6	1.57E-5	0.0034	6.12E-6	2.46E-6	0.0217	2.38E-6	1.09E-5	0.0217	1.99E-6	1.45E-5	0.0217	3.34E-6	1.02E-5
rs16984239	0.0124	0.0256	1.27E-5	0.0250	0.0106	1.45E-5	0.0279	0.0098	0.0002	0.0312	0.0128	7.09E-6	0.0208	0.0133	8.93E-6
rs12680546	0.0026	3.90E-5	2.17E-5	0.0841	0.0001	0.0024	0.0796	0.0001	0.0003	0.0187	4.43E-5	3.62E-5	0.0038	4.17E-5	7.33E-5
rs6013382	0.0031	0.0003	0.0002	0.0563	0.0003	0.0028	0.0442	0.0003	0.0017	0.0402	0.0003	0.0007	0.0115	0.0002	0.0004
rs2782931	0.1039	0.2543	0.0005	0.3178	0.2679	0.2854	0.3178	0.2706	0.0062	0.1510	0.2762	0.0014	0.1768	0.2719	0.0014
rs7976059	0.0012	0.0303	0.0190	0.0250	0.0772	0.0019	0.0142	0.0441	0.0390	0.0098	0.0150	0.0218	0.0012	0.0175	0.0437
rs10773543	5.98E-5	2.98E-6	2.23E-5	0.0006	2.49E-6	9.14E-6	0.0003	6.06E-7	0.0002	0.0003	8.63E-7	0.0002	0.0001	1.32E-6	4.71E-5
rs332389	0.0011	1.74E-6	1.01E-5	0.0281	2.67E-6	3.86E-5	0.0184	2.61E-6	3.15E-5	0.0056	1.68E-6	2.37E-5	0.0095	1.78E-6	2.38E-5
rs2767584	0.0025	0.0004	0.0005	0.0564	0.0005	0.0091	0.0458	0.0006	0.0008	0.0166	0.0008	0.0008	0.0056	0.0006	0.0008

	Logistic: Combining SSU p-values											
	L2			L3			L4			L5		
SNP	minP	Fisher	TPM	minP	Fisher	TPM	minP	Fisher	TPM	minP	Fisher	TPM
rs4363506	3.30E-5	<1E-6	<1E-6	3.00E-5	<1E-6	<1E-6	2.20E-5	<1E-6	<1E-6	2.60E-5	<1E-6	<1E-6
rs16984239	0.0204	0.0027	0.0015	1.00E-4	<1E-6	<1E-6	7.00E-5	<1E-6	<1E-6	4.30E-5	1.30E-5	7.00E-6
rs12680546	0.0001	2.00E-6	2.00E-6	0.0003	<1E-6	<1E-6	0.0002	<1E-6	<1E-6	0.0001	5.00E-6	5.00E-6
rs6013382	0.0007	2.00E-6	1.00E-6	0.0017	<1E-6	<1E-6	0.0013	3.00E-6	2.00E-6	0.0006	5.90E-5	4.00E-5
rs2782931	0.4423	0.2502	0.0930	0.7686	0.5915	0.2188	0.6914	0.6752	0.1794	0.4454	0.5113	0.0933
rs7976059	0.0581	0.0416	0.0584	0.0300	8.80E-5	3.60E-5	0.0241	5.20E-5	1.50E-5	0.0120	0.0018	0.0010
rs10773543	7.00E-6	<1E-6	<1E-6	1.10E-5	<1E-6	<1E-6	8.00E-6	<1E-6	<1E-6	7.00E-6	<1E-6	<1E-6
rs332389	1.00E-5	<1E-6	<1E-6	1.90E-5	<1E-6	<1E-6	1.30E-5	<1E-6	<1E-6	9.00E-6	<1E-6	<1E-6
rs2767584	0.0011	1.10E-5	9.00E-6	0.0026	0.0003	0.0001	0.0020	0.0170	0.0086	0.0010	0.0036	0.0022

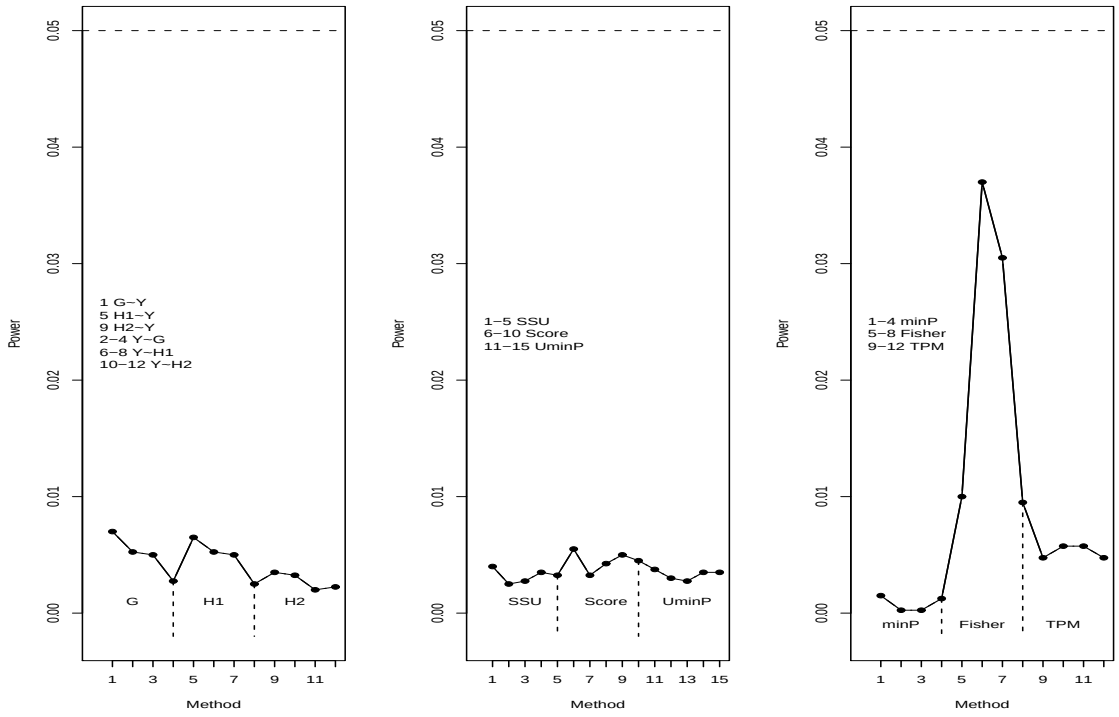

Figure 1: Empirical Type I error rates of various tests at the nominal level of 0.05 for simulated data. In the left panel, methods 2-4 (or 6-8 or 10-12) correspond to the SSU, multivariate score, and UminP tests respectively; in the middle panel, methods 1-5 (or 6-10 or 11-15) correspond to the SSU (or score or UminP) test applied to models L1-L5 respectively; in the right panel, methods 1-4 (or 5-8 or 9-12) correspond to the minP (or Fisher or TMP) for combining SSU p-values for models L2-L5 respectively.

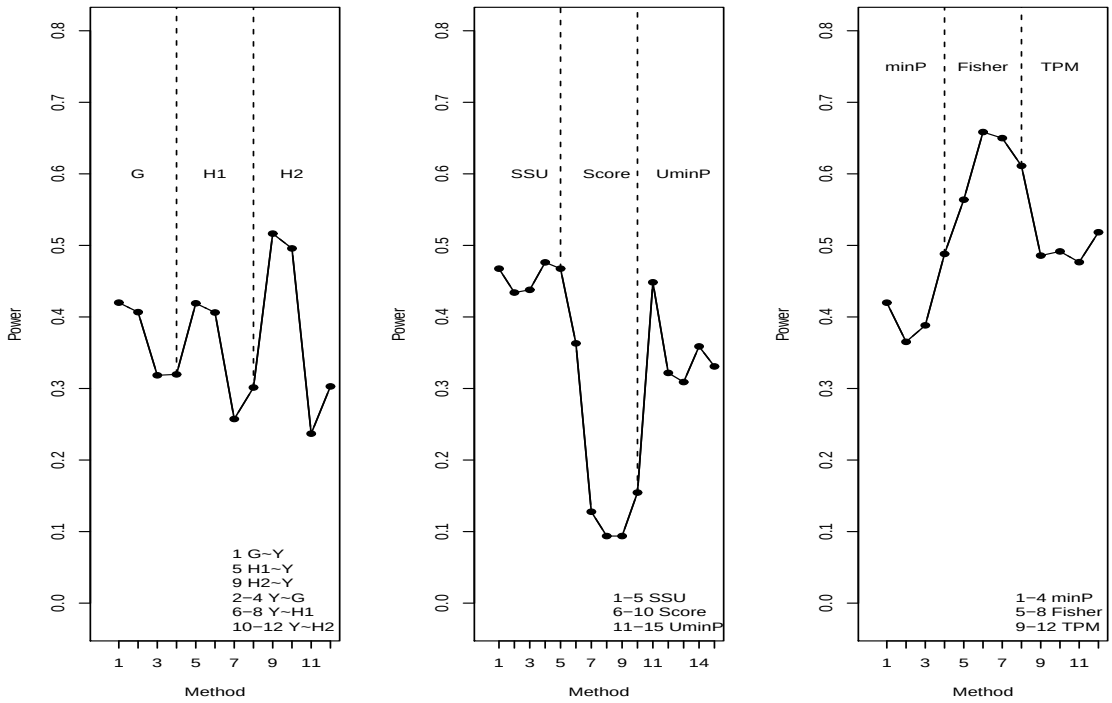

Figure 2: Empirical power of various tests from simulated data. In the left panel, methods 2-4 (or 6-8 or 10-12) correspond to the SSU, multivariate score, and UminP tests respectively; in the middle panel, methods 1-5 (or 6-10 or 11-15) correspond to the SSU (or score or UminP) test applied to models L1-L5 respectively; in the right panel, methods 1-4 (or 5-8 or 9-12) correspond to the minP (or Fisher or TMP) for combining SSU p-values for models L2-L5 respectively.

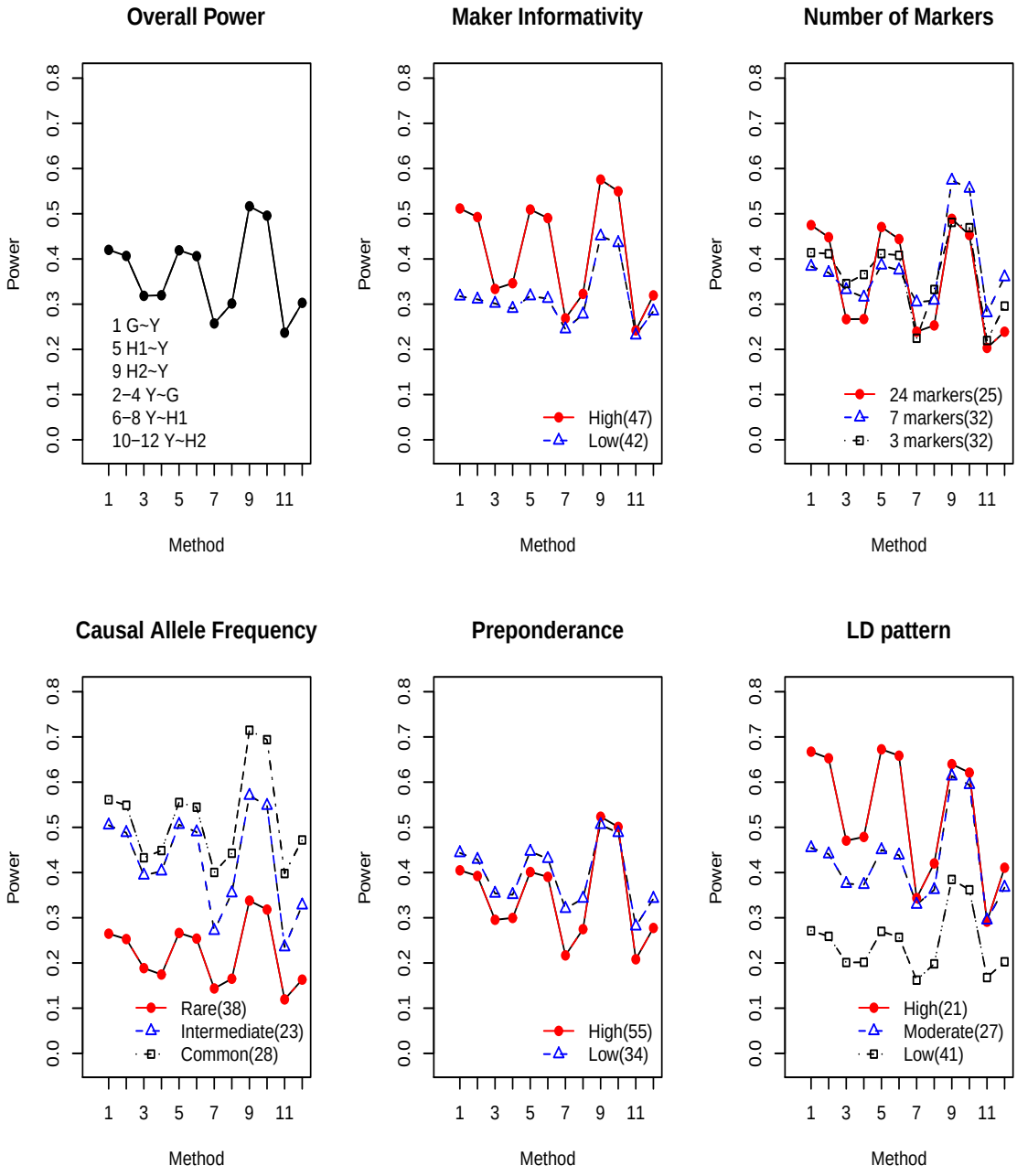

Figure 3: Empirical power of various tests from simulated data. Methods 2-4 (or 6-8 or 10-12) correspond to the SSU, multivariate score, and UminP tests respectively.

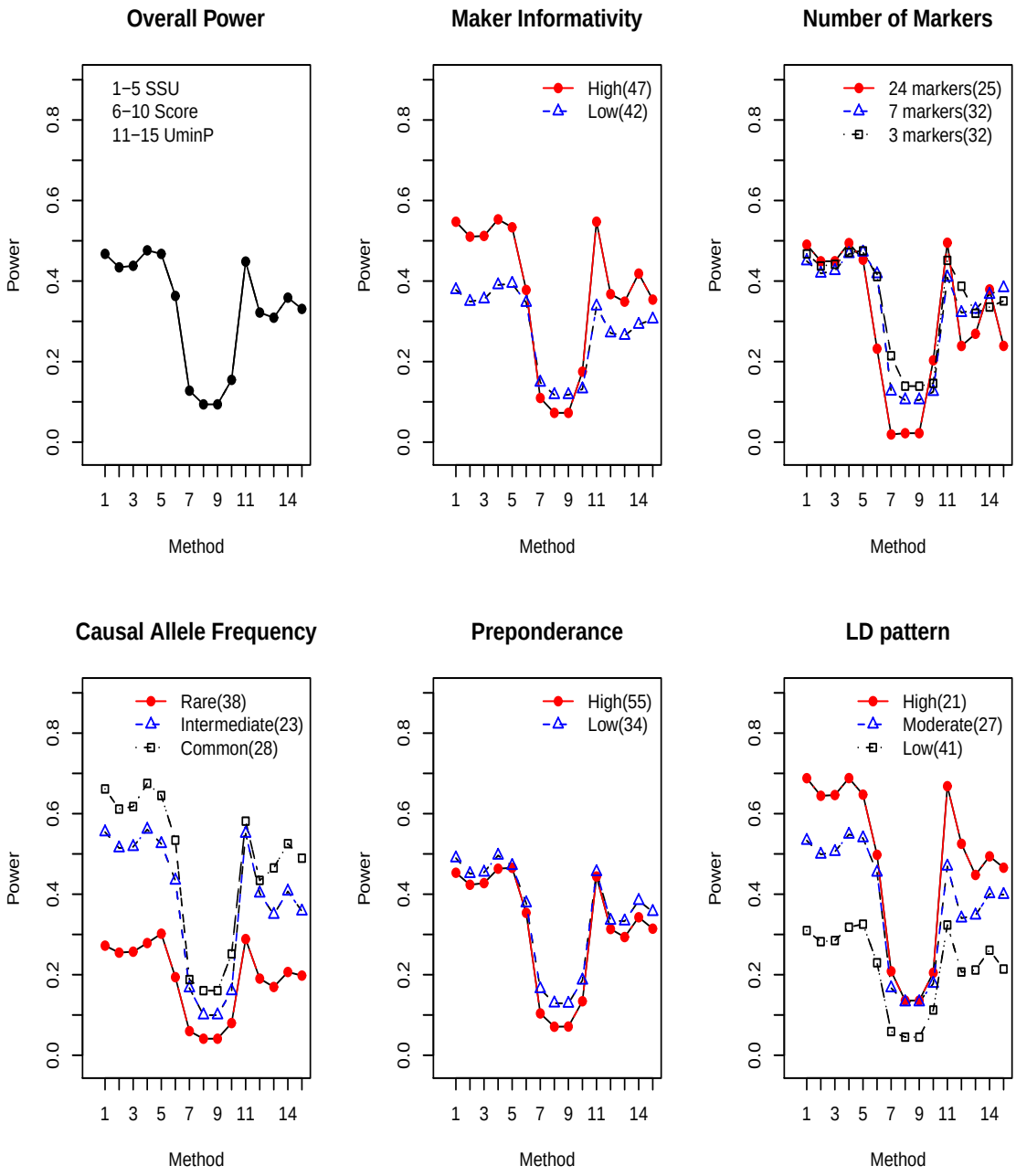

Figure 4: Empirical power of various tests from simulated data. Methods 1-5 (or 6-10 or 11-15) correspond to the SSU (or score or UminP) test applied to models L1-L5 respectively.

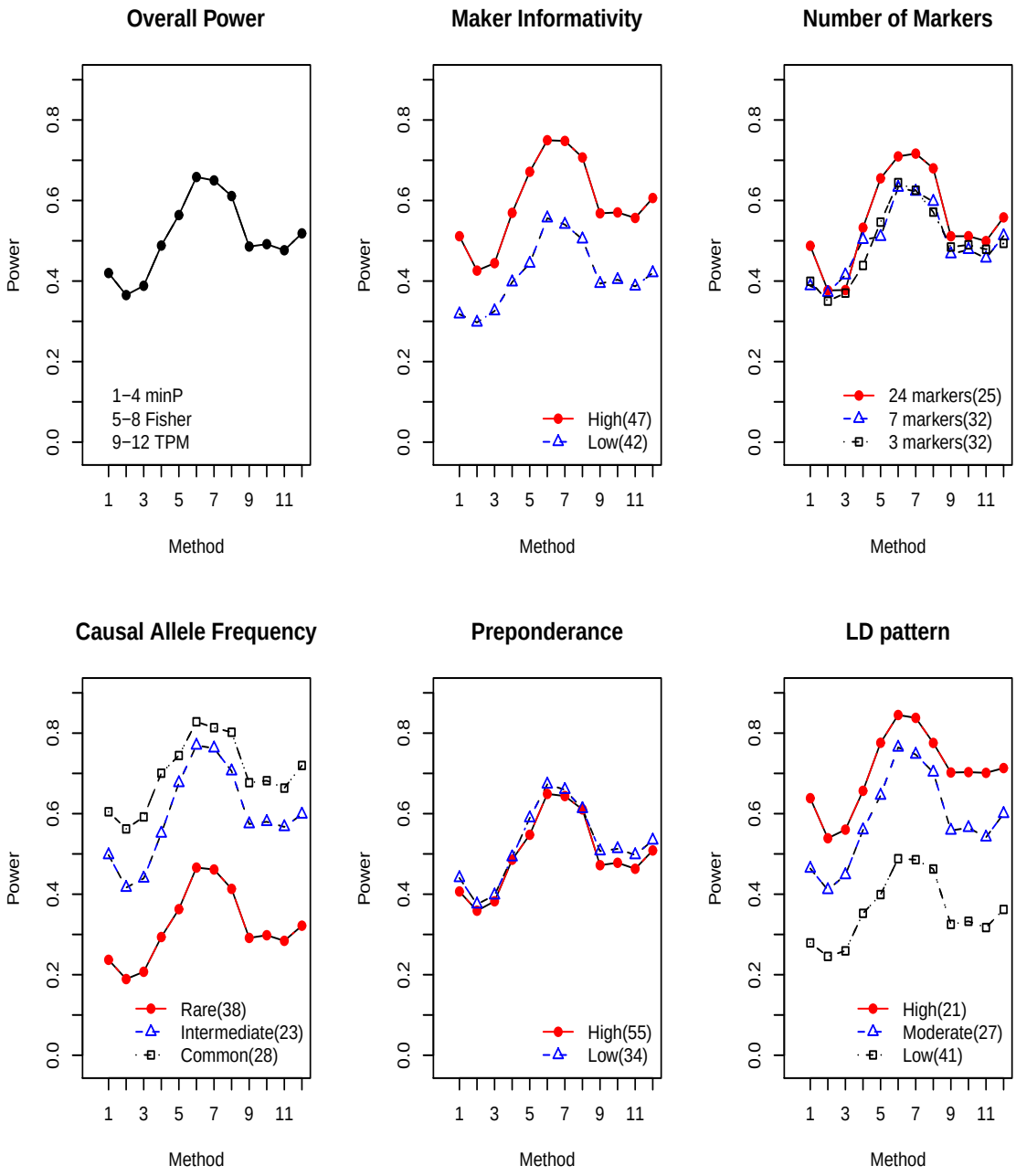

Figure 5: Empirical power of various tests from simulated data. Methods 1-4 (or 5-8 or 9-12) correspond to the minP (or Fisher or TMP) for combining SSU p-values for models L2-L5 respectively.