Supplementary Note

Marginal effects for two-locus models

Here we derive the marginal effect size of the three models given in Figure 1 of the main text. For each model we assume the two loci (*A* and *B*) are unlinked and have population allele frequencies π_A and π_B. The marginal odds at locus *A* are given by

$$\frac{p(D|g_A)}{p(\overline{D}|g_A)} = \frac{\sum_{g_B} p(D|g_A, g_B)p(g_B)}{\sum_{g_B} p(\overline{D}|g_A, g_B)p(g_B)},$$

where $p(D|g_A)$ is the probability that an individual has the disease given that they have genotype g_A at locus A and g_B. $p(\overline{D}|g_A)$ is the probability that an individual does not have the disease given that they have genotype g_A at locus A. For each model we give an expression for the parameter λ that is used in the power simulations and is defined as

$$\lambda = \frac{p(D|1_A)}{p(\overline{D}|1_A)} \Bigg/ \frac{p(D|0_A)}{p(\overline{D}|0_A)} - 1. \tag{1}$$

The resulting expression for λ will involve the allele frequencies (π_A and π_B) and the disease model parameters (typically α and θ in what follows). In our simulations we fix the allele frequencies, λ and the prevalence of disease p defined as

$$p = p(D) = \sum_{g_A, g_B} p(D|g_A, g_B)p(g_A, g_B). \tag{2}$$

We use numerical techniques to solve Eq.1 and Eq.2 for α and θ.

Model 1

$$\begin{aligned}
\frac{p(D|0_A)}{p(\overline{D}|0_A)} &= \frac{p(D|0_A,0_B)p(0_B)+p(D|0_A,1_B)p(1_B)+p(D|0_A,2_B)p(2_B)}{p(\overline{D}|0_A,0_B)p(0_B)+p(\overline{D}|0_A,1_B)p(1_B)+p(\overline{D}|0_A,2_B)p(2_B)} \\
&= \frac{\frac{\alpha}{1+\alpha}(1-\pi_B)^2+\frac{\alpha(1+\theta_B)}{1+\alpha(1+\theta_B)}2\pi_B(1-\pi_B)+\frac{\alpha(1+\theta_B)^2}{1+\alpha(1+\theta_B)^2}\pi_B^2}{\frac{1}{1+\alpha}(1-\pi_B)^2+\frac{1}{1+\alpha(1+\theta_B)}2\pi_B(1-\pi_B)+\frac{1}{1+\alpha(1+\theta_B)^2}\pi_B^2} \\
&= \alpha\left(1+\frac{\frac{2\theta_B\pi_B(1-\pi_B)}{1+\alpha(1+\theta_B)}+\frac{\theta_B(2+\theta_B)\pi_B^2}{1+\alpha(1+\theta_B)^2}}{\frac{(1-\pi_B)^2}{1+\alpha}+\frac{2\pi_B(1-\pi_B)}{1+\alpha(1+\theta_B)}+\frac{\pi_B^2}{1+\alpha(1+\theta_B)^2}}\right) \\
&= \alpha(1+\mu_1).
\end{aligned}$$

$$\begin{aligned}
\frac{p(D|1_A)}{p(\overline{D}|1_A)} &= \frac{p(D|1_A,0_B)p(0_B)+p(D|1_A,1_B)p(1_B)+p(D|1_A,2_B)p(2_B)}{p(\overline{D}|1_A,0_B)p(0_B)+p(\overline{D}|1_A,1_B)p(1_B)+p(\overline{D}|1_A,2_B)p(2_B)} \\
&= \frac{\frac{\alpha(1+\theta_A)}{1+\alpha(1+\theta_A)}(1-\pi_B)^2+\frac{\alpha(1+\theta_A)(1+\theta_B)}{1+\alpha(1+\theta_A)(1+\theta_B)}2\pi_B(1-\pi_B)+\frac{\alpha(1+\theta_A)(1+\theta_B)^2}{1+\alpha(1+\theta_A)(1+\theta_B)^2}\pi_B^2}{\frac{1}{1+\alpha(1+\theta_A)}(1-\pi_B)^2+\frac{1}{1+\alpha(1+\theta_A)(1+\theta_B)}2\pi_B(1-\pi_B)+\frac{1}{1+\alpha(1+\theta_A)(1+\theta_B)^2}\pi_B^2} \\
&= \alpha(1+\theta_A)\left(1+\frac{\frac{2\theta_B\pi_B(1-\pi_B)}{1+\alpha(1+\theta_A)(1+\theta_B)}+\frac{\theta_B(2+\theta_B)\pi_B^2}{1+\alpha(1+\theta_A)(1+\theta_B)^2}}{\frac{(1-\pi_B)^2}{1+\alpha(1+\theta_A)}+\frac{2\pi_B(1-\pi_B)}{1+\alpha(1+\theta_A)(1+\theta_B)}+\frac{\pi_B^2}{1+\alpha(1+\theta_A)(1+\theta_B)^2}}\right) \\
&= \alpha(1+\theta_A)(1+\mu_2).
\end{aligned}$$

$$\begin{aligned}
\frac{p(D|2_A)}{p(\overline{D}|2_A)} &= \frac{p(D|2_A,0_B)p(0_B)+p(D|2_A,1_B)p(1_B)+p(D|2_A,2_B)p(2_B)}{p(\overline{D}|2_A,0_B)p(0_B)+p(\overline{D}|2_A,1_B)p(1_B)+p(\overline{D}|2_A,2_B)p(2_B)} \\
&= \frac{\frac{\alpha(1+\theta_A)^2}{1+\alpha(1+\theta_A)^2}(1-\pi_B)^2+\frac{\alpha(1+\theta_A)^2(1+\theta_B)}{1+\alpha(1+\theta_A)^2(1+\theta_B)}2\pi_B(1-\pi_B)+\frac{\alpha(1+\theta_A)^2(1+\theta_B)^2}{1+\alpha(1+\theta_A)^2(1+\theta_B)^2}\pi_B^2}{\frac{1}{1+\alpha(1+\theta_A)^2}(1-\pi_B)^2+\frac{1}{1+\alpha(1+\theta_A)^2(1+\theta_B)}2\pi_B(1-\pi_B)+\frac{1}{1+\alpha(1+\theta_A)^2(1+\theta_B)^2}\pi_B^2} \\
&= \alpha(1+\theta_A)^2\left(1+\frac{\frac{2\theta_B\pi_B(1-\pi_B)}{1+\alpha(1+\theta_A)^2(1+\theta_B)}+\frac{\theta_B(2+\theta_B)\pi_B^2}{1+\alpha(1+\theta_A)^2(1+\theta_B)^2}}{\frac{(1-\pi_B)^2}{1+\alpha(1+\theta_A)^2}+\frac{2\pi_B(1-\pi_B)}{1+\alpha(1+\theta_A)^2(1+\theta_B)}+\frac{\pi_B^2}{1+\alpha(1+\theta_A)^2(1+\theta_B)^2}}\right) \\
&= \alpha(1+\theta_A)^2(1+\mu_3).
\end{aligned}$$

Thus, the marginal odds can be expressed as

	Genotype	Odds
(aa)	0	$\alpha(1+\mu_1)$
(Aa)	1	$\alpha(1+\mu_2)(1+\theta_A)$
(AA)	2	$\alpha(1+\mu_3)(1+\theta_A)^2$

and $\lambda = \frac{(1+\mu_2)(1+\theta_A)}{(1+\mu_1)} - 1$.

The marginal odds ratios of Model 1 are similar to those for a model which involves just one locus i.e. each disease allele multiples the odds by a factor $(1+\theta_A)$. Moreover, when the prevalence of disease is low e.g. $p = 0.01$ and the marginal effect sizes are small (i.e. $\theta_A = \theta_B = 0.5$) the conditional probabilities of genotype given disease for the marginalised Model 2 are approximately the same as those for a single locus model. This implies that, when evaluating one locus at a time, the genetic effects for two interacing multiplicative loci appear just like those for a single locus. In this situation we might expect that the advantage in searching over models with more than one locus is minimized as the relevant loci exhibit a strong signal that is not greatly diminished by the action of the interaction.

Model 2

$$\begin{aligned}
\frac{p(D|0_A)}{p(\overline{D}|0_A)} &= \frac{p(D|0_A,0_B)p(0_B) + p(D|0_A,1_B)p(1_B) + p(D|0_A,2_B)p(2_B)}{p(\overline{D}|0_A,0_B)p(0_B) + p(\overline{D}|0_A,1_B)p(1_B) + p(\overline{D}|0_A,2_B)p(2_B)} \\
&= \frac{\frac{\alpha}{1+\alpha}(1-\pi_B)^2 + \frac{\alpha}{1+\alpha}2\pi_B(1-\pi_B) + \frac{\alpha}{1+\alpha}\pi_B^2}{\frac{1}{1+\alpha}(1-\pi_B)^2 + \frac{1}{1+\alpha}2\pi_B(1-\pi_B) + \frac{1}{1+\alpha}\pi_B^2} \\
&= \alpha.
\end{aligned}$$

$$\begin{aligned}
\frac{p(D|1_A)}{p(\overline{D}|1_A)} &= \frac{p(D|1_A,0_B)p(0_B) + p(D|1_A,1_B)p(1_B) + p(D|1_A,2_B)p(2_B)}{p(\overline{D}|1_A,0_B)p(0_B) + p(\overline{D}|1_A,1_B)p(1_B) + p(\overline{D}|1_A,2_B)p(2_B)} \\
&= \frac{\frac{\alpha}{1+\alpha}(1-\pi_B)^2 + \frac{\alpha(1+\theta_B)}{1+\alpha(1+\theta_B)}2\pi_B(1-\pi_B) + \frac{\alpha(1+\theta_B)^2}{1+\alpha(1+\theta_B)^2}\pi_B^2}{\frac{1}{1+\alpha}(1-\pi_B)^2 + \frac{1}{1+\alpha(1+\theta_B)}2\pi_B(1-\pi_B) + \frac{1}{1+\alpha(1+\theta_B)^2}\pi_B^2} \\
&= \alpha\left(1 + \frac{\frac{2\theta_B\pi_B(1-\pi_B)}{1+\alpha(1+\theta_B)} + \frac{\theta_B(2+\theta_B)\pi_B^2}{1+\alpha(1+\theta_B)^2}}{\frac{(1-\pi_B)^2}{1+\alpha} + \frac{2\pi_B(1-\pi_B)}{1+\alpha(1+\theta_B)} + \frac{\pi_B^2}{1+\alpha(1+\theta_B)^2}}\right) \\
&= \alpha(1+\lambda_1).
\end{aligned}$$

$$\begin{aligned}
\frac{p(D|2_A)}{p(\overline{D}|2_A)} &= \frac{p(D|2_A,0_B)p(0_B)+p(D|2_A,1_B)p(1_B)+p(D|2_A,2_B)p(2_B)}{p(\overline{D}|2_A,0_B)p(0_B)+p(\overline{D}|2_A,1_B)p(1_B)+p(\overline{D}|2_A,2_B)p(2_B)} \\
&= \frac{\frac{\alpha}{1+\alpha}(1-\pi_B)^2+\frac{\alpha(1+\theta)^2}{1+\alpha(1+\theta)^2}2\pi_B(1-\pi_B)+\frac{\alpha(1+\theta)^4}{1+\alpha(1+\theta)^4}\pi_B^2}{\frac{1}{1+\alpha}(1-\pi_B)^2+\frac{1}{1+\alpha(1+\theta)^2}2\pi_B(1-\pi_B)+\frac{1}{1+\alpha(1+\theta)^4}\pi_B^2} \\
&= \alpha\left(1+\frac{\frac{2\theta_B(\theta_B+2)\pi_B(1-\pi_B)}{1+\alpha(1+\theta_B)^2}+\frac{((1+\theta_B)^4-1)\pi_B^2}{1+\alpha(1+\theta_B)^4}}{\frac{(1-\pi_B)^2}{1+\alpha}+\frac{2\pi_B(1-\pi_B)}{1+\alpha(1+\theta_B)^2}+\frac{\pi_B^2}{1+\alpha(1+\theta_B)^4}}\right) \\
&= \alpha(1+\lambda_2).
\end{aligned}$$

Thus the marginal odds can be expressed as

	Genotype	Odds
(aa)	0	α
(Aa)	1	$\alpha(1+\lambda_1)$
(AA)	2	$\alpha(1+\lambda_2)$

where

$$\lambda = \lambda_1 = \frac{\frac{2\theta\pi_B(1-\pi_B)}{1+\alpha(1+\theta)}+\frac{\theta(2+\theta)\pi_B^2}{1+\alpha(1+\theta)^2}}{\frac{(1-\pi_B)^2}{1+\alpha}+\frac{2\pi_B(1-\pi_B)}{1+\alpha(1+\theta)}+\frac{\pi_B^2}{1+\alpha(1+\theta)^2}}. \tag{3}$$

This result shows the size of effect we can expect to see marginally (at locus A) for a Model 2 interaction parameterized by θ that involves an unobserved locus (locus B) with allele frequency π_B. Supplementary Figure 1 illustrates the size of the marginal effect for a range of different values of θ and π_B.

Model 3

$$\begin{aligned}
\frac{p(D|0_A)}{p(\overline{D}|0_A)} &= \frac{p(D|0_A,0_B)p(0_B)+p(D|0_A,1_B)p(1_B)+p(D|0_A,2_B)p(2_B)}{p(\overline{D}|0_A,0_B)p(0_B)+p(\overline{D}|0_A,1_B)p(1_B)+p(\overline{D}|0_A,2_B)p(2_B)} \\
&= \frac{\frac{\alpha}{1+\alpha}(1-\pi_B)^2+\frac{\alpha}{1+\alpha}2\pi_B(1-\pi_B)+\frac{\alpha}{1+\alpha}\pi_B^2}{\frac{1}{1+\alpha}(1-\pi_B)^2+\frac{1}{1+\alpha}2\pi_B(1-\pi_B)+\frac{1}{1+\alpha}\pi_B^2} \\
&= \alpha.
\end{aligned}$$

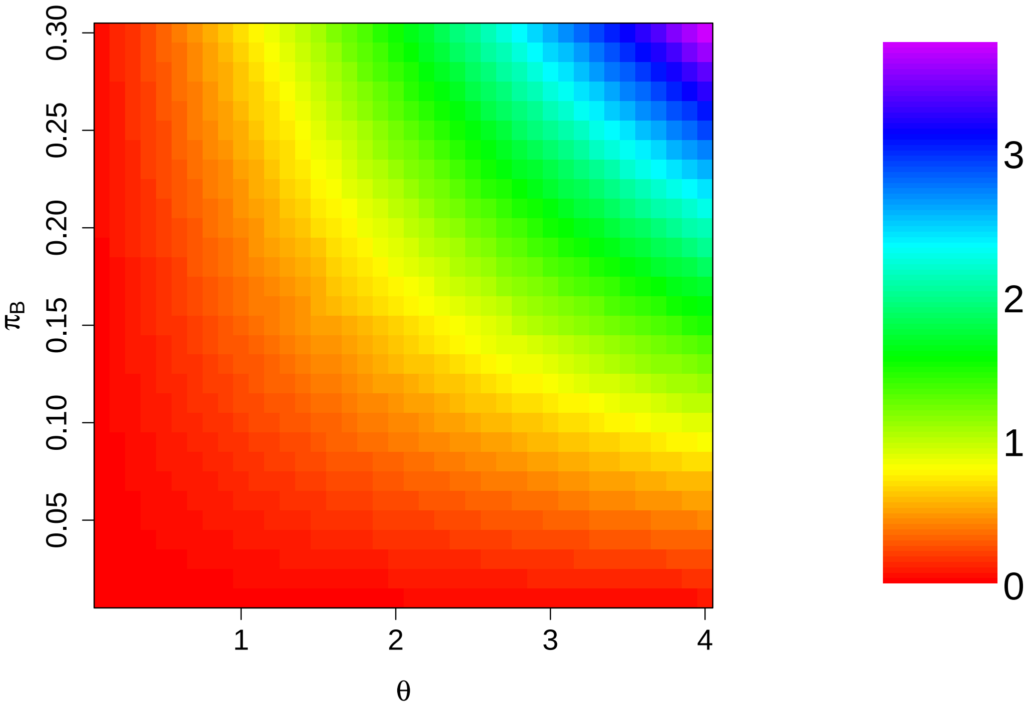

Figure 1: Marginal effect size for Model 2 (λ in equation (3)) for a grid of values of interaction parameter (θ) and unobserved marker allele frequency (π_B).

$$\begin{aligned}
\frac{p(D|1_A)}{p(\overline{D}|1_A)} &= \frac{p(D|1_A,0_B)p(0_B)+p(D|1_A,1_B)p(1_B)+p(D|1_A,2_B)p(2_B)}{p(\overline{D}|1_A,0_B)p(0_B)+p(\overline{D}|1_A,1_B)p(1_B)+p(\overline{D}|1_A,2_B)p(2_B)} \\
&= \frac{\frac{\alpha}{1+\alpha}(1-\pi_B)^2+\frac{\alpha(1+\theta_B)}{1+\alpha(1+\theta_B)}\pi_B(2-\pi_B)}{\frac{1}{1+\alpha}(1-\pi_B)^2+\frac{1}{1+\alpha(1+\theta_B)}\pi_B(2-\pi_B)} \\
&= \alpha\left(1+\frac{\frac{\theta_B\pi_B(2-\pi_B)}{1+\alpha(1+\theta_B)}}{\frac{(1-\pi_B)^2}{1+\alpha}+\frac{\pi_B(2-\pi_B)}{1+\alpha(1+\theta_B)}}\right) \\
&= \alpha(1+\lambda_1).
\end{aligned}$$

$$\begin{aligned}
\frac{p(D|2_A)}{p(\overline{D}|2_A)} &= \frac{p(D|2_A,0_B)p(0_B)+p(D|2_A,1_B)p(1_B)+p(D|2_A,2_B)p(2_B)}{p(\overline{D}|2_A,0_B)p(0_B)+p(\overline{D}|2_A,1_B)p(1_B)+p(\overline{D}|2_A,2_B)p(2_B)} \\
&= \frac{\frac{\alpha}{1+\alpha}(1-\pi_B)^2+\frac{\alpha(1+\theta_B)}{1+\alpha(1+\theta_B)}\pi_B(2-\pi_B)}{\frac{1}{1+\alpha}(1-\pi_B)^2+\frac{1}{1+\alpha(1+\theta_B)}\pi_B(2-\pi_B)} \\
&= \alpha\left(1+\frac{\frac{\theta_B\pi_B(2-\pi_B)}{1+\alpha(1+\theta_B)}}{\frac{(1-\pi_B)^2}{1+\alpha}+\frac{\pi_B(2-\pi_B)}{1+\alpha(1+\theta_B)}}\right) \\
&= \alpha(1+\lambda_2).
\end{aligned}$$

Thus the marginal odds can be expressed as

	Genotype	Odds
(aa)	0	α
(Aa)	1	$\alpha(1+\lambda_1)$
(AA)	2	$\alpha(1+\lambda_2)$

and $\lambda = \lambda_1$.

Incorporating linkage disequilibrium

To avoid the assumption that we will have typed the causative loci in a given study we consider the general situation in which we have data at two loci (*X* and *Y*) that are in LD with the two disease loci (*A* and *B*). In this situation we can calculate the odds of disease at the two loci *X* and *Y* if we have the odds at the disease loci *A* and *B* and details of the LD between the pairs (*X*, *A*) and (*Y*, *B*). The odds can be calculated as

$$\begin{aligned}
\frac{p(D|g_X, g_Y)}{p(D'|g_X, g_Y)} &= \frac{\sum_{g_A,g_B} p(D|g_A, g_B)p(g_A, g_B|g_X, g_Y)}{\sum_{g_A,g_B} p(D'|g_A, g_B)p(g_A, g_B|g_X, g_Y)} \\
&= \frac{\sum_{g_A,g_B} p(D|g_A, g_B)p(g_X, g_Y|g_A, g_B)p(g_A, g_B)}{\sum_{g_A,g_B} p(D'|g_A, g_B)p(g_X, g_Y|g_A, g_B)p(g_A, g_B)} \\
&= \frac{\sum_{g_A,g_B} p(D|g_A, g_B)p(g_X|g_A)p(g_Y|g_B)p(g_A)p(g_B)}{\sum_{g_A,g_B} p(D'|g_A, g_B)p(g_X|g_A)p(g_Y|g_B)p(g_A)p(g_B)},
\end{aligned}$$

where $p(g_A)$ and $p(g_B)$ are the genotype probabilities at loci *A* and *B* and are specified by the allele frequencies π_A and π_B. The conditional genotype probabilities, $p(g_X|g_A)$ and $p(g_Y|g_B)$, are specified by the conditional haplotype probabilities $p(X|A), p(X|a), p(Y|B)$ and $p(Y|b)$. For example, the conditional genotype probabilities $p(g_X|g_A)$ are given by

$$\begin{aligned}
p(0_X|0_A) &= p(x|a)^2 \\
p(0_X|1_A) &= p(x|a)p(x|A) \\
p(0_X|2_A) &= p(x|A)^2 \\
p(1_X|0_A) &= 2p(x|a)p(X|a) \\
p(1_X|1_A) &= p(x|a)p(X|A) + p(X|a)p(x|A) \\
p(1_X|2_A) &= 2p(x|A)p(X|A) \\
p(2_X|0_A) &= p(X|a)^2 \\
p(2_X|1_A) &= p(X|a)p(X|A) \\
p(2_X|2_A) &= p(X|A)^2
\end{aligned}$$

The conditional haplotype probabilities also specify the allele frequencies at loci A and D (π_A and π_D). Thus we can calculate the LD within each pair of loci using the squared correlation coefficient r^2 (Pritchard and Przeworski, 2001). That is

$$\begin{aligned}
r^2_{BA} &= \Big[p(A|B) - p(A|b)\Big]^2 \frac{\pi_B(1-\pi_B)}{\pi_A(1-\pi_A)} \\
r^2_{CD} &= \Big[p(D|C) - p(D|c)\Big]^2 \frac{\pi_C(1-\pi_C)}{\pi_D(1-\pi_D)}
\end{aligned}$$

Alternatively, we can specify the value of r^2 and calculate the value of the conditional haplotype probabilities. To precisely determine these probabilities we need to impose an additional constraint. We have used 3 different constraints for $p(X|A)$ and $p(X|a)$ (and similarly for $p(Y|B)$ and $p(Y|b)$):

(Constraint 1) $\quad p(A|B) = q \quad p(A|b) = 1-q \quad \Rightarrow \pi_B \leq \pi_A \leq 1-\pi_B$

(Constraint 2) $\quad p(A|B) = 1 \quad p(A|b) = q \quad \Rightarrow \pi_B \leq \pi_A \leq 1$

(Constraint 3) $\quad p(A|B) = q \quad p(A|b) = 0 \quad \Rightarrow 0 \leq \pi_A \leq \pi_B$

Constraint 1 attempts to mimic the situation in which the typed allele, X, is at a higher frequency (and likely to be older) than the causative allele, A, and when it is possible for all 4 haplotypes to be present in a given sample, which will only occur if at least 1 visible recombination event has occurred between the markers (assuming no repeat mutations or genotyping error). Constraint 2 attempts to mimic the situation in which the typed allele, X, is at a higher frequency (and likely to be older) than the causative allele, A, but there are no xA haplotypes and thus no inferable evidence of recombination. Constraint 3 attempts to mimic the situation in which the typed allele, X, is at a lower frequency (and likely to be younger) than the causative allele, A, but there are no Xa haplotypes and thus no directly inferable evidence of recombination.

To investigate the most realistic constraint we carried out a coalescent simulation study (Hudson, 1990). We simulated 1000 chromosomes from a neutral coalescent model over a region of 100kb. We set the scaled recombination rate ($\rho = 4N_e r$) for the region to the genome-wide average of approximately 0.4/kb. We set the scaled mutation rate ($\theta = 4N_e \mu$) to be 100 and ascertained SNPs by simulating an additional 4 sequences and ascertaining the SNP if it segregated on these 4 sequences. For each simulated dataset we considered all pairs of SNPs at distance less than 15kb apart. We found in 43$\%$ of such pairs there was a detectable recombination event. Thus, constraints 2 and 3 are slightly more likely than constraint 1.

A numerical example

To illustrate the process of marginalization and the incorporation of the effects of LD we constructed a numerical example based on Model 2. The model consists of 2 unlinked causative loci (A and B) linked to 2 typed markers (X and Y). The marginal effect size at the two interaction loci was fixed at $\lambda = 0.5$, minor allele frequencies were fixed at 0.1, haplotype probabilities were specified by the conditional probabilities $p(X|A) = p(Y|B) = 0.95$ and $p(X|a) = p(Y|b) = 0.05$ (Constraint 1). This resulted in minor allele frequencies at the typed SNPs (X and Y) of $\pi_X = \pi_Y = 0.14$ and correlations between typed and causative markers of $r^2_{AX} = r^2_{BY} = 0.606$. The table below shows the odds for the pair of causative loci, the pair of typed loci and the marginal models for loci *X* and *A*.

A	B	Odds
0	0	α
0	1	α
0	2	α
1	0	α
1	1	$\alpha(1+2.305)$
1	2	$\alpha(1+9.928)$
2	0	α
2	1	$\alpha(1+9.928)$
2	2	$\alpha(1+118.44)$

$\Rightarrow$

X	Y	Odds
0	0	α'
0	1	$\alpha'(1+0.0179)$
0	2	$\alpha'(1+0.0606)$
1	0	$\alpha'(1+0.0179)$
1	1	$\alpha'(1+1.093)$
1	2	$\alpha'(1+3.746)$
2	0	$\alpha'(1+0.0606)$
2	1	$\alpha'(1+3.746)$
2	2	$\alpha'(1+19.019)$

A	Odds
0	α
1	$\alpha(1+0.5)$
2	$\alpha(1+2.259)$

$\Rightarrow$

X	Odds
0	α
1	$\alpha(1+0.339)$
2	$\alpha(1+1.238)$

The effects of allele frequency differences

In addition to our investigations of the power of different detection strategies we carried out a quantitative assessment of the suggestion that allele frequency differences in an initial detection population and a subsequent replicate population provide a potential explanation for the lack of replication of some association studies.

The expression for the marginal effect of Model 2 (equation (3) in this Supplementary Note) suggests that the marginal effect size at an observed causative locus depends upon the allele frequency of the unobserved locus. Thus if the allele frequency of the unobserved locus is small we will have low power to detect the disease loci. If we were to carry out two studies in two separate populations that differed in allele frequency at the unobserved locus then we may find a significant effect at the observed locus in one population but not in the other. This effect is a potential explanation for the lack of replication in association studies. An analogous situation will occur if the disease model includes an interaction between a causative locus and an unobserved environmental factor. In order to examine this effect we carried out a simulation study to assess how often this effect will occur.

We considered the situation in which an interaction occurs between two loci (*A* and *B*) according to Model 2 with $n = 1000, \lambda = 2.0, p = 0.01$. At both loci we assumed that the allele frequencies had diverged from an ancestral frequency of p using a Beta distribution (Balding, 2003; Nicholson et al., 2002; Marchini and Cardon, 2002; Balding and Nichols, 1995) such that

$$\pi_A^{(i)} \sim \beta\Big(0.1\frac{1-c}{c}, 0.9\frac{1-c}{c}\Big) \qquad \pi_B^{(i)} \sim \beta\Big(\pi\frac{1-c}{c}, (1-\pi)\frac{1-c}{c}\Big)\ i = 1, 2,$$

where $\pi_A^{(i)}$ and $\pi_B^{(i)}$ are the *A* and *B* allele frequencies in population i, π is the global/ancestral allele frequency at locus *B* and c specifies the amount of divergence around the 'ancestral' allele frequency (0.1 at locus *A* and π at locus *B*) in the two populations. In this model c is equivalent to F_{ST} (Balding, 2003). We have recently validated this model on several large datasets (Marchini et al., 2004).

After generating the subpopulation allele frequencies we simulated datasets for the 2 populations and tested for association at both loci in both populations assuming a total of 300,000 loci in the study. We repeated this simulation 1000 times and calculated the percentage of times in which the results of the tests at locus *A* in the two populations differed i.e. locus *A* was identified in only one population. We chose to focus on locus *A* as it has (on average) the same allele frequency in both populations thus any differences in detection will be mostly due to the different allele frequencies at the interacting locus, *B*. The results in Figure 3 of the main text show that as the extent of differentiation increases (c increases) we see a higher percentage of simulations showing a difference between populations. As c increases the allele frequencies at the effectively unobserved locus *B* will become more distinct in the two populations causing a larger difference in effect size at locus *A*. Also, in general as π increases the percentage of simulations showing a difference between populations increases.

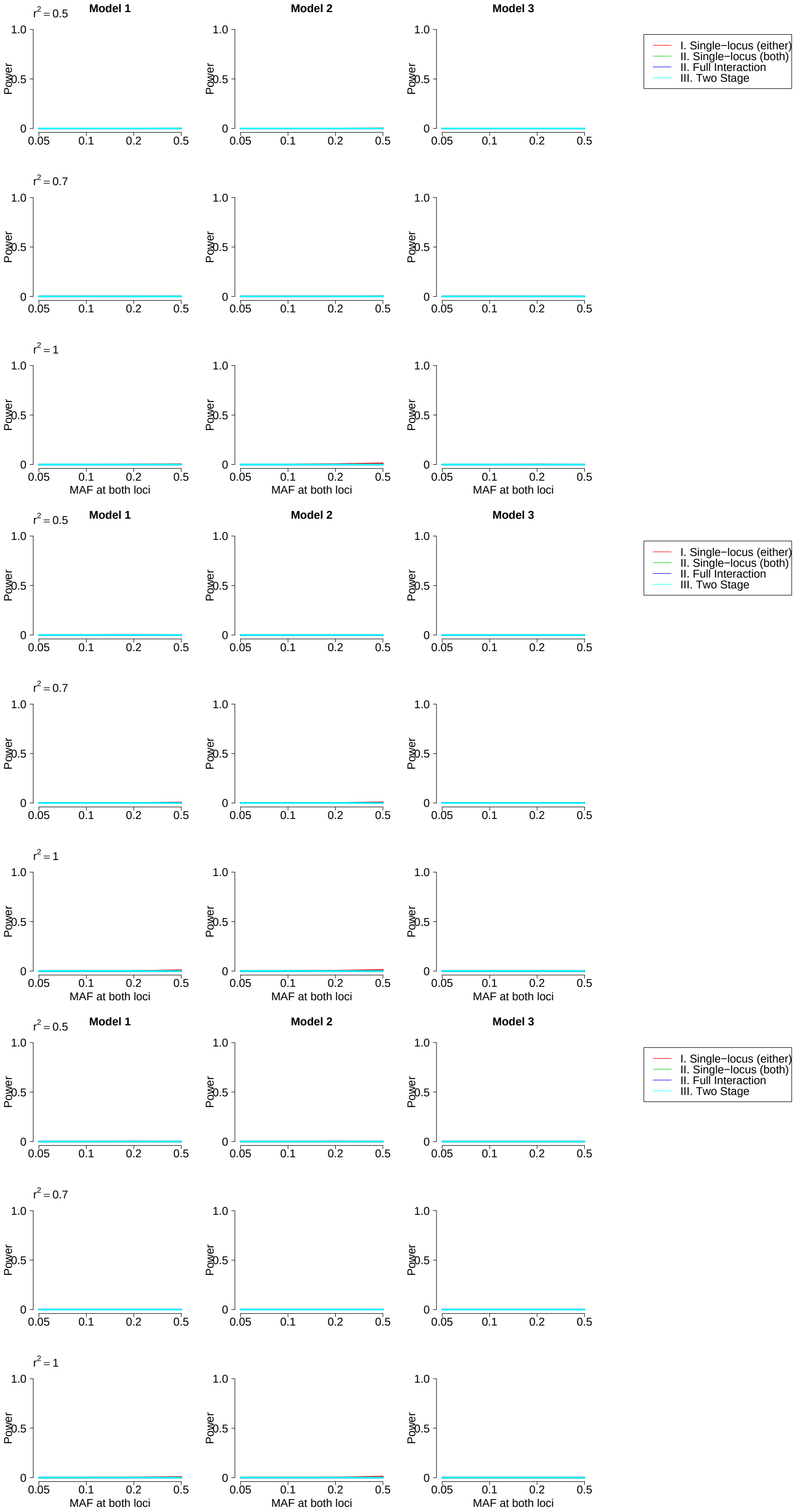

Figure 2: Power of search strategies for parameters $n = 1000, \lambda = 0.2$ and for Constraint 1 (first 9 plots), Constraint 2 (second 9 plots) and Constraint 3 (last 9 plots

.

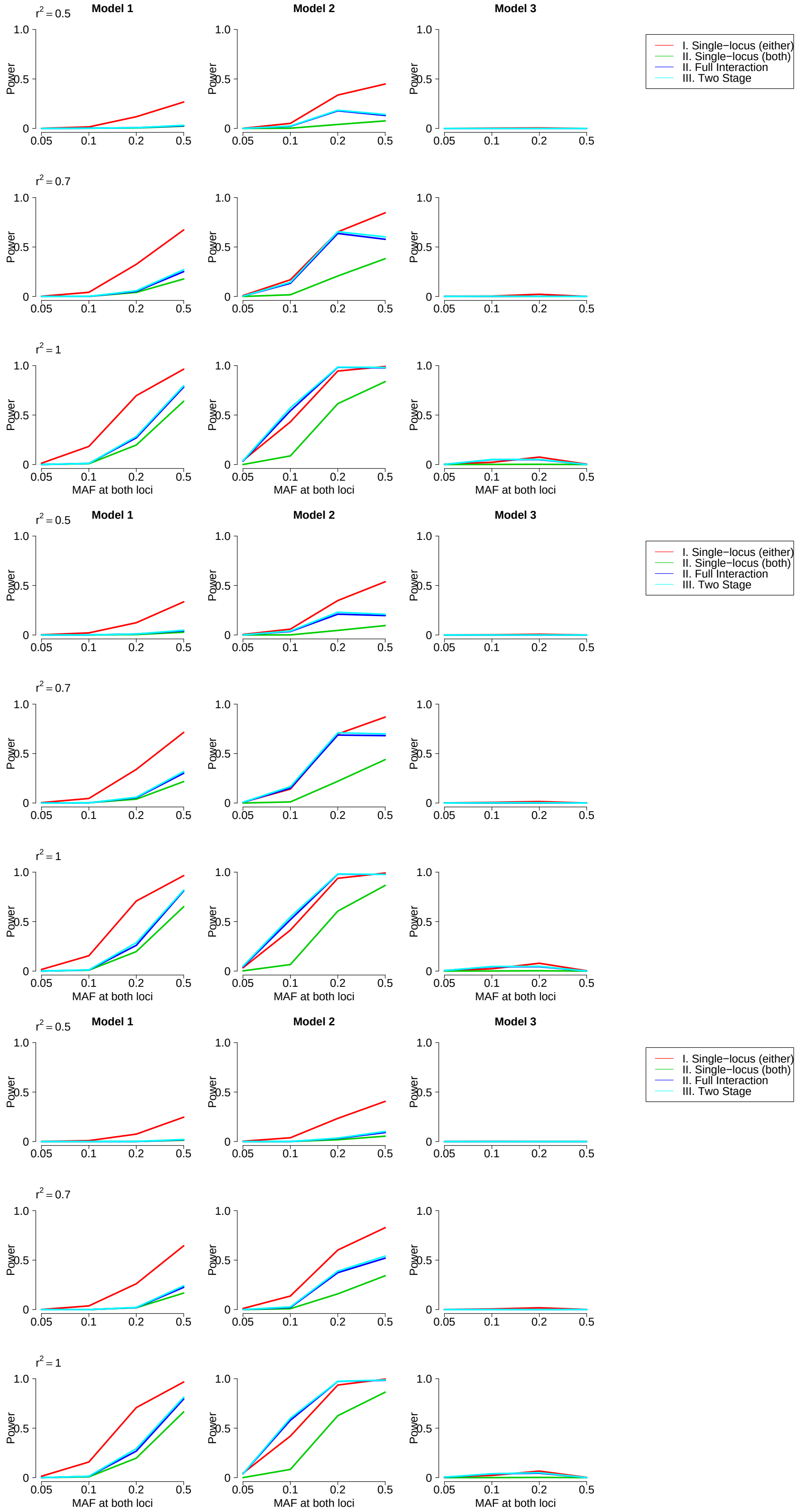

Figure 3: Power of search strategies for parameters $n = 1000, \lambda = 0.5$ and for Constraint 1 (first 9 plots), Constraint 2 (second 9 plots) and Constraint 3 (last 9 plots

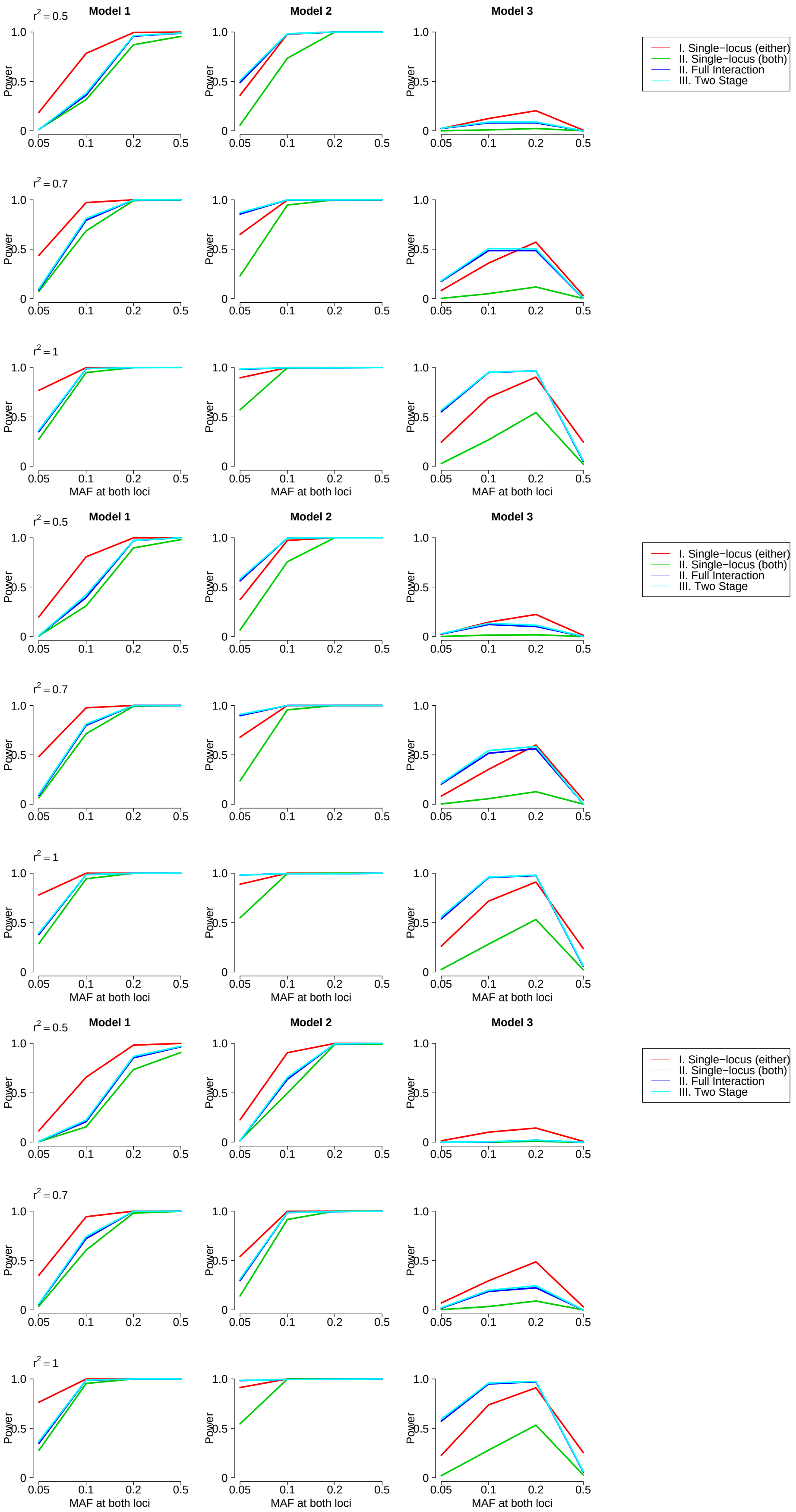

Figure 4: Power of search strategies for parameters $n = 1000, \lambda = 1.0$ and for Constraint 1 (first 9 plots), Constraint 2 (second 9 plots) and Constraint 3 (last 9 plots

.

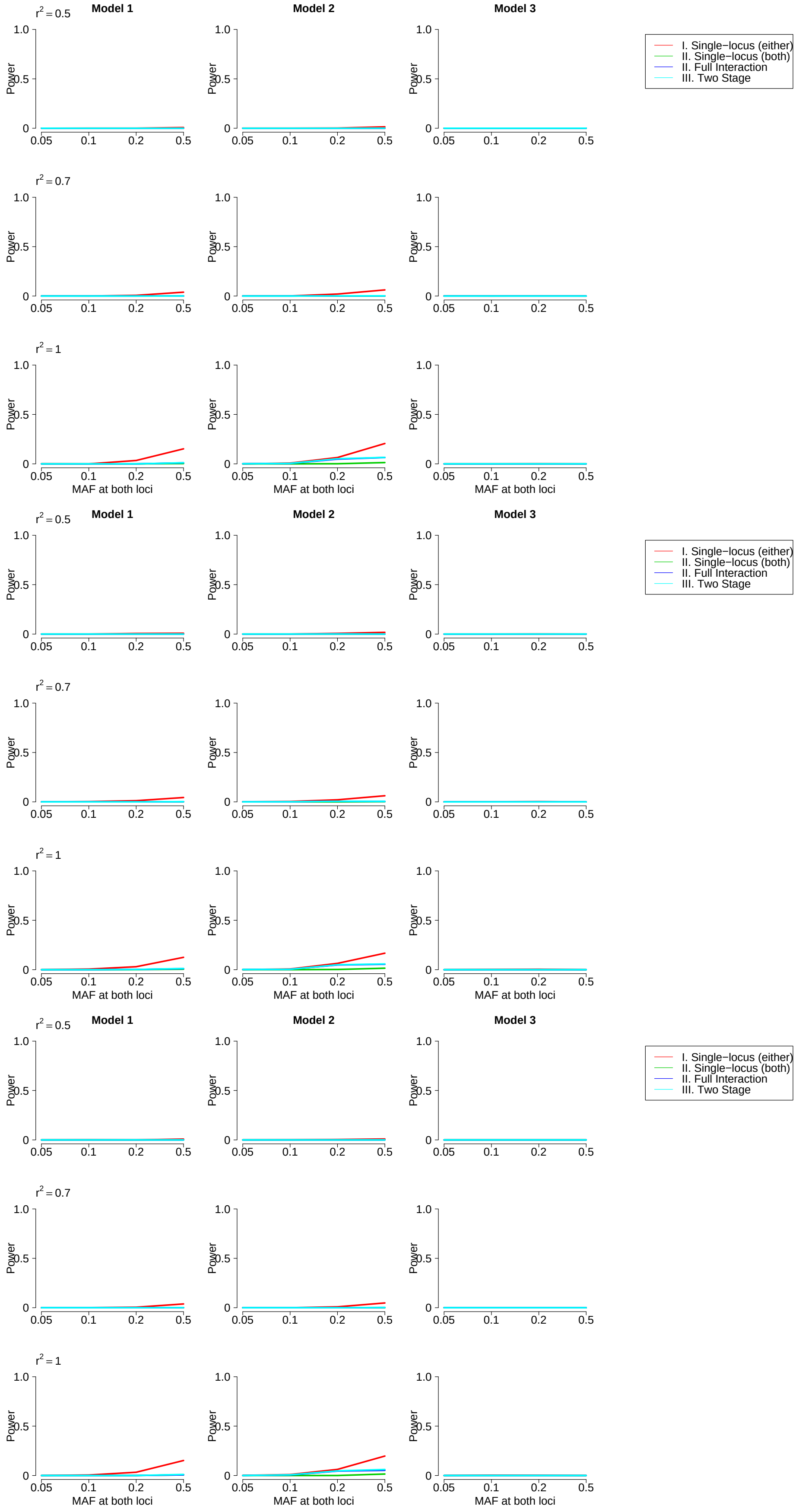

Figure 5: Power of search strategies for parameters $n = 2000, \lambda = 0.2$ and for Constraint 1 (first 9 plots), Constraint 2 (second 9 plots) and Constraint 3 (last 9 plots

.

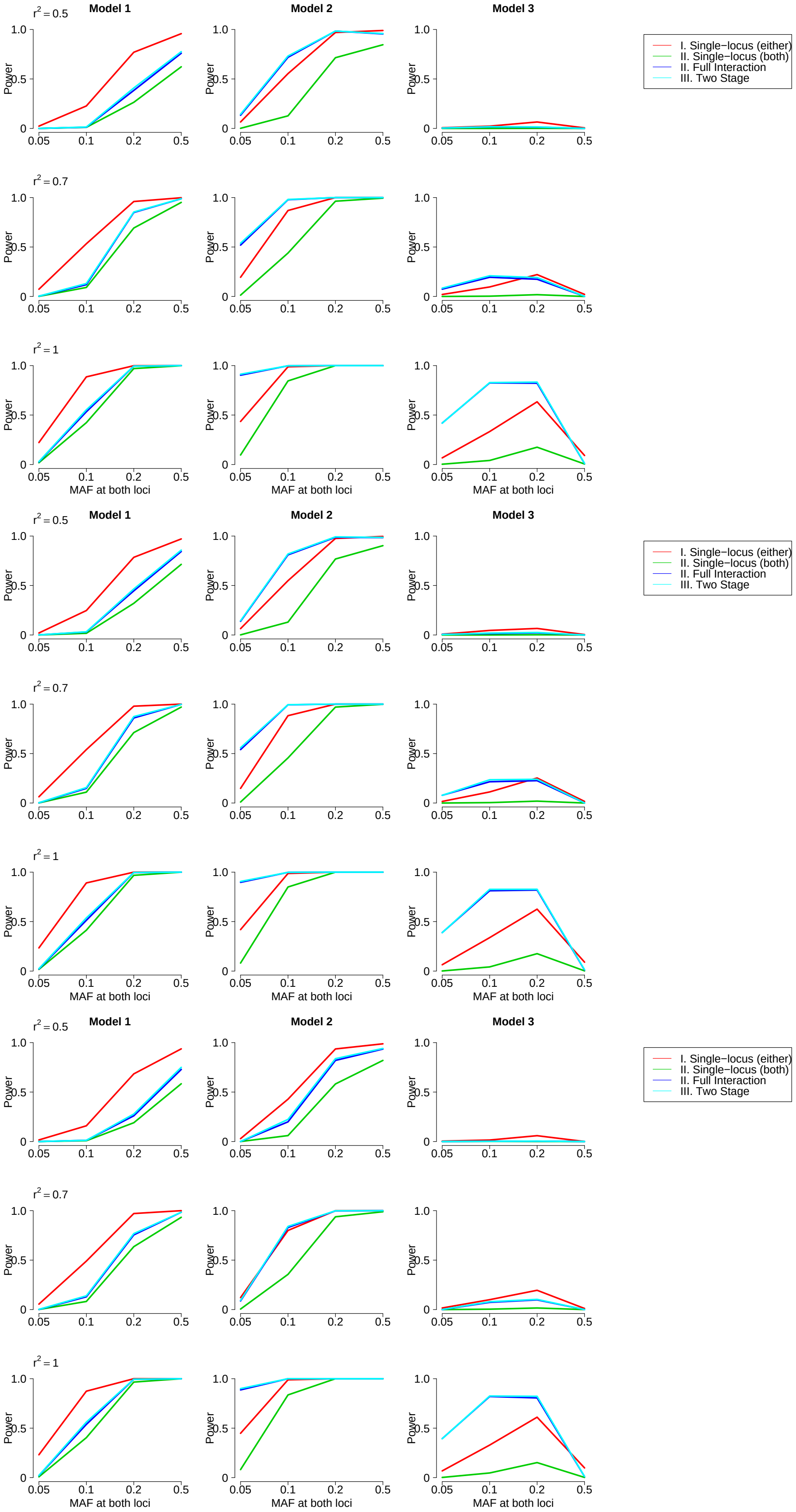

Figure 6: Power of search strategies for parameters $n = 2000, \lambda = 0.5$ and for Constraint 1 (first 9 plots), Constraint 2 (second 9 plots) and Constraint 3 (last 9 plots

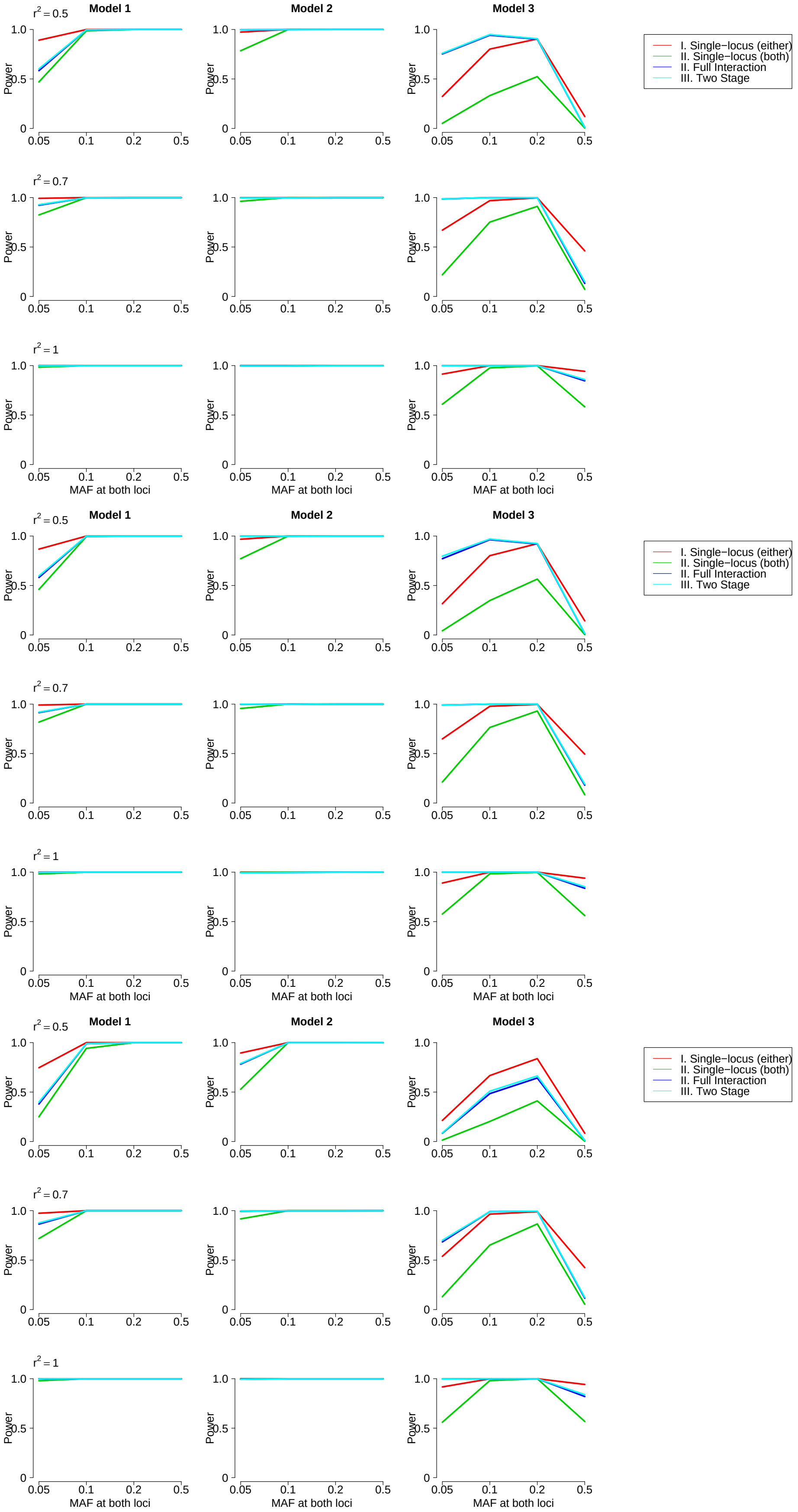

Figure 7: Power of search strategies for parameters $n = 2000, \lambda = 1.0$ and for Constraint 1 (first 9 plots), Constraint 2 (second 9 plots) and Constraint 3 (last 9 plots

.

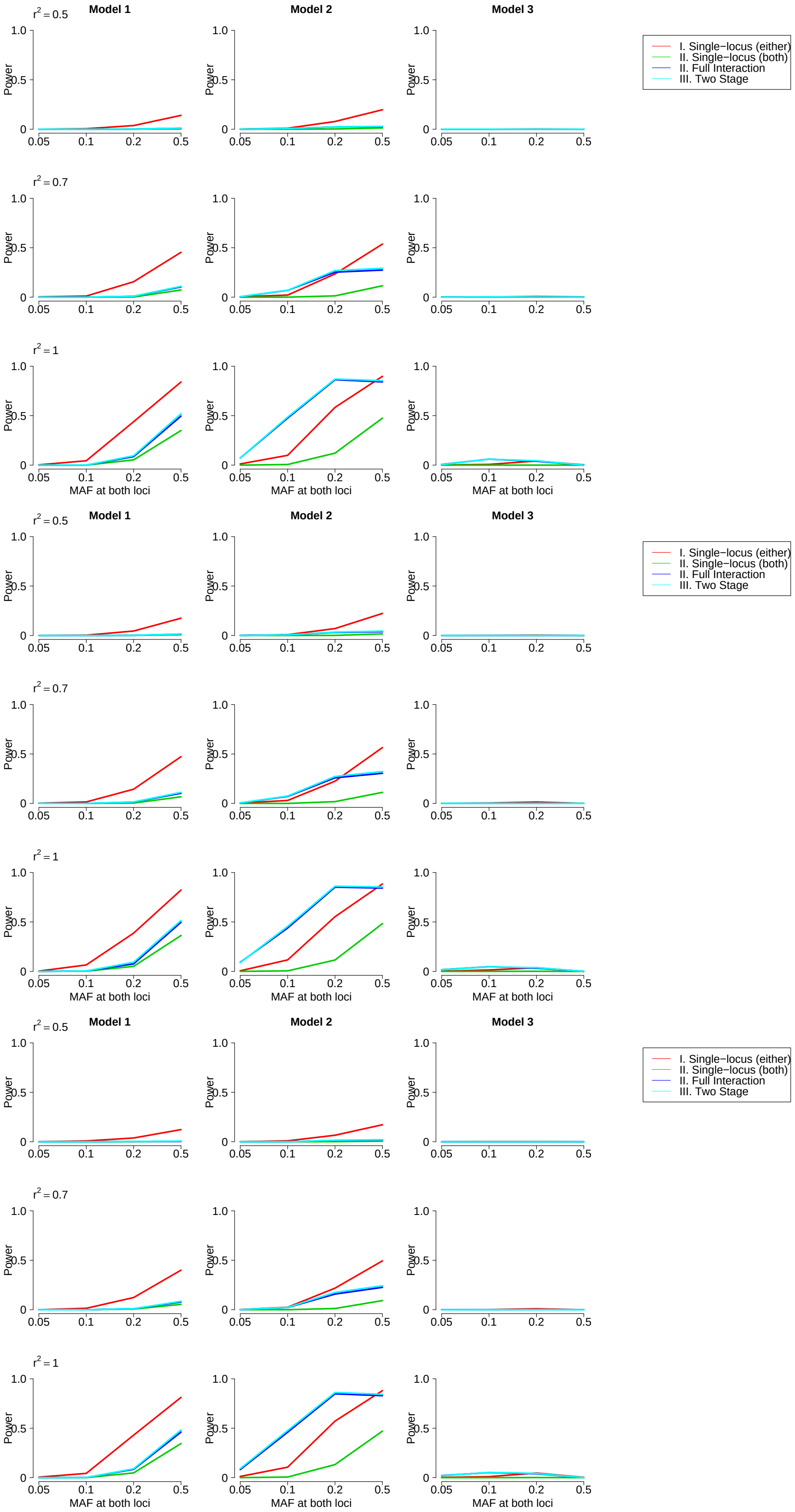

Figure 8: Power of search strategies for parameters $n = 4000, \lambda = 0.2$ and for Constraint 1 (first 9 plots), Constraint 2 (second 9 plots) and Constraint 3 (last 9 plots

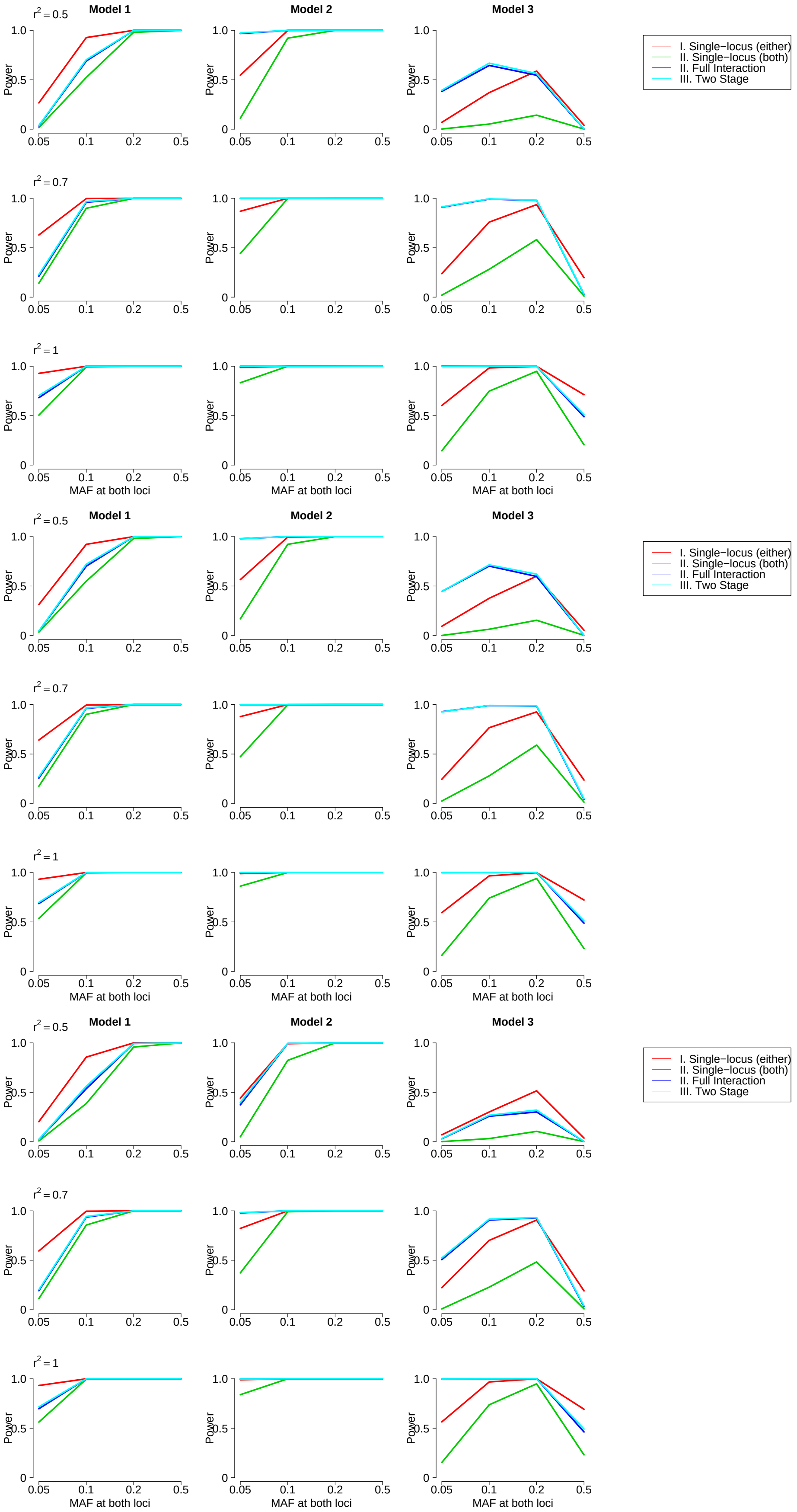

Figure 9: Power of search strategies for parameters $n = 4000, \lambda = 0.5$ and for Constraint 1 (first 9 plots), Constraint 2 (second 9 plots) and Constraint 3 (last 9 plots

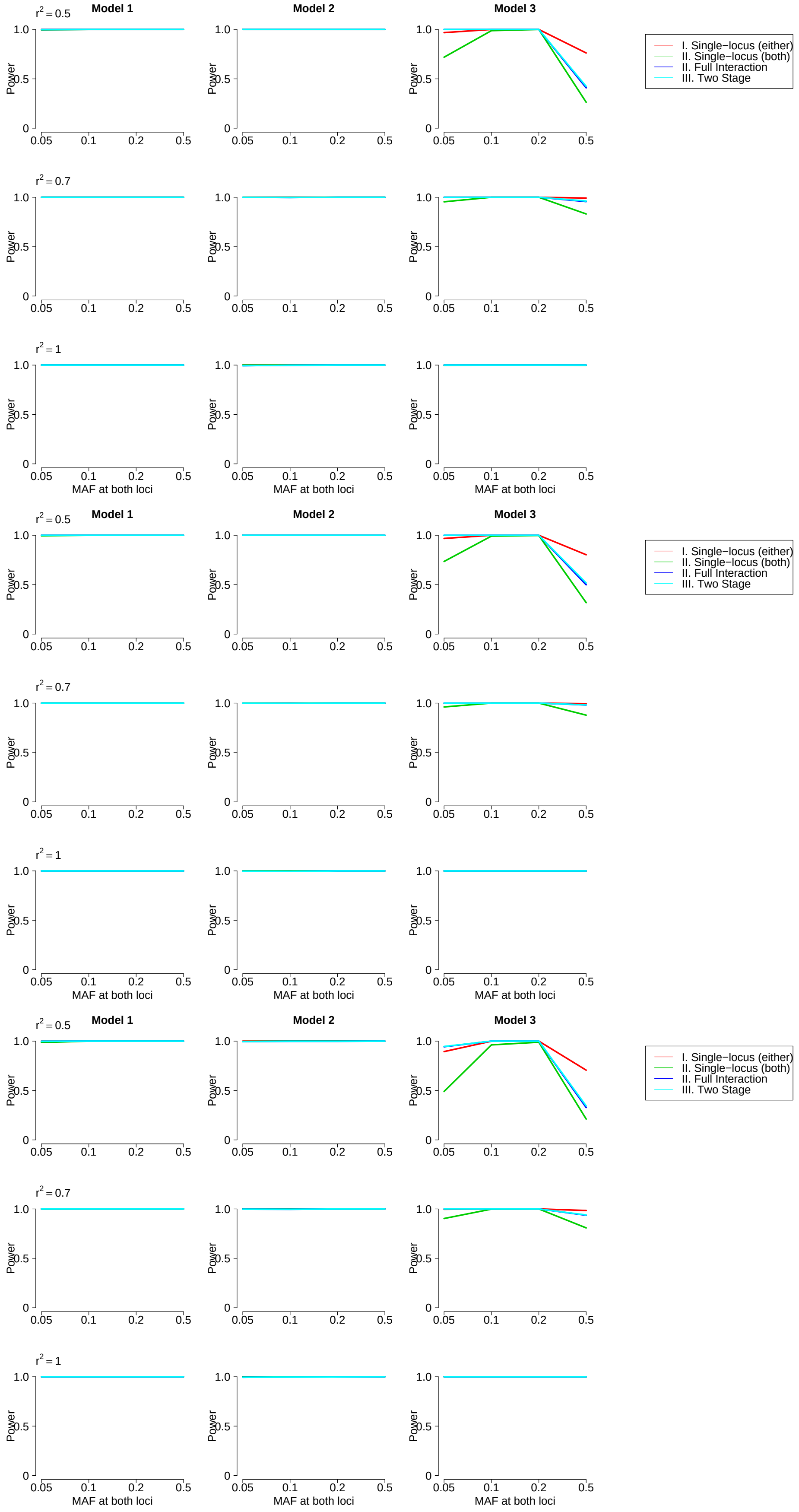

Figure 10: Power of search strategies for parameters $n = 4000, \lambda = 1.0$ and for Constraint 1 (first 9 plots), Constraint 2 (second 9 plots) and Constraint 3 (last 9 plots

.

Computational Feasibility

Fitting interaction models to genome-wide marker data imposes serious computational demands due to the large numbers of tests (Hoh and Ott, 2003; Moore and Ritchie, 2004; Carlson et al., 2004). Feasibility depends crucially on the type of models being fitted and the fine details of the computational calculations. We investigated the computational feasibility of Strategies 2 and 3 within our own model fitting framework. Since these strategies fit the full interaction model for a given pair of loci, the maximum likelihood estimates of the parameters can be obtained in closed form and thus are extremely fast to compute. Using test datasets with 1000 cases and 1000 controls we were able to fit full 2-locus models at a rate of approximately 38,000 per second on a AMD Opteron 244 (1.8GHz) processor using code written in C. At this speed, a dataset consisting of 300,000 typed loci could be searched for all two locus interactions (Strategy II) in approximately 33 hours on 10 processors. Strategy III would require just 19.8 minutes on 10 processors.

Multi-locus models

In order to investigate the power of different strategies for multi-locus models with more than 2 loci we considered the natural 3-locus extensions of our models 1, 2 and 3. More specifically, if we denote the 3-locus genotype as (G_A, G_B, G_C) where $G_A, G_B, G_C \in \{0, 1, 2\}$ are numbers of risk alleles at each disease locus A, B and C then

Model 1 odds$(G_A, G_B, G_C) = \alpha(1+\theta)^{G_A+G_B+G_C}$

Model 2 odds$(G_A, G_B, G_C) = \alpha(1+\theta)^{G_A*I(G_A>0)+G_B*I(G_B>0)+G_C*I(G_C>0)}$

Model 3 odds$(G_A, G_B, G_C) = \alpha(1+\theta)^{I(G_A>0\cap G_B>0\cap G_C>0)}$.

For each of these models we chose the parameters α and θ so that the marginal effect size (λ) was fixed. We then simulated 1000 sets of 3-locus genotypes in cases and controls for all combinations of the following parameter sets : $n \in$

$(1000, 2000, 4000)$, $\pi_A = \pi_B = \pi_C \in (0.05, 0.1, 0.2, 0.5)$ and $\lambda \in (0.2, 0.5, 1.0)$. For each set of these parameters we assessed the power of the following 3 strategies

Strategy I Single locus scan using Bonferroni correction. We measured the power of this strategy in 2 ways : (Ia) Detection of all 3 loci, (Ib) Detection of at least one locus.

Strategy II A 2-stage strategy in which (i) all loci with a marginal p-value less than 0.1 are selected, (ii) a full model for all pairs of loci selected in stage (i) is fitted and the significance is assessed taking into account stage (i). A Bonferroni correction is used to account for the multiple testing. We measured the power of this strategy in 2 ways : (IIa) At least 2 pairs of loci are detected (IIb) At least 1 pair of loci are detected.

Strategy III A 2-stage strategy in which (i) all loci with a marginal p-value less than 0.1 are selected, (ii) a full model for all triples of loci selected in stage (i) is fitted and the significance is assessed taking into account stage (i). A Bonferroni correction is used to account for the multiple testing.

The results of these simulations are shown in Figure 11. As described in the main text we feel that are two questions of interest in this context

- how well do our current one and two locus search strategies perform when there are 3 interacting loci?

- are there any better strategies for uncovering all 3 loci?

To answer the first question it seems appropriate to measure power by requiring that we make any detection and thus the strategies that we need to compare are Ib and IIb. The results are analogous to those for the 2-locus models considered in the main paper in that we see that the single locus search (Ib) does well for Model 1 but the strategy that fits 2-locus models (IIb) has more power for Models 2 and 3.

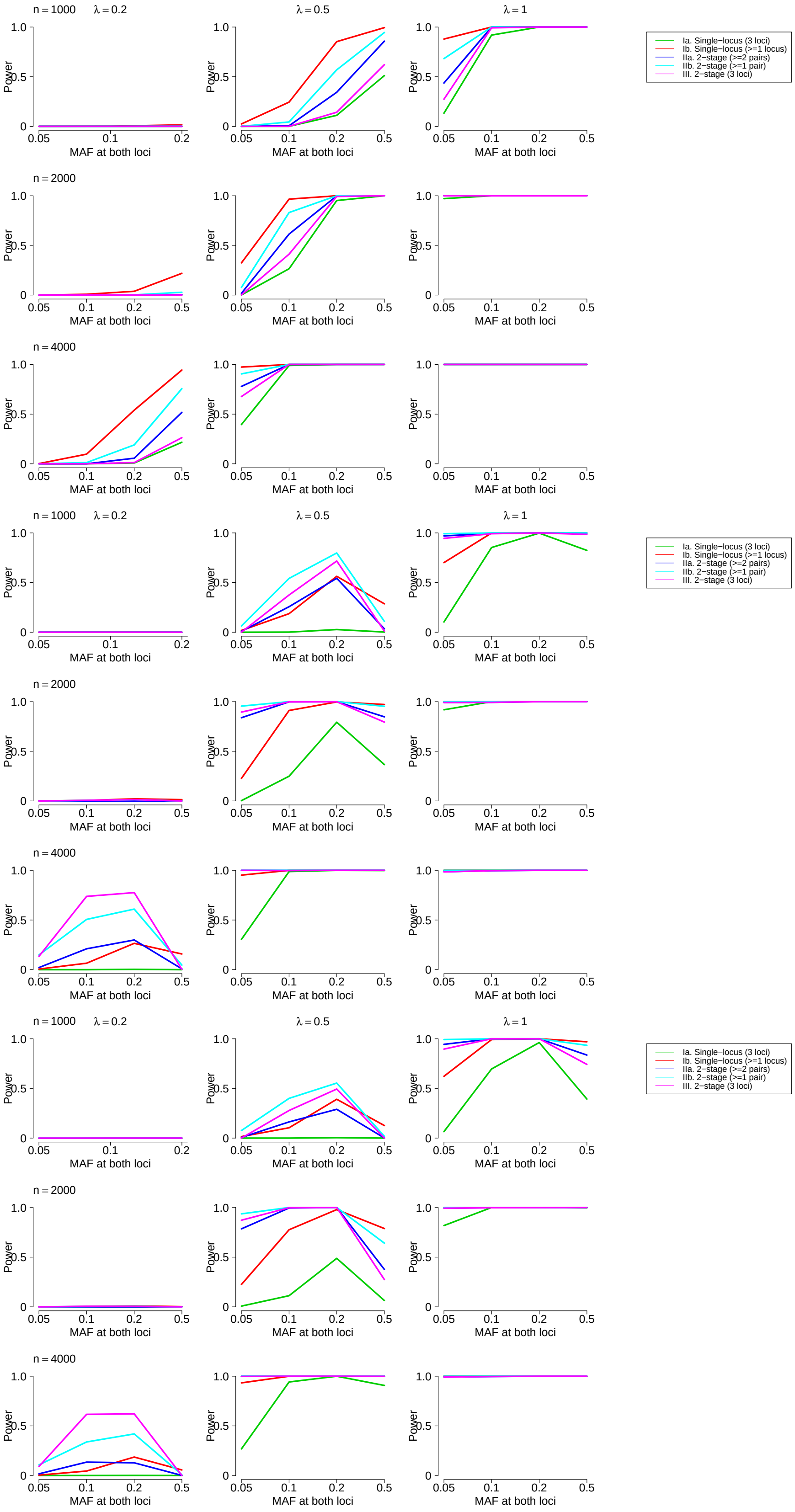

Figure 11: Power of search strategies for 3-locus models : Model 1 (first 9 plots), Model 2 (second 9 plots) and Model 3 (last 9 plots).

The appropriate strategies to compare to answer the second question are Ia, IIa and III. The results indicate that the worst strategy overall is the single locus search strategy Ia. Also, the relative power of the strategies that fit 2 and 3 locus models (IIa and III) depends upon both model and sample size.

References

Balding, D. J. (2003) Likelihood-based inference for genetic correlation coefficients. *Theor Popul Biol*, **63**, 221–230.

Balding, D. J. and Nichols, R. A. (1995) A method for quantifying differentiation between populations at multi-allelic loci and it's implications for investigating identity and paternity. *Genetica*, **96**, 3–12.

Carlson, C. S., Eberle, M. A., Kruglyak, L. and Nickerson, D. A. (2004) Mapping complex disease loci in whole-genome association studies. *Nature*, **429**, 446–452.

Hoh, J. and Ott, J. (2003) Mathematical multi-locus approaches to localizing complex human trait genes. *Nature Reviews Genetics*, **4**, 701–709.

Hudson, R. R. (1990) Gene genealogies and the coalescent process. In *Oxford Surveys in Evolutionary Biology*, vol. 7, 1–44. OUP.

Marchini, J., Cardon, L. R., Phillips, M. S. and Donnelly, P. (2004) The effects of human population structure on large genetic association studies. *Nat Genet*, **36**, 512–517.

Marchini, J. L. and Cardon, L. (2002) Discussion of Nicholson et al. (2002). *JRSS (B)*, **64**, 1–21.

Moore, J. and Ritchie, M. (2004) The challenges of whole-genome approaches to common diseases. *JAMA*, **291**, 1642–3.

Nicholson, G., Smith, A. V., Jobsson, F., Gustfasson, O., Stefansson, K. and Donnelly, P. (2002) Assessing population differentiation and isolation from single nucleotide polymorphism data. *JRSS (B)*, **64**, 695–716.

Pritchard, J. K. and Przeworski, M. (2001) Linkage disequilibrium in humans: models and data. *American Journal of Human Genetics*, **69**, 1–14.