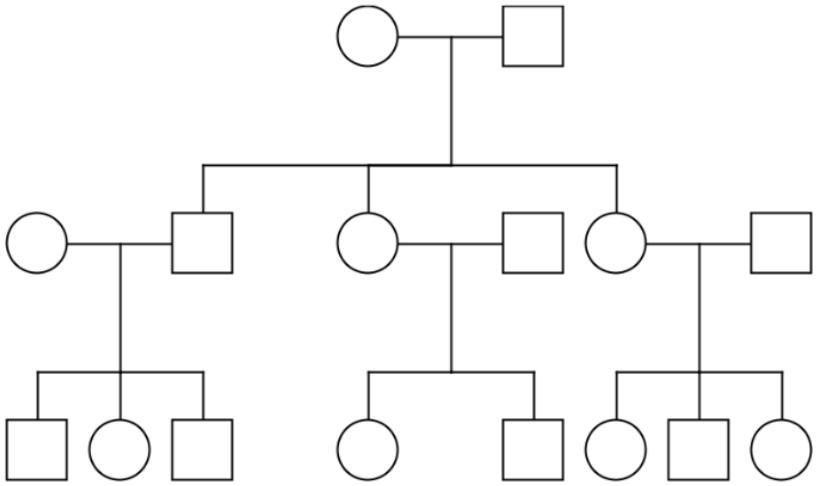

Figure S1. Pedigree Structure Used in Simulation Studies.

Pedigree structure used in simulations in which the study design consists of 60 outbred, three-generation pedigrees.

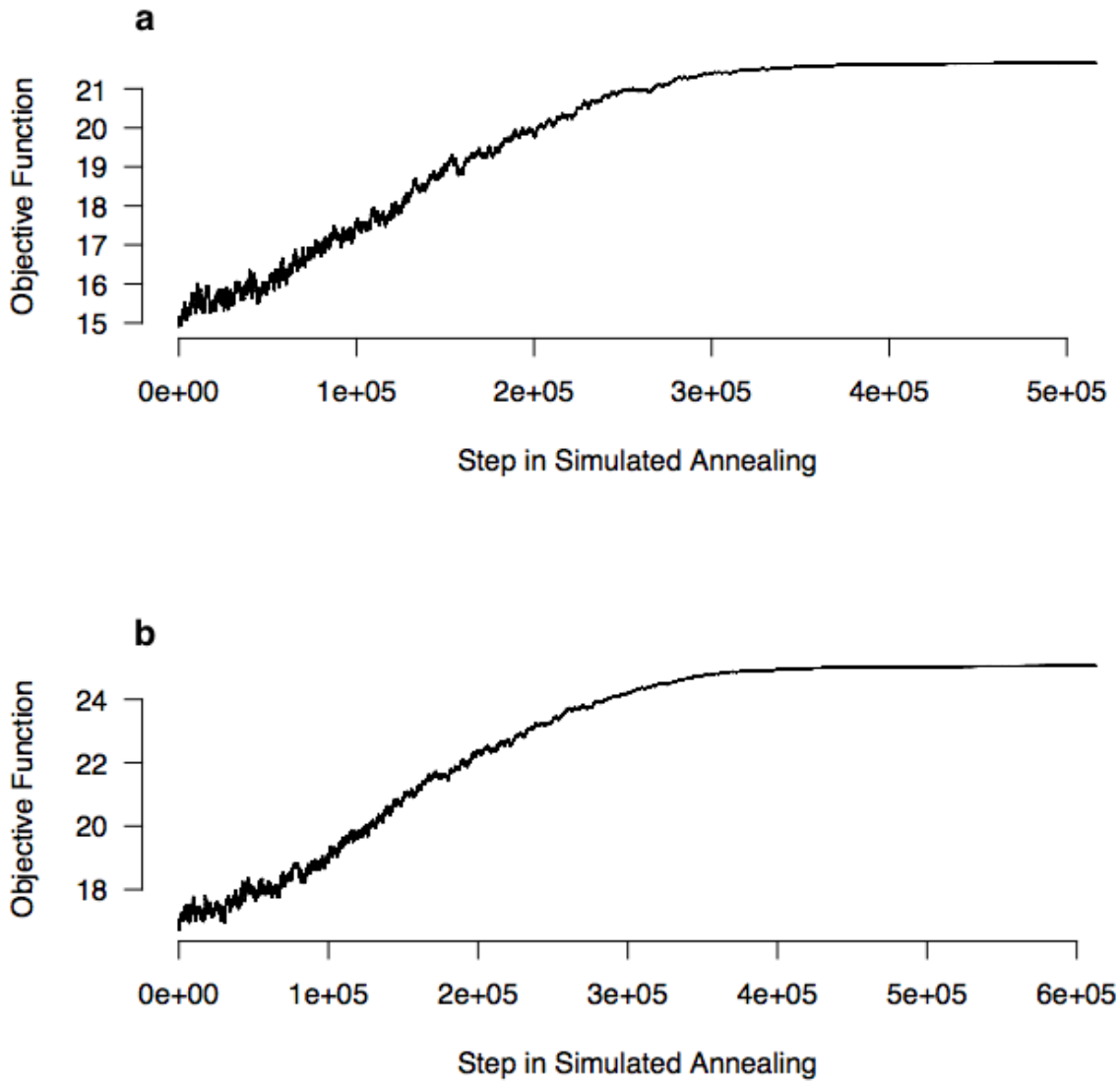

Figure S2. Trajectory Plots of G-STRATEGY Objective Function in Simulated Annealing Runs when Applied to HDL Data.

The AGES-REFINE sample consists of 8,030 individuals, among whom $n_0 = 3{,}134$ have been previously genotyped. Panel (A) is based on a G-STRATEGY run to select $n_a = 500$ additional individuals to genotype; Panel (B) is based on a G-STRATEGY run to select $n_a = 1,000$ additional individuals to genotype.

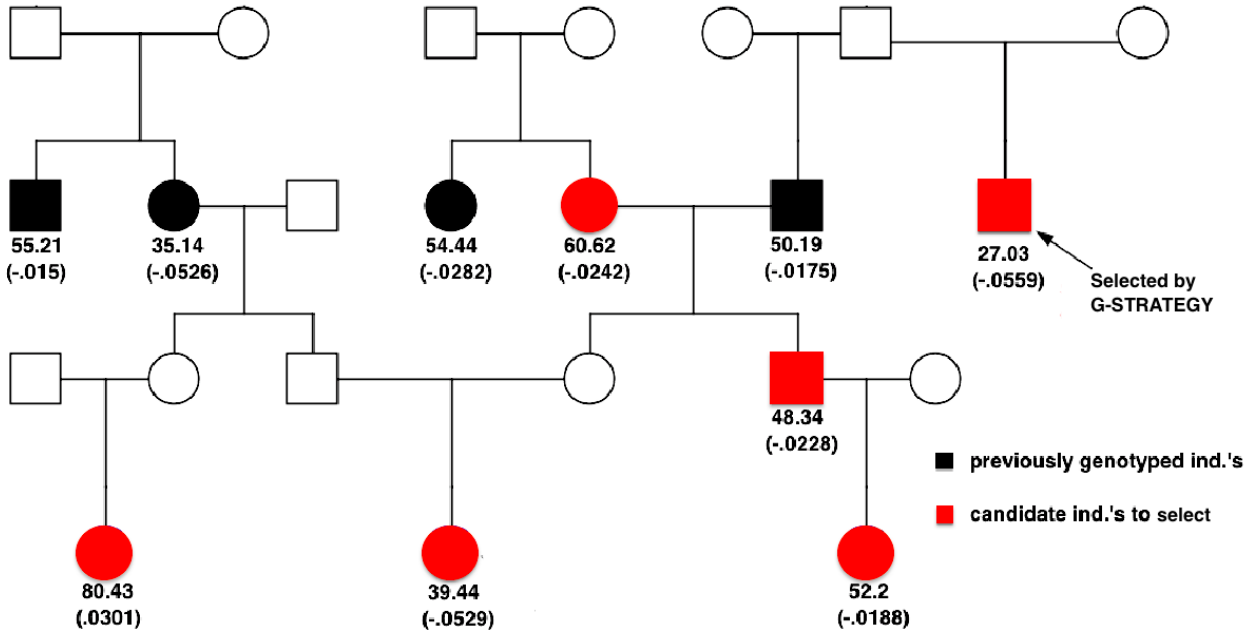

Figure S3. Application of G-STRATEGY to HDL Data: Pedigree 1 when 500 Individuals are Selected.

G-STRATEGY is implemented to select $n_a = 500$ additional individuals to genotype from the AGES-REFINE study. Only one pedigree from the sample is as depicted. Unshaded individuals are neither genotyped nor phenotyped; individuals in black are phenotyped and previously genotyped; individuals in red are phenotyped and available to be selected for additional genotyping. HDL value and enrichment value (in parentheses) are annotated under each phenotyped individual.

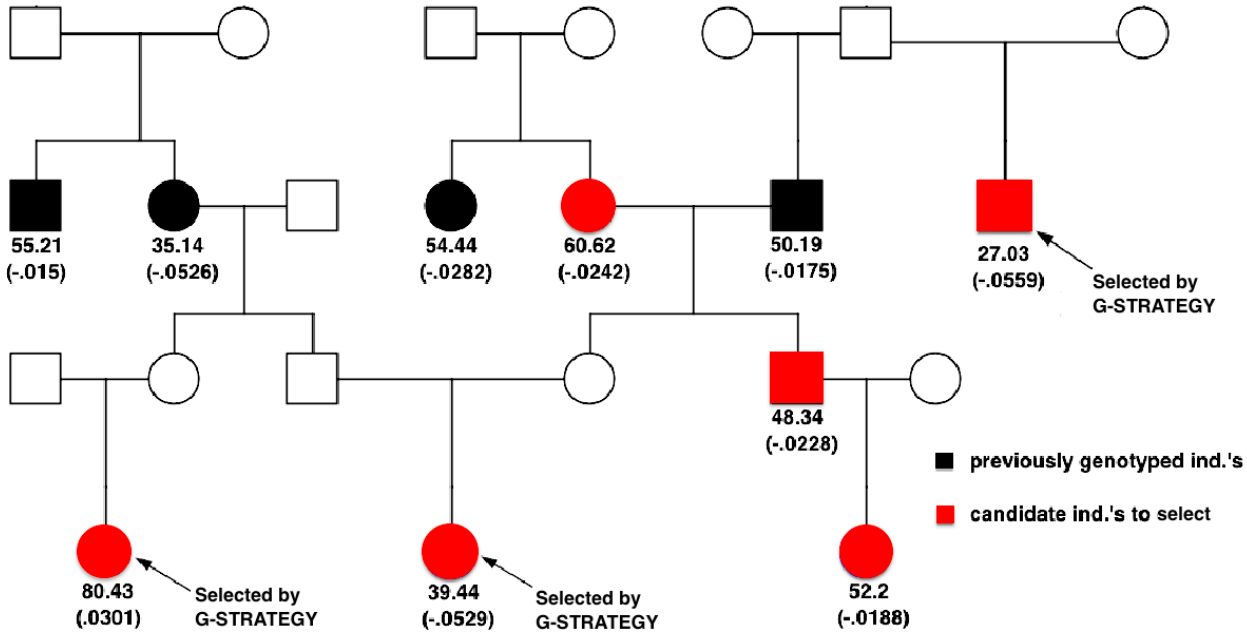

Figure S4. Application of G-STRATEGY to HDL Data: Pedigree 1 when 1,000 Individuals are Selected.

G-STRATEGY is implemented to select $n_a = 1,000$ additional individuals to genotype from the AGES-REFINE study. The same pedigree as in Figure S3 is depicted. Unshaded individuals are neither genotyped nor phenotyped; individuals in black are phenotyped and previously genotyped; individuals in red are phenotyped and available to be selected for additional genotyping. HDL value and enrichment value (in parentheses) are annotated under each phenotyped individual.

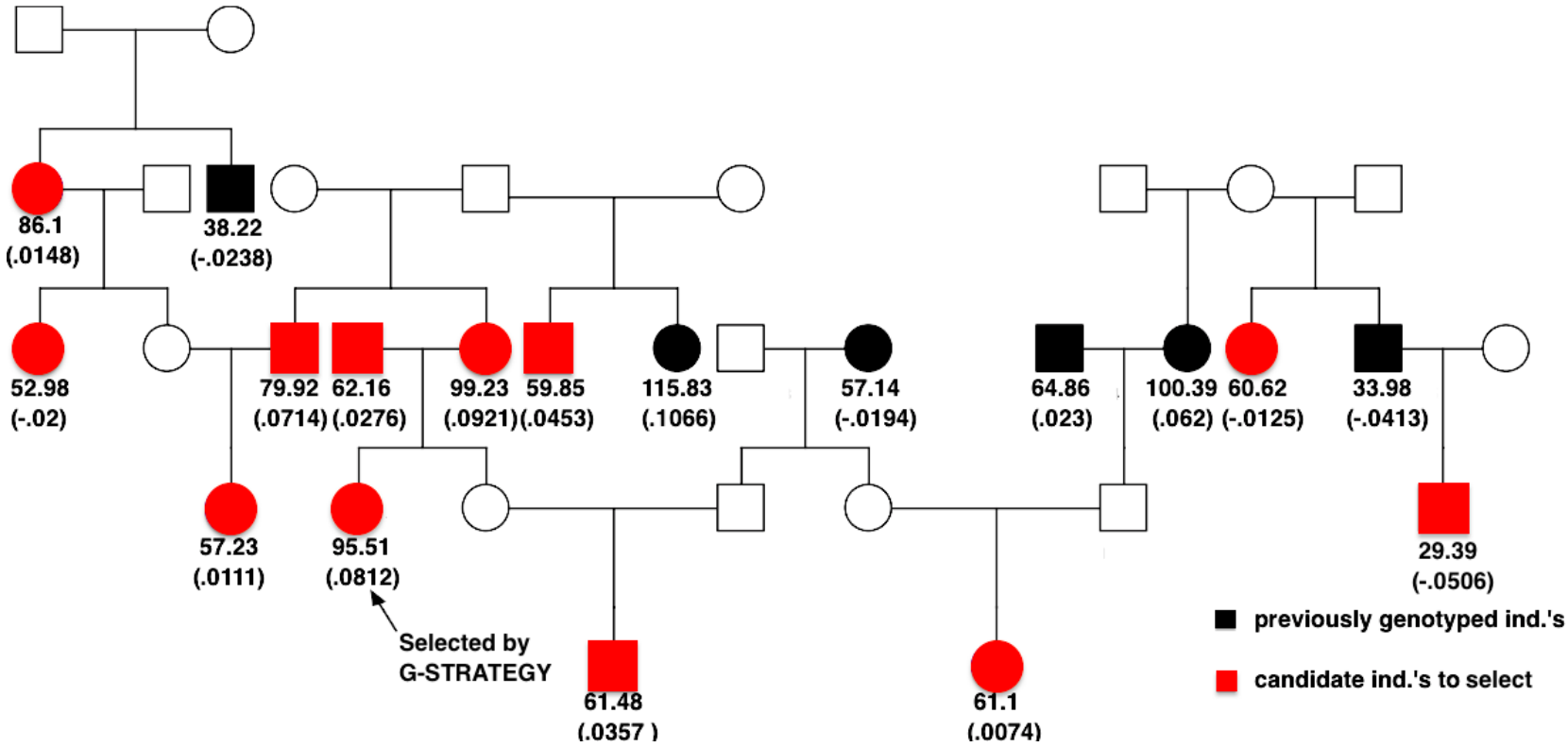

Figure S5. Application of G-STRATEGY to HDL Data: Pedigree 2 when 500 Individuals are Selected.

G-STRATEGY is implemented to select $n_a = 500$ additional individuals to genotype from the AGES-REFINE study. Only one pedigree from the sample is as depicted, different from that in Figures S2 and S3. Unshaded individuals are neither genotyped nor phenotyped; individuals in black are phenotyped and previously genotyped; individuals in red are phenotyped and available to be selected for additional genotyping. HDL value and enrichment value (in parentheses) are annotated under each phenotyped individual.

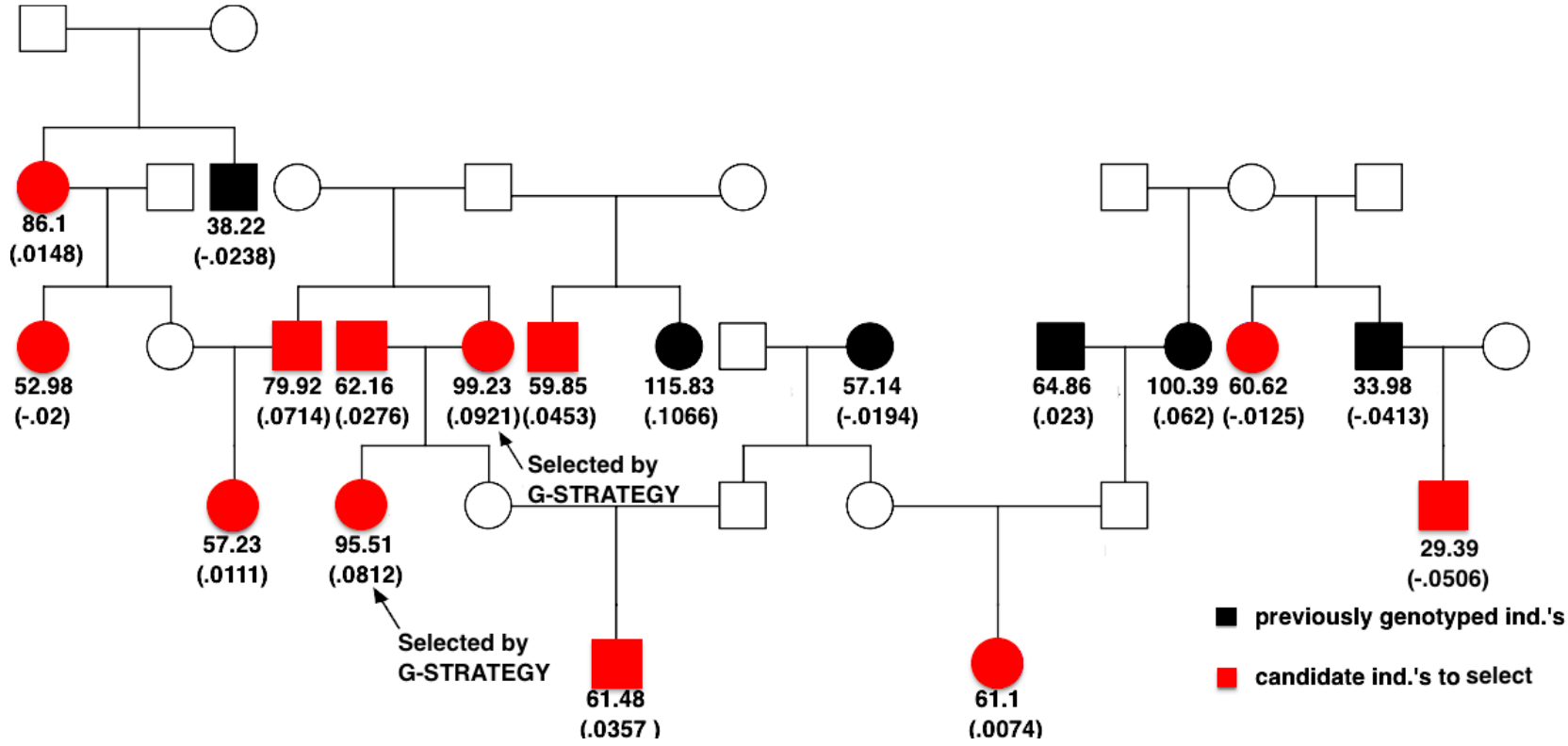

Figure S6. Application of G-STRATEGY to HDL Data: Pedigree 2 when 1,000 Individuals are Selected.

G-STRATEGY is implemented to select $n_a = 1,000$ additional individuals to genotype from the AGES-REFINE study. The same pedigree as in Figure S5 is depicted. Unshaded individuals are neither genotyped nor phenotyped; individuals in black are phenotyped and previously genotyped; individuals in red are phenotyped and available to be selected for additional genotyping. HDL value and enrichment value (in parentheses) are annotated under each phenotyped individual.

Table S1. Allele frequencies and penetrance parameters for simulated binary trait models 1, 2 and 3

Model	p_1	p_2	p_3	f_1	f_2	f_3	f_4	K_p	K_s	λ_s
1a	.1	.5	...	.2	.05	...	...	.071	.087	1.224
1b	.4	.25	...	.35	.20	...	...	.242	.249	1.030
2a	.3	.2	...	.5	.45	.30	.01	.041	.139	3.381
2b	.3	.4	...	.5	.15	.10	.08	.090	.097	1.082
3a	.1	.2	.3	.4	.03	...	...	.078	.156	2.001
3b	.05	.4	.2	.5	.05	...	...	.084	.154	1.844

Note: 1, 2 and 3 indicate the model class (described in the text), and the letter (a or b) that follows the model class is used to distinguish among the different settings of allele frequencies and penetrance parameters used for that model class. p_i is the frequency of allele 1 at SNP i, f_i, $i = 1, 2, 3, 4$ are penetrance parameters described in the text, K_p is the population prevalence of the disease in outbreds, K_s is the probability that an outbred individual is affected, given that the individual has a sibling who is affected, and $\lambda_s = \dfrac{K_s}{K_p}$ is the sibling risk ratio.

Table S2: Parameter settings for simulated quantitative trait models used in simulations

Model	β_0	β_1	p_1	p_2	p_3	p_4	$g(x_4)$			σ_a^2	σ_e^2
							$x_4 = 0$	$x_4 = 0.5$	$x_4 = 1$		
4a	1.5	.5	.15	.2	.3	.1	.1	1.25	1.5	16	49
4b	2	2	.4	.3	.25	.2	.5	.5	1	25	4
5	1	2.5	.2	.1	.3	.3	.5	1	3	25	4
6	1	2.5	.2	.2	.3	.2	.5	1	4	25	4

Note: Models 4a, 4b, 5 and 6 are described in the text. β_0 is the intercept, β_1 is the sex effect, p_i is the frequency of allele 1 at SNP i, $g(x_4)$ is the genotypic effect of SNP 4, where x_4 is 0, .5, or 1 according to whether the individual has 0, 1, or 2 copies of a given allele at locus 4, and σ_a^2 and σ_e^2 are the additive polygenic and environmental variance components, respectively.

Table S3: Combined genotypic effect, $f(x_1, x_2, x_3)$, of SNPs 1, 2, and 3 in the quantitative trait models used in simulations

Model		$x_3 = 0$			$x_3 = .5$			$x_3 = 1$		
		$x_2 = 0$	$x_2 = .5$	$x_2 = 1$	$x_2 = 0$	$x_2 = .5$	$x_2 = 1$	$x_2 = 0$	$x_2 = .5$	$x_2 = 1$
4a	$x_1 = 0$	.5	.5	.5	.75	.75	.75	1	1	1
	$x_1 = .5$	.5	3	3	3.5	3.5	3.5	5	5	5
	$x_1 = 1$	.5	3	3	3.5	3.5	3.5	5	5	5
4b	$x_1 = 0$	.1	.5	1	.25	.5	1	1	1.5	3
	$x_1 = .5$	.2	2	2	.25	3	3	1.5	3.5	6
	$x_1 = 1$	2	2	2.5	3	4	6	3	6	6
5	$x_1 = 0$	.1	.25	1	.2	1.5	2	1.5	2	2
	$x_1 = .5$	.5	1.5	1.5	1.5	2	3.5	3	3.5	4
	$x_1 = 1$	1	2	3.5	2.5	4	5	4	4.5	6
6	$x_1 = 0$	.1	.25	1.5	.2	1.5	2	1.5	3	3.5
	$x_1 = .5$	.5	1.5	1.5	1.5	2	3.5	3	3.5	4
	$x_1 = 1$	1	2	2.5	2	3	4	3	3.5	4

Note: Combined genotypic effect of SNPs 1, 2, and 3 is given as a function, $f(x_1, x_2, x_3)$, where x_i is 0, .5, or 1 according to whether the individual has 0, 1, or 2 copies of a given allele at locus i.