Online Supporting Material

Model Selection

Screening of the SNPs

Two methods were used to screen the single nucleotide polymorphisms (SNP). The first method used was the stepwise backward selection based on Akaike Information Criterion (AIC) and the second method was the Least Angle Regression (LASSO). To use AIC variable selection, we first fitted a preliminary model with only main effects, including blood sample type (whole blood, WB, versus plasma-deplete cells, PDC) used for folate analysis, age, gender, BMI, and 64 SNP, with no interaction terms. Group, gender, and 14 SNPs were selected by AIC stepwise selection and formed the preliminary model 1 (**Supplemental Table S1**). The *P* values were based on two-sided *t*-tests and we also computed *P* values adjusted by the false discovery rate (1, 2).

Supplemental Table 1: Results from the first screening of predictors for main effects using the AIC criterion

Predictor	SNP	Estimate	Standard Error	*t* value	*P* value	Adjusted *P* value[1]
Intercept	-	-5.298	0.136	-38.906	0.000	0.000
SPTLC1	rs11790991	-0.066	0.037	-1.764	0.079	0.115
BCMO1	rs12934922	-0.051	0.032	-1.607	0.109	0.115
CD36	rs1358337	0.051	0.032	1.616	0.107	0.115
MTHFR	rs1801131	0.052	0.034	1.514	0.131	0.131
CRBP2	rs2118981	-0.104	0.032	-3.236	0.001	0.006
ABCA1	rs2230806	0.070	0.034	2.075	0.039	0.087
APOA1	rs2727784	-0.052	0.032	-1.661	0.098	0.115
BCMO2	rs35361223	0.150	0.075	1.994	0.047	0.094
BHMT	rs3733890	0.077	0.035	2.171	0.031	0.079
α-TTP	rs4501570	0.683	0.299	2.288	0.023	0.068
α-TTP	rs4587328	-0.689	0.299	-2.307	0.022	0.068
CETP	rs5882	0.060	0.033	1.793	0.074	0.115
FOLH1	rs61886492	-0.167	0.086	-1.939	0.053	0.096
SCARB1	rs838892	0.054	0.032	1.691	0.092	0.115

[1] *P*-values are adjusted by the false discovery rate (1, 2).

Online Supporting Material

For the LASSO variable selection, we only included 64 SNPs as predictors. We used the ***Lars*** program in R to implement the variable selection (3, 4). The response is log(nHcy), the predictors are 64 SNP, which need to be normalized to mean zero and variance 1. A tuning parameter lambda controls the fraction of non-zero coefficients (number of SNP in the model). The solution path is shown in **Supplemental Table 2**, which indicates the order in which the SNP entered the model when decreasing the tuning parameter lambda. The SNP that enter the model in earlier steps tend to be more important and significant.

Supplemental Table 2: The LASSO path.

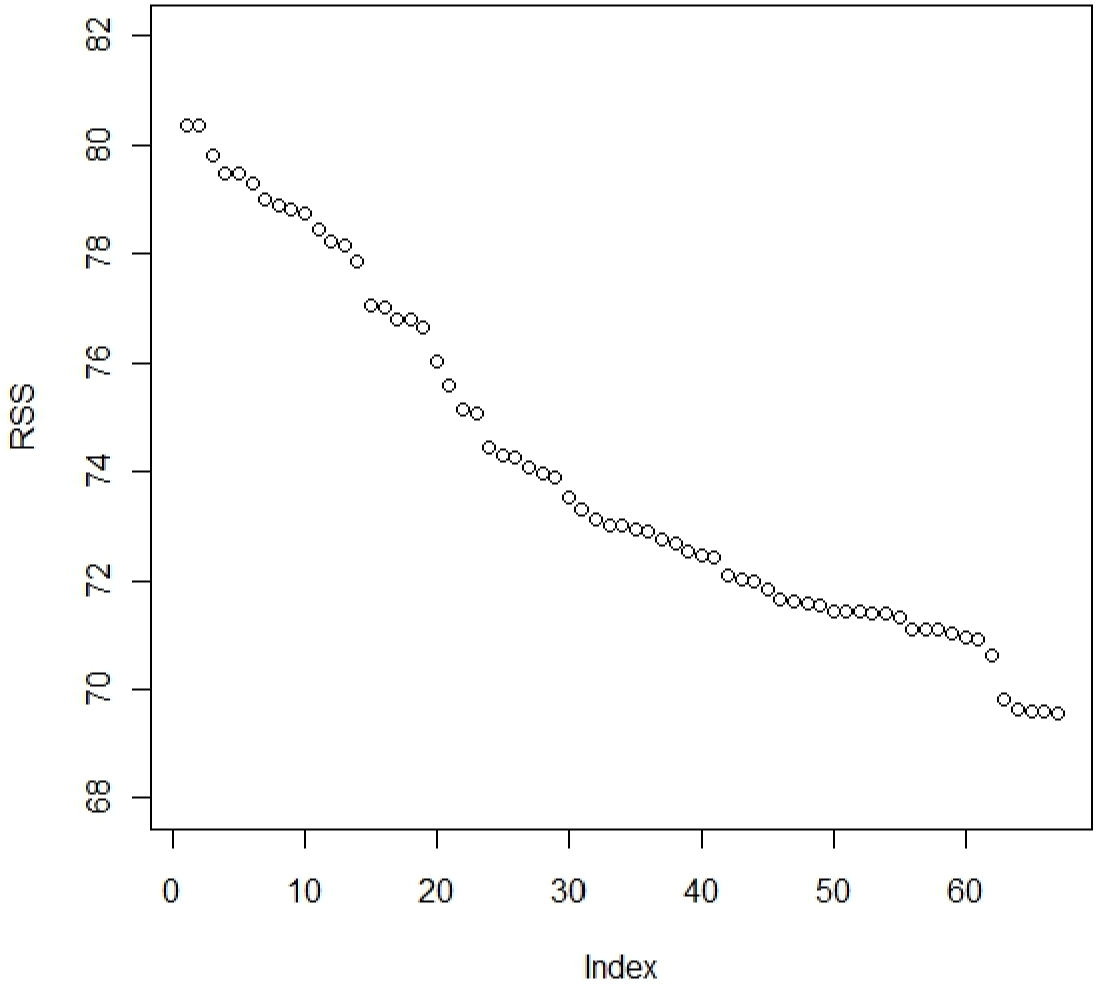
Step	Action	Predictor	SNP	lambda
1	+	*CRBP2*	rs3772875	0.928
2	+	*SPTLC1*	rs11790991	0.919
3	+	*CETP*	rs5882	0.761
4	+	*FOLH1*	rs61886492	0.688
5	+	*BHMT*	rs3733890	0.684
6	+	*BCMO1*	rs12934922	0.662
7	+	*RBP2*	rs2118981	0.624
8	+	*CD36*	rs3211931	0.613
9	+	*SCARB1*	rs838892	0.602
10	+	*SLC46A1*	rs17719944	0.595
11	+	*CYP4F2*	rs2108622	0.571
12	+	*MTHFR*	rs1801131	0.552
13	+	*ABCA1*	rs2230806	0.547
14	+	*APOA1*	rs2727784	0.524
15	+	*BCMO1*	rs7501331	0.468
16	+	*SLC46A1*	rs11080058	0.463
17	+	*SCARB1*	rs7306660	0.450
18	+	*α-TTP*	rs4606052	0.450
19	+	*SCARB1*	rs10773105	0.439
20	+	*BMCO2*	rs35361223	0.399
21	+	*MTR*	rs1805087	0.368
22	+	*SCARB1*	rs7967521	0.340
23	+	*MTTP*	rs3828542	0.334
24	+	*CETP*	rs7205804	0.291
25	+	*TCII*	rs1801198	0.279
26	+	*RBP2*	rs35674260	0.277

27	+	*NPC1L1*	rs217420	0.262
28	+	*CETP*	rs5880	0.252
29	+	*MTRR*	rs1801394	0.247
30	+	*SLC19A1*	Rs111702537	0.219
31	+	*MTHFR*	rs1801133	0.200
32	+	*BCMO2*	rs2250417	0.187
33	+	*CYP4F2*	rs3093194	0.177
34	+	*ABCA1*	rs4149327	0.176
35	+	*SLC46A1*	rs1128162	0.171
36	+	*APOA5*	rs12272004	0.167
37	+	*SPTLC1*	rs7858659	0.156
38	+	*SLC46A1*	rs739439	0.151
39	+	*NPC1L1*	rs217430	0.140
40	+	*SLC19A1*	rs13050920	0.136
41	+	*BCMO1*	rs6564851	0.133
42	+	*NPC1L1*	rs217432	0.108
43	+	*ABCA1*	rs2230808	0.103
44	+	*CETP*	rs7499892	0.102
45	+	*CD36*	rs3211956	0.092
46	+	*SCARB1*	rs12582221	0.077
47	+	*MTTP*	rs1057613	0.074
48	+	*α-TTP*	rs4501570	0.070
49	+	*BCMO2*	rs11214139	0.068
50	+	*CYP4F2*	rs3093168	0.061
51	+	*CBS*	rs5742905	0.061
52	+	*NPC1L1*	rs11763759	0.061
53	+	*SLC19A1*	rs3788199	0.058
54	+	*ABCA1*	rs4149267	0.058
55	+	*CD36*	rs3211834	0.056
56	-	*CYP4F2*	rs3093194	0.045
57	+	*APOA5*	rs3135506	0.045
58	+	*SLC46A1*	rs35714695	0.045
59	+	*SLC19A1*	rs3788205	0.043
60	+	*CD36*	rs1358337	0.041
61	+	*α-TTP*	rs4587328	0.040
62	+	*SPTLC1*	rs2297568	0.036
63	+	*APOA5*	rs662799	0.017
64	+	*MTTP*	rs10516445	0.009
65	+	*CYP4F2*	rs3093194	0.008
66	+	*MTTP*	rs881890	0.006

Online Supporting Material

The sum of residual squares plot for the LASSO method (**Supplemental Figure 1**) indicates that there is a drop in the sum of residual squares at step 15. Therefore, we decided to include the first 15 SNP with corresponding lambda of 0.468 (**Supplemental Table 2**).

Online Supporting Material



Supplemental Figure 1: The sum of residual squares plot for the LASSO method

Online Supporting Material

Adding interactions and fitting the final model

To check for interactions between gender and SNP, we first started with the preliminary model 1 (**Supplemental Table 1**) that was obtained by AIC variable selection. All fourteen predictors in the preliminary model and all interaction terms between gender and fourteen SNP were included before using AIC stepwise selection again. The summary of the results are in **Supplemental Table 3**.

Supplemental Table 3: The summary of the model obtained by adding interactions between gender and 14 SNP to the model screened by AIC.

Predictor	**SNP**	**Estimate**	**Standard Error**	***t* value**	***P* value**	**Adjusted *P* value[1]**
Intercept	-	-4.986	0.092	-54.144	0.000	0.000
Blood groups, WB vs. PDC	-	-0.270	0.054	-5.015	0.000	0.000
SPTLC1	rs11790991	-0.075	0.036	-2.090	0.037	0.069
BCMO1	rs12934922	-0.051	0.031	-1.619	0.106	0.123
MTHFR	rs1801131	-0.033	0.044	-0.759	0.448	0.469
CRBP2	rs2118981	-0.181	0.043	-4.247	0.000	0.000
ABCA1	rs2230806	0.065	0.033	1.973	0.049	0.083
APOA1	rs2727784	-0.099	0.042	-2.342	0.020	0.043
BCMO2	rs35361223	0.119	0.073	1.623	0.106	0.123
BHMT	rs3733890	0.083	0.035	2.386	0.018	0.043
α-TTP	rs4501570	0.602	0.289	2.086	0.038	0.069
α-TTP	rs4587328	-0.543	0.289	-1.881	0.061	0.085
CETP	rs5882	0.111	0.043	2.575	0.010	0.038
FOLH1	rs61886492	-0.145	0.084	-1.733	0.084	0.109
SCARB1	rs838892	0.000	0.044	0.003	0.998	0.998
Gender	-	-0.542	0.122	-4.453	0.000	0.000
Gender × *MTHFR*	rs1801131	0.162	0.068	2.387	0.018	0.043
Gender × *CRBP2*	rs2118981	0.164	0.061	2.669	0.008	0.035
Gender × *APOA1*	rs2727784	0.116	0.062	1.874	0.062	0.085
Gender × *α-TTP*	rs4501570	-0.122	0.063	-1.934	0.054	0.085
Gender × *CETP*	rs5882	-0.094	0.064	-1.459	0.146	0.160

Gender × *SCARB1*	rs838892	0.154	0.063	2.431	0.016	0.043

[1] *P*-values are adjusted by the false discovery rate (1, 2).

The F test for the interaction terms with df = 6 and 350 gave a value of $P < 0.0001$, which indicates that the interaction effects are highly significant. We performed backward selection based on *P* values to obtain the final model. The results of the final model are included in **Table 2** of the manuscript.

We also used the preliminary model (screened by the LASSO method, which includes the first 15 SNP (labeled 1 – 15, **Supplemental Table 2**) as well as gender, blood group (WB versus PDC), and any interaction terms between gender and the 15 SNP. We then used AIC stepwise selection again. The summary of the resulting model is in **Supplemental Table 4**. We further performed backward selection based on *P* values, which led to the same final model as the model obtained when using the preliminary model by AIC selection (**Supplemental Table 1**).

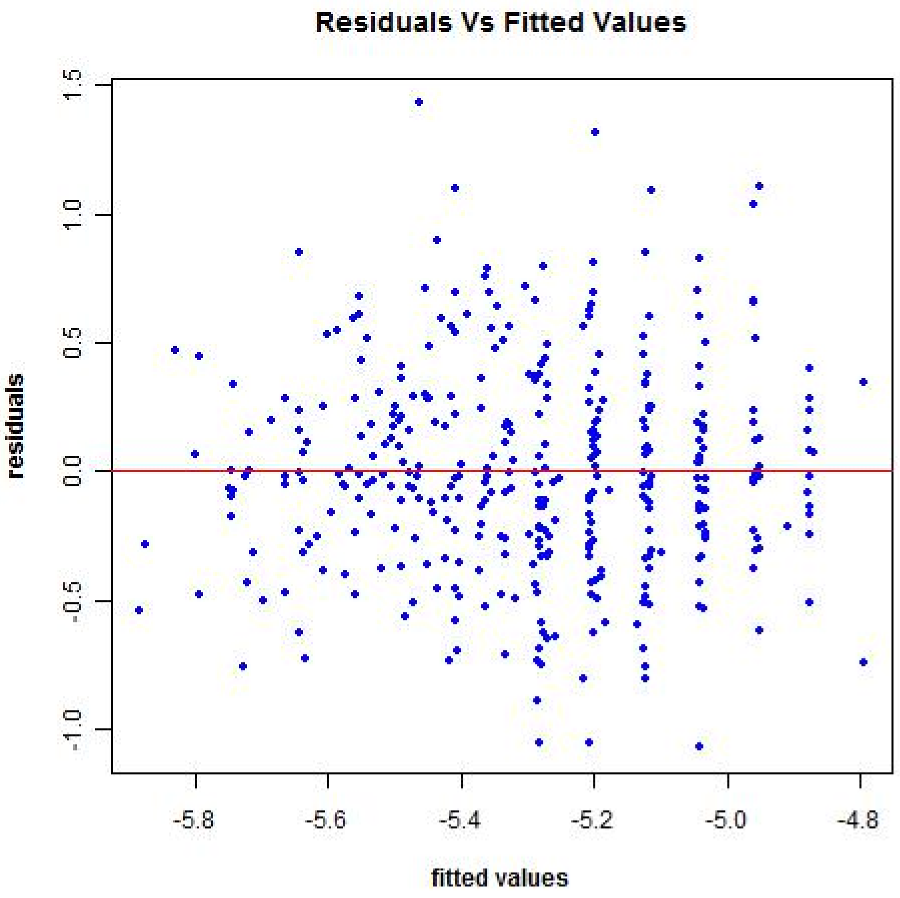
For this final model (**Table 2**, manuscript), the residuals versus fitted values were plotted (**Supplemental Figure 2**) and we also obtained the Q-Q plot of residuals from the final model (**Supplemental Figure 3**). The model assumptions of normality, independence, and homoscedasticity were found not to be seriously violated as indicated by **Supplemental Figure 2** and **Supplemental Figure 3**.

Supplemental Table 4: Summary of the model obtained by adding interactions between gender and 15 SNPs to the model screened by LASSO.

Predictor	**SNP**	**Estimate**	**Standard Error**	***t* value**	***P* value**	**Adjusted *P* value**[1]
Intercept	-	-4.929	0.083	-59.615	0.000	0.000
SPTLC1	rs11790991	-0.081	0.036	-2.226	0.027	0.044
BCMO1	rs12934922	-0.045	0.032	-1.430	0.154	0.173
MTHFR	rs1801131	-0.039	0.044	-0.888	0.375	0.397
CRBP2	rs2118981	-0.173	0.043	-4.026	0.000	0.000
ABCA1	rs2230806	0.069	0.033	2.078	0.038	0.058
APOA1	rs2727784	-0.100	0.043	-2.347	0.019	0.039
BHMT	rs3733890	0.080	0.035	2.289	0.023	0.041
CETP	rs5882	0.113	0.043	2.614	0.009	0.024
FOLH1	rs61886492	-0.149	0.084	-1.761	0.079	0.109
SCARB1	rs838892	-0.008	0.044	-0.183	0.855	0.855
Blood groups, WB vs. PDC	-	-0.254	0.054	-4.720	0.000	0.000
Gender	-	-0.652	0.106	-6.135	0.000	0.000
Gender × *MTHFR*	rs1801131	0.186	0.068	2.748	0.006	0.023
Gender × *CRBP2*	rs2118981	0.161	0.062	2.599	0.010	0.024
Gender × *APOA1*	rs2727784	0.104	0.062	1.693	0.091	0.117
Gender × *CETP*	rs5882	-0.100	0.064	-1.547	0.123	0.147
Gender × *SCARB1*	rs838892	0.163	0.064	2.564	0.011	0.024

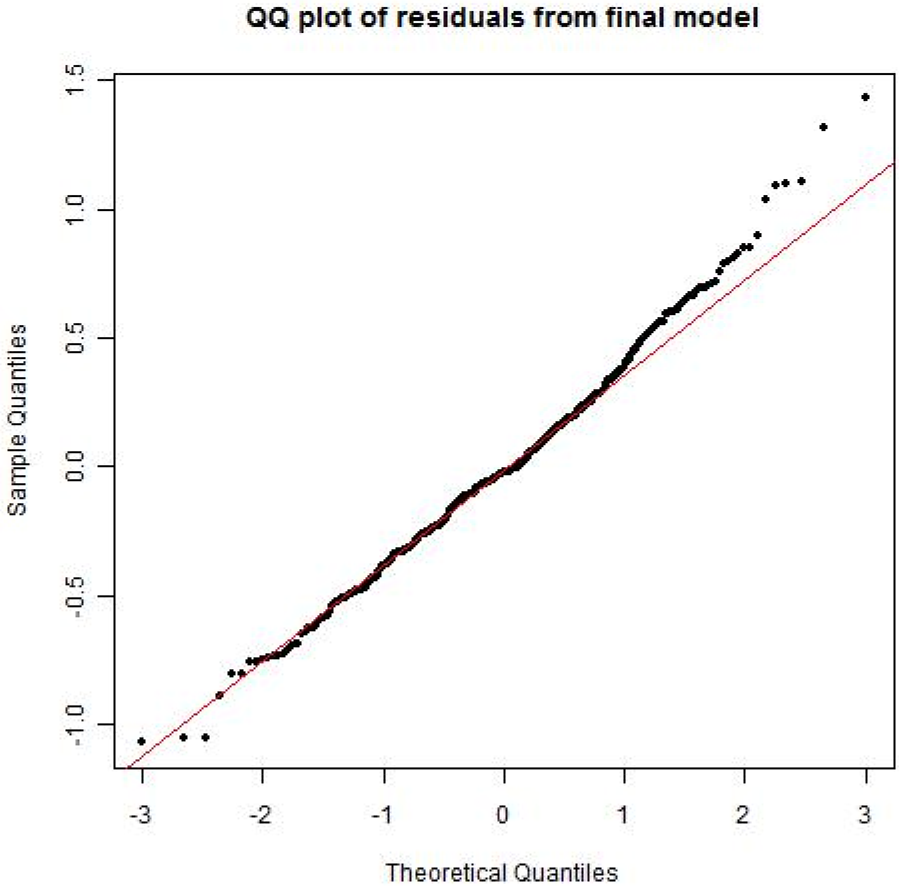
[1] *P*-values are adjusted by the false discovery rate (1, 2).

Online Supporting Material



Supplemental Figure 2: The residual plot for the final model

Online Supporting Material



Supplemental Figure 3: The Q-Q plot of the residuals for the final model

Supplemental Literature Cited

1. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Series B*. 1995;57:239-300.
2. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*. 2001;29:1165-88.
3. Efron B, Hastie T, Johnstone I, Tibshirani R. Least angle regression. Annals of Statistics. 2004;32:407-99.

4. Hastie T, Tibshirani R, Friedman J. Elements of Statistical Learning. New York, NY: Springer; 2002.