Exploring genome-wide – dietary heme iron intake interactions and the risk of type 2 diabetes

Louis R. Pasqualetti, Stephanie J. Loomis, Hugues Aschard, Jae H. Kang, Marilyn C. Cornelis, Lu Qi, Peter Kraft, and Frank B. Hu

1 Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
2 Glaucoma Service, Massachusetts Eye and Ear Infirmary, Boston, MA, USA
3 Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA
4 Department of Nutrition, Harvard School of Public Health, Boston, MA, USA

Aims/hypothesis: Genome-wide association studies have identified over 50 new genetic loci for type 2 diabetes (T2D). Several studies conclude that higher dietary heme iron intake increases the risk of T2D. Therefore we assessed whether the relation between genetic loci and T2D is modified by dietary heme iron intake.

Methods: We used Affymetrix Genome-Wide Human 6.0 array data [681,770 single nucleotide polymorphisms (SNPs)] and dietary information collected in the Health Professionals Follow-up Study (n = 725 cases; n = 1,273 controls) and the Nurses' Health Study (n = 1,081 cases; n = 1,692 controls). We assessed whether genome-wide SNPs or iron metabolism SNPs interacted with dietary heme iron intake in relation to T2D, testing for associations in each cohort separately and then meta-analyzing to pool the results. Finally, we created 1,000 synthetic pathways matched to an iron metabolism pathway on number of genes, and number of SNPs in each gene. We compared the iron metabolic pathway SNPs with these synthetic SNP assemblies in their relation to T2D to assess if the pathway as a whole interacts with dietary heme iron intake.

Results: Using a genomic approach, we found no significant gene–environment interactions with dietary heme iron intake in relation to T2D at a Bonferroni corrected genome-wide significance level of 7.33 × 10^{-8}. Furthermore, no SNP in the iron metabolic pathway significantly interacted with dietary heme iron intake at a Bonferroni corrected significance level of 2.10 × 10^{-4}. Finally, neither the main genetic effects nor gene–dietary heme–iron interactions (pooled empirical p-value for the interactions = 0.72) were significant for the iron metabolic pathway as a whole.

Conclusions: We found no significant interactions between dietary heme iron intake and common SNPs in relation to T2D.

Keywords: type 2 diabetes, gene environment interactions, dietary heme iron, pathway analysis

INTRODUCTION

Type 2 diabetes (T2D) is a multifactorial condition whereby insulin resistance and beta-cell dysfunction produce glucose metabolism alterations, most notably hyperglycemia, resulting in microvascular and macrovascular complications. T2D affects over 25 million individuals (greater than 8% of the U.S. adult population; American Diabetes Association, 2011). Discovery of the combination of genetic and environmental factors contributing to T2D is essential so that more targeted preventive and management strategies can be devised.

Abbreviations: 1-df test, one degree of freedom gene–environment interaction test; 2-df test, two degree of freedom gene–environment interaction test; HH, hereditary hemochromatosis; HPFS, Health Professionals Follow-Up Study; NHS, Nurses' Health Study; QC, quality control SNP; single nucleotide polymorphism; T2D, type 2 diabetes.
obesity, excess iron stores produced by dietary behaviors may render pancreatic beta cells particularly vulnerable to oxidative stress because they possess a weak antioxidative stress system (Tiedge et al., 1997).

While dietary heme iron intake plays a role in T2D, genome-wide association studies have revealed approximately 50 genetic loci for T2D (Vocche et al., 2012), raising the question of whether gene environment interactions focused on dietary heme iron intake exist in T2D.

In fact, a previous candidate gene association study demonstrated a significant interaction between dietary iron intake and the H63D or C282Y risk variants for HFE, the gene for hereditary hemochromatosis (HH), in relation to T2D risk (Qi et al., 2005). HH is a condition where secondary diabetes mellitus is a well-known complication that results from high iron stores in the pancreas (Utzschneider and Kowdley, 2010). In this work, we assessed whether the relation between dietary heme iron intake and T2D is modified by genome-wide single nucleotide polymorphisms (SNPs) or iron metabolic pathway SNPs.

### MATERIALS AND METHODS

#### STUDY POPULATION

This study included participants from two longitudinal cohort studies, the Nurses’ Health Study (NHS) and the Health Professionals Follow-up Study (HPFS; Barton et al., 1980; Rimm et al., 1990). The HPFS and the NHS are two populations of men and women, respectively, for whom stored blood, DNA samples, and dietary heme iron intake data are available. The NHS began in 1976 with 121,700 female registered nurses aged 30–55, and the HPFS started in 1986 with 51,529 male health professionals aged 40–75. Both cohorts have been followed up biennially through mailed questionnaires that gather data on new diseases, diet, and other lifestyle factors.

#### ASCERTAINMENT OF TYPE 2 DIABETES:

For participants with a self-report of diabetes on biennial questionnaires, we mailed supplemental questionnaires inquiring about the diagnosis and treatment of their condition, as well as a history of ketoacidosis to corroborate the self-report and to differentiate between type 1 diabetes mellitus and T2D. We transformed heme to be centered on the mean value for each cohort. We also tested models that treated the dietary heme iron intake term as a dichotomous variable divided at the median value in controls.

#### DATA ANALYSIS

To assess the interaction between dietary heme iron intake and gene variants in relation to T2D, we created the following logistic regression models:

1. $T2D = \beta_0 + \beta_1 (age) + \beta_2 (body\ mass\ index, BMI) + \beta_3 (heme)$ – Environmental model
2. $T2D = \beta_0 + \beta_1 (SNP) + \beta_2 (age) + \beta_3 (BMI) + \beta_4 (heme)$ – Genetic model
3. $T2D = \beta_0 + \beta_1 (SNP) + \beta_2 (age) + \beta_3 (BMI) + \beta_4 (heme) + \beta_5 (SNP \times heme)$ – one-degree of freedom gene-environment (GxE) interaction model

For Model 1 (which we generated in SAS version 9.3, Cary, NC, USA), we formulated these models in each cohort separately using PLINK. Heme iron intake (mg/day), BMI (kg/m²), as of 1986 for HPFS and 1980 for NHS and age (years, as of 1986 for HPFS and 1990 for NHS) were treated as continuous variables. We transformed heme to be centered on the mean value for each cohort. SNPs were coded as 0, 1, or 2 minor alleles. We controlled for the top three eigenvectors in NHS and the top four eigenvectors in HPFS. We also tested models that treated the dietary heme iron intake term as a dichotomous variable divided at the median value in controls.

To test for interactions between dietary heme intake and SNP genotypes in relation to T2D, we utilized a one degree of freedom

### Table 1 | Characteristics of Health Professionals Follow-up Study (HPFS) and Nurses’ Health Study (NHS) cohorts.

<table>
<thead>
<tr>
<th>Cohort</th>
<th>N</th>
<th>Mean heme, SD (mg/day)</th>
<th>Mean body mass index, SD (kg/m²)</th>
<th>Mean age at baseline, SD (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPFS Cases/Controls</td>
<td>19687251273</td>
<td>129.0 (43.1</td>
<td>89.43.123.0 42)</td>
<td>25.05 (3.462</td>
</tr>
<tr>
<td>NHS Cases/Controls</td>
<td>277310811692</td>
<td>136.0 (45.4</td>
<td>1.41.44.133.45)</td>
<td>26.08 (5.0028.66</td>
</tr>
</tbody>
</table>

SD, standard deviation.
We identified 237 SNPs in genes coding for enzymes in this path- 
mean dietary heme iron intake than controls in men and women 
(1-df) test, which is represented as $\beta_3$ of Model 3. Models 1 and 2 
tested the marginal effects of dietary heme iron intake and SNPs 
respectively, adjusting for age and BMI. We also performed a two 
degree of freedom joint test (2-df) by comparing the fit of the null 
model containing the environment exposure only (Model 1) to 
the model with gene and gene–environment covariates (Model 3) 
as an alternate test of the gene–environment interaction (Cornelis 
et al., 2011; Manning et al., 2011). After running these models sep- 
ately in NHS and HFFPS, we performed tests for heterogeneity of 
the cohort specific results to check for appropriateness of pooling 
the data. We conducted an inverse variance-weighted fixed-effects 
meta-analysis of estimates from the two cohorts using the METAL 
software. Only SNPs with genotypes available in both cohorts 
were included in the meta-analysis (N = 683,770).

Next we limited our gene–environment interaction analyses 
to SNPs in the iron metabolic pathway created using the KEGG 
database and other sources (Michal, 1999; Andrews and Schmidt, 
We identified 237 SNPs in genes coding for enzymes in this path- 
way that were present on the Affymetrix 6.0 platform (Table 2). We 
repeated the logistic regression analysis of Models 2 and 3 using 
the iron metabolic pathway SNPs. We used a Bonferroni correc- 
tion based on the number of SNPs analyzed to establish statistical 
significance in genome-wide ($p = 0.05/181,770 = 7.33 \times 10^{-8}$) 
and pathway analyses ($p = 0.05/237 = 2.10 \times 10^{-4}$). These esti- 
mates of the correction for multiple comparisons are somewhat 
liberal in that they do not account for the secondary analyses we 
performed. To evaluate whether the iron metabolic pathway as a whole 
might interact with dietary heme intake, we assessed whether the 
iron metabolic pathway SNP panel was enriched with variants 
strongly associated with T2D. We compared the distribution of 
p-values in the SNPs from the iron metabolic pathway with the 
distribution of SNPs from 1,000 permuted “synthetic pathways” 
generated by randomly picking SNPs available on the platform 
that were not in the iron metabolic pathway. To ensure an adequate 
comparison, all synthetic pathways were constructed such that 
they had the same number of genes and the same number of SNPs 
per gene ± 10% as the heme pathway. Enrichment of SNPs with 
low p-values was evaluated by comparing $C_{237}$, the count of SNPs with 
a p-value below a given significance threshold $T$ in the heme 
metabolic pathway, to $C_{237}$, the corresponding count derived in 
the synthetic pathways. The empirical p-value (for both the main 
genetic effect and the gene–environment interaction term) related 
to enrichment for a given $T$ was derived as the number of times 
$C_{237}$ was higher than $C_{237}$ divided by 1,000, the total number of 
synthetic pathways. 

RESULTS 
Type 2 diabetes cases were similar in age compared to controls 
in men (overall mean ± SD = 53.3 ± 8.4 years) and women (47.5 ± 6.8 years). As expected, cases had higher BMI and higher 
mean dietary heme iron intake than controls in men and women 
(Table 1). Dietary heme iron intake was adversely associated 
with T2D (OR = 1.36 (1.17, 1.58); pooled $p = 7.51 \times 10^{-7}$; 
Model 1). As expected, the top SNP associated with T2D 
was in TCF7L2 (rs7901695; pooled $p$-value = $1.88 \times 10^{-14}$) 
(Model 2). 

Using the 1-df test (Model 3), no gene–environment interac- 
tion achieved genome-wide significance level. The most signif- 
ificant interaction with continuous dietary heme iron intake was 
rs10988058 (pooled $p = 1.03 \times 10^{-6}$; an intergenic SNP between 
muscle, skeletal, receptor tyrosine kinase (MUSK) and Sushi, von 
Willebrand factor type A, EGF, and pentraxin domains-containing 
1 (SVEP1) Table 3). The 2-df test revealed that SNPs in TCF7L2 
had genome-wide margin association with T2D but did not reveal 
new marginal gene effects of genome-wide significance; nor was 
there significant interaction between TCF7L2 and dietary heme 
iron intake in T2D (data not shown). When we generated mod- 
els substituting dietary heme intake with red meat, processed 
meat, and total meat, we found similar results with top marginal 
genetic effects in TCF7L2; yet, Model 3 did not yield significant 
gene–environment interactions (data not shown).

No significant iron metabolism SNP – dietary heme iron intake 
interaction was detected with the 1-df test (Model 3) in relation 
to T2D (top SNP rs1805313; in ALAD (delta-aminolevulinic 
dehydrogenase); pooled $p = 1.14 \times 10^{-7}$; Bonferroni corrected 
significance level $p = 2.10 \times 10^{-4}$). The 2-df test of gene and 
genome–environment interactions also did not reveal any signifi- 
cant interactions between dietary heme iron intake and SNPs 
in the iron metabolic pathway SNP in pooled analyses (data 
not shown).

Compared with synthetic pathways, the iron metabolic pathway 
was not associated with T2D when we performed the analyses by 
SNP (pooled empirical $p$ by SNP = 0.41). Similar null results were 
obtained when interactions with dietary heme iron intake were 
considered (pooled empirical $p$-value for the interactions =0.72). 
Interactions between various forms of dietary meat intake and the 
iron metabolic pathway were also not significant. 

DISCUSSION 
Neither the 1-df test nor the 2-df test revealed any genome- 
wide significant interactions between dietary heme iron intake 
and genomic SNPs in T2D. Furthermore, the relation between an 
iron metabolic pathway SNP panel and T2D was not modified by 
dietary heme iron intake. Finally the iron metabolic pathway 
was not enriched with SNPs related to T2D.

There could be several possible reasons for the null results 
reported here. First, despite its large size ($n = 4,771$) our study 
could be underpowered to find modest interaction terms. In fact, 
only in a log additive model would we achieve 80% power to detect 
a genome-wide environmental interaction effect of 1.8 (assumes 
minor allele frequency = 0.4; genetic relative risk = 1.2, and rela-
tive risk of the highest tertile of dietary heme iron intake = 1.3). 
 Nonetheless we had ~80% power to discover an interaction effect 
of 1.5 between iron metabolic SNPs and dietary heme iron intake 
in relation to T2D using similar assumptions in a dominant 
haplotype model. Second, power could be compromised due 
to inherent error in measuring dietary heme iron intake. Third, 
self-reported iron intake, while collected using a validated food 
frequency questionnaire, could be prone to recall bias. Finally, the
<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosome</th>
<th>SNPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACO1</td>
<td>24</td>
<td>rs1028932, rs10435792, rs10318318, rs13293491, rs10318318, rs2026738, rs7032871, rs10738885, rs16918276, rs3780473, rs3780474, rs4495514, rs13302577, rs7866419, rs7022554, rs4879683, rs10579061, rs7033149, rs10709732, rs10709794, rs7015925, rs1235240, rs1233491, rs3780474</td>
</tr>
<tr>
<td>ALAD</td>
<td>7</td>
<td>rs1177504, rs1888884, rs226189, rs1877612, rs1695313, rs1695316</td>
</tr>
<tr>
<td>ALAS1</td>
<td>6</td>
<td>rs527169, rs386722, rs386726, rs1322164, rs392279, rs6813468</td>
</tr>
<tr>
<td>ALB</td>
<td>8</td>
<td>rs2034557, rs489761, rs699510, rs491362, rs489765, rs10233887, rs10268545, rs1315716</td>
</tr>
<tr>
<td>ALBRA</td>
<td>11</td>
<td>rs2613843</td>
</tr>
<tr>
<td>CP</td>
<td>17</td>
<td>rs1776798, rs3816893, rs7372906, rs6881598, rs16861634, rs16861550, rs4934389, rs7071565, rs3628238, rs13072552, rs16861557, rs378047</td>
</tr>
<tr>
<td>HAMP</td>
<td>5</td>
<td>rs8101606</td>
</tr>
<tr>
<td>HFE</td>
<td>3</td>
<td>rs1800562, rs2071303, rs2858996</td>
</tr>
<tr>
<td>HFE2</td>
<td>2</td>
<td>rs10218795, rs7540883</td>
</tr>
<tr>
<td>HMBS</td>
<td>7</td>
<td>rs1799993, rs1784304, rs1006195, rs494046, rs114401, rs17075, rs6498693</td>
</tr>
<tr>
<td>HMOX1</td>
<td>7</td>
<td>rs9306300, rs1406699, rs1403700, rs2281126, rs6985907, rs2266533</td>
</tr>
<tr>
<td>HMOX2</td>
<td>16</td>
<td>rs4785600, rs1605651, rs1362626, rs10326781, rs3786324, rs748651, rs2270366, rs10505325, rs8063864, rs4786569, rs993475, rs7105251, rs8063552, rs6555519</td>
</tr>
<tr>
<td>IL6</td>
<td>7</td>
<td>rs2069832, rs2066995, rs206895, rs2426572, rs2245678, rs1474947</td>
</tr>
<tr>
<td>IL6R</td>
<td>16</td>
<td>rs11265618, rs4945626, rs4537546, rs1072561, rs4245611, rs1126616, rs6466816, rs4846618, rs10192386, rs6686893, rs1206250, rs1239262, rs1922292, rs4575245, rs105752641</td>
</tr>
<tr>
<td>IREB2</td>
<td>18</td>
<td>rs13110, rs268491, rs920421, rs248440, rs1748392, rs1696899, rs2650671, rs1696898, rs17483721, rs1009542, rs8043272, rs10510198, rs7381446, rs5668483, rs2650673, rs8044016, rs11636431, rs25864952</td>
</tr>
<tr>
<td>SLC11A2</td>
<td>8</td>
<td>rs1712512, rs2286693, rs2245672, rs2245678, rs2245689</td>
</tr>
<tr>
<td>SLC25A2</td>
<td>33</td>
<td>rs7652641, rs24857176, rs78695032, rs2942194, rs78695394, rs7033754, rs7084538, rs17085533, rs11758797, rs2508672, rs732778, rs7946025, rs3763032, rs7834883, rs7830209, rs2920665, rs9168284, rs10505725, rs7868981, rs9246204, rs2974747, rs2974745, rs17089332, rs17182122, rs7051818, rs12544753, rs7629004, rs13269590, rs17089358, rs2046444, rs4971981, rs4971980, rs7684536</td>
</tr>
<tr>
<td>SLC40A1</td>
<td>10</td>
<td>rs4667282, rs304704, rs1439816, rs3052373, rs1231110, rs3790207, rs11565835, rs13439138, rs1121639, rs13404407</td>
</tr>
<tr>
<td>SMAD4</td>
<td>5</td>
<td>rs7344817, rs10522913, rs6892790, rs12457540, rs9463886</td>
</tr>
<tr>
<td>STAT1</td>
<td>16</td>
<td>rs8109279, rs3052690, rs4766544, rs1780802, rs3052690, rs1362626, rs2026738, rs7032871, rs10738885, rs16918276, rs3780473, rs3780474, rs4495514, rs13302577, rs7866419, rs7022554, rs4879683, rs10579061, rs7033149, rs10709732, rs10709794, rs7015925, rs1235240, rs1233491, rs3780474</td>
</tr>
<tr>
<td>TFR2</td>
<td>1</td>
<td>rs4521655</td>
</tr>
<tr>
<td>TFRC</td>
<td>6</td>
<td>rs3933, rs3914411, rs3904412, rs4927866, rs4867860, rs12332245</td>
</tr>
<tr>
<td>UROS</td>
<td>12</td>
<td>rs123348, rs5014097, rs1572197, rs2281598, rs10794025, rs3814663, rs1244653, rs1006460, rs2027515, rs3497019, rs7071553, rs2149009, rs12251135</td>
</tr>
</tbody>
</table>

Table 2 | Genes and single nucleotide polymorphisms (SNPs) in the heme iron metabolic pathway on the Affymetrix 6.0 array that passed quality control.
**Table 3** | Top 10 p-values from genome-wide SNP-heme interactions predicting type 2 diabetes adjusting for age, body mass index, dietary hem iron intake, and eigenvectors using a degree of freedom test from meta-analysis of the Health Professionals Follow-up Study and Nurses’ Health Study.1

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>SNP</th>
<th>Beta, p-value*</th>
<th>Beta, p-value*</th>
<th>p-Value, pooled**</th>
<th>p-Value, pooled**</th>
<th>Associated gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>rs10865508</td>
<td>0.96</td>
<td>0.64</td>
<td>1.81 x 10^-4</td>
<td>1.23 x 10^-3</td>
<td>1.03 x 10^-3</td>
</tr>
<tr>
<td>9</td>
<td>rs12378245</td>
<td>0.96</td>
<td>0.61</td>
<td>1.77 x 10^-4</td>
<td>1.76 x 10^-3</td>
<td>1.51 x 10^-3</td>
</tr>
<tr>
<td>9</td>
<td>rs1817049</td>
<td>0.92</td>
<td>0.61</td>
<td>3.09 x 10^-4</td>
<td>1.61 x 10^-3</td>
<td>2.16 x 10^-3</td>
</tr>
<tr>
<td>9</td>
<td>rs7048110</td>
<td>0.96</td>
<td>0.58</td>
<td>1.80 x 10^-4</td>
<td>2.75 x 10^-3</td>
<td>2.53 x 10^-3</td>
</tr>
<tr>
<td>9</td>
<td>rs1817052</td>
<td>0.90</td>
<td>0.60</td>
<td>4.43 x 10^-5</td>
<td>1.93 x 10^-3</td>
<td>3.53 x 10^-5</td>
</tr>
<tr>
<td>16</td>
<td>rs7177078</td>
<td>1.48</td>
<td>0.83</td>
<td>1.50 x 10^-4</td>
<td>5.66 x 10^-3</td>
<td>5.06 x 10^-3</td>
</tr>
<tr>
<td>9</td>
<td>rs10980495</td>
<td>0.91</td>
<td>0.57</td>
<td>3.60 x 10^-4</td>
<td>3.33 x 10^-3</td>
<td>5.49 x 10^-3</td>
</tr>
<tr>
<td>7</td>
<td>rs7048127</td>
<td>−0.58</td>
<td>−0.44</td>
<td>7.85 x 10^-4</td>
<td>2.11 x 10^-3</td>
<td>6.26 x 10^-3</td>
</tr>
<tr>
<td>9</td>
<td>rs1448627</td>
<td>0.94</td>
<td>0.64</td>
<td>2.58 x 10^-4</td>
<td>4.95 x 10^-3</td>
<td>6.58 x 10^-3</td>
</tr>
<tr>
<td>22</td>
<td>rs470089</td>
<td>−0.59</td>
<td>−0.55</td>
<td>2.90 x 10^-3</td>
<td>9.43 x 10^-4</td>
<td>8.64 x 10^-4</td>
</tr>
</tbody>
</table>

1Meta controls for BMI (kg/m^2) and age (years), both as continuous variables, continuous mg dietary heme intake per day, centered on the mean heme value for each cohort and eigenvectors 1-3 for NHLS or eigenvectors 1-4 for HPFS. SNPp are coded as G1 or 2 minor alleles.
2Beta and p-values are from the β1 (gene variant) × dietary heme iron intake term in Model 3. T2D = β1 (BMI) + β2 (age) + β3 (dietary heme iron intake) + β4 (gene variant) + ε. Gene × dietary heme iron intake = eigenvectors.
3p for heterogeneity > 0.05 (lowest p-value is 0.4).
4HPFS, Health Professionals Follow-up Study; NHLS, Nurses’ Health Study; MUSK, muscle, skeletal, receptor tyrosine kinase; SVEP1, Sushi, von Willebrand factor type A, EGF, and pentraxin domains-containing 1; TP53, tumor protein p53 gene; AGX, anterior gradient 3; AGR2, anterior gradient 2; SULT4A1, sulfotransferase family 4A, member 1.

reason why a higher dietary heme iron intake increases the risk of T2D could be solely related to environmental influences (Lee et al., 2004).

A prior study, using a candidate approach, did find a marginal interaction between HFE and dietary heme iron intake in T2D (p = 0.03; Qi et al., 2005). While we did not discover any new gene–dietary heme interactions, more studies using serum biomarkers as surrogates of dietary heme iron intake might point to new gene–iron intake interactions in T2D.

**ACKNOWLEDGMENTS**

This work was supported by grants CA87969, CA49449, CA053075, EY09611, EVO15473, DK58843, UO1 HG004728 and U10HG004728-0251 from the National Institutes of Health.

**REFERENCES**

2115.
Visscher, P. M., Brown, M. A., McCarthy, M. I., and Yang, J. (2012). Five years of GWAS dis-
Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any com-
mercial or financial relationships that could be construed as a potential con-
flict of interest.