Haplotype Structure of the ENPP1 Gene and Nominal Association of the K121Q Missense Single Nucleotide Polymorphism With Glycemic Traits in the Framingham Heart Study

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.2337/db08-0266</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:4745730">http://nrs.harvard.edu/urn-3:HUL.InstRepos:4745730</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Haplotype Structure of the ENPP1 Gene and Nominal Association of the K121Q Missense Single Nucleotide Polymorphism With Glycemic Traits in the Framingham Heart Study

Elliot S. Stolerman, Alisa K. Manning, Jarred B. McAteer, Josée Dupuis, Caroline S. Fox, L. Adrienne Cupples, James B. Meigs, and Jose C. Florez

OBJECTIVE—A recent meta-analysis demonstrated a nominal association of the ectonucleotide pyrophosphatase phosphodiesterase 1 (ENPP1) K→Q missense single nucleotide polymorphism (SNP) at position 121 with type 2 diabetes. We set out to confirm the association of ENPP1 K121Q with hyperglycemia, expand this association to insulin resistance traits, and determine whether the association stems from K121Q or another variant in linkage disequilibrium with it.

RESEARCH DESIGN AND METHODS—We characterized the haplotype structure of ENPP1 and selected 39 tag SNPs that captured 96% of common variation in the region (minor allele frequency ≥5%) with an r2 value ≥0.80. We genotyped the SNPs in 2,511 Framingham Heart Study participants and used age- and sex-adjusted linear mixed effects (LME) models to test for association with quantitative metabolic traits. We also examined whether interaction between K121Q and BMI affected glycemic trait levels.

RESULTS—The Q allele of K121Q (rs1044498) was associated with increased fasting plasma glucose (FPG), A1C, fasting insulin, and insulin resistance by homeostasis model assessment (HOMA-IR; all P = 0.01–0.006). Two noncoding SNPs (rs7775386 and rs7773477) demonstrated similar associations, but LME models indicated that their effects were not independent from K121Q. We found no association of K121Q with obesity, but interaction models suggested that the effect of the Q allele on FPG and HOMA-IR was stronger in those with a higher BMI (P = 0.008 and 0.01 for interaction, respectively).

CONCLUSIONS—The Q allele of ENPP1 K121Q is associated with hyperglycemia and insulin resistance in whites. We found an adiposity-SNP interaction, with a stronger association of K121Q with diabetes-related quantitative traits in people with a higher BMI. Diabetes 57:1971–1977, 2008

Ectonucleotide pyrophosphatase phosphodiesterase 1 (ENPP1), also known as plasma cell membrane glycoprotein 1 (PC-1), is a transmembrane glycoprotein that down-regulates insulin signaling in cells by inhibiting the tyrosine kinase activity of the insulin receptor, perhaps by interaction with its α-subunit (1). Within the coding region of ENPP1, a K→Q missense single nucleotide polymorphism (SNP) at position 121 (K121Q; rs1044498) has been previously associated with insulin resistance and related abnormalities in some studies (2–7). The molecular mechanism thought to be responsible for the role of the Q121 variant is a “gain of function” of the ENPP1 protein inhibitory activity on the insulin receptor (8). It has also been reported that insulin receptor autophosphorylation in fibroblasts is decreased in Q allele carriers compared with KK homozygotes (2). Thus, the ENPP1 gene is considered to be a likely candidate gene for insulin resistance and type 2 diabetes (9).

Multiple studies have shown both positive and negative evidence of association between variants in ENPP1 and obesity, type 2 diabetes, and related traits. Most recently, Meyre et al. (5) found that a haplotype formed by three SNPs in ENPP1 (one of which was K121Q) was associated with childhood and adult obesity and type 2 diabetes. Three subsequent large association studies of variants in ENPP1 detected no association with type 2 diabetes or obesity: Grarup et al. (10) found no association of K121Q with type 2 diabetes in a Danish population; Lyon et al. (11) found no significant association between three ENPP1 SNPs (K121Q, rs1799774, and rs7754561) and BMI or type 2 diabetes; and Weedon et al. (12) found no association with variants in ENPP1 and type 2 diabetes or obesity in a study involving 8,089 subjects in the U.K. In a comprehensive meta-analysis, we have recently shown that the ENPP1 K121Q variant confers a modestly increased risk of type 2 diabetes under a recessive genetic model in whites (P = 0.005), an effect that appears to be modulated by BMI (13). Although these studies were informative, only a handful of SNPs were genotyped, and common variation in the ENPP1 locus has not been examined comprehensively. Given the conflicting body of evidence in the literature and incomplete evaluation of the ENPP1 gene, we set out to confirm the association of ENPP1 K121Q with hyperglycemia, expand
We targeted a genomic region from \(\text{ENPP1} \rightarrow H11006\) to \(\text{ENPP1} \rightarrow H11350\). We first downloaded Phase 2 HapMap (\(\text{www.hapmap.org}\)) genotypes for the CEU population (accessed January 2006). We genotyped the previously reported non-HapMap SNP rs1790774 in the HapMap CEU samples to include it in its haplotype structure. We examined these 167 HapMap SNPs in Haploview (\(\text{http://www.broad.mit.edu/edu/mpg/haploview/}\)) and found that 95 passed quality control filters, including genotyping call rate >75%, Hardy-Weinberg equilibrium (\(P > 0.001\)), and absence of Mendelian errors. We attempted to capture all remaining working variants with minor allele frequency (MAF) >5% using Tagger (\(\text{http://www.broad.mit.edu/mpg/tagger}\)), using a pairwise approach, setting an \(r^2 > 0.8\), and forcing in the three previously reported SNPs (K121Q, rs1790774, and rs7773477). This procedure yielded 39 SNPs tagging the region, including rs1790774. We genotyped these 39 SNPs (and an additional SNP, rs7773477, whose MAF is 4.5% in Framingham and is located at the 3’ intron-exon junction of exon 2) in the FHS samples. Because seven SNPs failed genotyping in these initial FHS samples, we repeated the tagging procedure forcing in all successful 32 SNPs as tags: this yielded seven new SNPs, which were genotyped in the FHS samples for a total of 40 SNPs (39 tags and rs7773477).

**Genotyping.** Genotyping was performed by allelic-specific primer extension of multiplex-amplified products with detection by matrix-assisted laser desorption ionization–time of flight mass spectroscopy on an iPLEX Sequenom platform. Average genotyping call rates were 97.7%, the minimum call rate was 94.7%, and the average consensus rate based on 254 duplicate samples was 99.5%.

**Statistical analysis.** The quantitative traits were regressed against covariates in order to produce Studentized residuals. Two models were used: the first with sex, age, and age\(^2\) covariate adjustment, and the second added BMI (kg/m\(^2\)) to the other covariates to examine the strength of the subsequent SNP associations when adjusted for obesity. We adjusted for both age and age\(^2\) to allow for a nonlinear trend over time. The covariates from offspring exam 5 were used for all the traits except for mean FPG, for which we used the 16-year average of each covariate.

The association between each individual and each SNP was assessed using a linear mixed effects (LME) model implemented in SOLAR (\(\text{H11350}\)) to account for the within-family correlation. Each SNP was included in a model as a fixed effect with additive coding, although additional dominant and recessive coding were evaluated to examine the association of rs1044498 (K121Q) with the quantitative traits. The models included random effects to account for the covariance between family members; the covariance structure was determined by the degree of relationship between each relative pair.

The role of BMI in the association between rs1044498 (K121Q) and the quantitative traits was further examined in LME models of the age-, age\(^2\), and sex-adjusted residual with K121Q, BMI, and an interaction term between K121Q and BMI as covariates.

To assess the association of each SNP with the type 2 diabetes phenotype, we used Cox proportional hazards survival analysis with diabetes as the outcome and the survival time as the age at the exam in which diabetes was determined. The survival time of individuals without diabetes was the age at their last exam. The model was implemented with the survival package in R (\(\text{H11350}\)), with the same adjustments as in the LME models with covariates taken at the first exam. Trait correlation among siblings was modeled with a frailty term in the survival model (\(\text{H11350}\)).

To assess whether previous association signals were due to linkage disequilibrium with K121Q or were independent, we added the SNPs to LME models already containing K121Q. If the signals were independent of K121Q, we expected that they would remain significant in these models. Alternatively, if both K121Q and the other SNPs became nonsignificant, we would conclude that the signal in the other SNPs was not independent from K121Q.

Our study was formulated around a single primary hypothesis: we intended to replicate the association of K121Q with hyperglycemia as captured by diabetes-related traits and, if such association was confirmed, perform further covariate adjustment for improved characterization and additional fine-mapping to determine the true source of the association signal. We have thus tested a unique SNP for association with several components of one composite trait (hyperglycemia), which is reflected in multiple correlated measures. We believe this primary hypothesis should not be seen as making multiple unrelated comparisons, and we therefore chose a nominal \(P\) value of 0.05 to indicate statistical significance.

**RESULTS**

A linkage disequilibrium plot showing the completed haplotype structure of the \(\text{ENPP1}\) locus is presented in the Supplemental Figure in the online appendix, which is available at http://dx.doi.org/10.2337/db08-0260. The \(\text{ENPP1}\) gene region contains a high degree of linkage disequilibrium. From the initial set of 167 SNPs, 96% of the 95 common variants with a MAF >5% were captured with an \(r^2\) value \(\geq 0.8\) (and 100% with an \(r^2\) value \(= 0.7\)) by a set of 39 tag SNPs using single-marker (pairwise) tests. The list of the 39 successful tag SNPs used in this analysis, with chromosomal position, major allele, and MAF in both the HapMap and Framingham population, is presented in Table 2. The list of 95 captured SNPs with their tags is listed in the online appendix (Supplemental Table 1).

We examined whether individual SNPs in \(\text{ENPP1}\) were associated with hyperglycemic and insulin resistance traits in the FHS population. We first focused our attention on our principal polymorphism of interest, rs1044498.
using three different genetic models: additive, dominant, and recessive. Table 3 presents mean trait levels for each genotypic group, with P values for association with K121Q before and after adjustment for BMI. Several associations with insulin resistance traits reached nominal levels of significance: specifically, under the additive model, the Q allele was associated with higher FPG (P = 0.01), A1C (P = 0.006), fasting insulin (P = 0.006), and HOMA-IR (P = 0.006); all of these associations remained significant after adjusting for BMI. Similar P values were obtained under the dominant genetic model. In this population sample, there were no differences in mean trait value across genotypic groups at ENPP1 K121Q for BMI (P = 0.32) or waist circumference (P = 0.64), both tested as continuous traits. When we compared the distribution of genotypes at this locus across individuals who were of normal weight (BMI <25 kg/m²), overweight (BMI 25–30 kg/m²), or obese (BMI >30 kg/m²), we found no association of K121Q with obesity as a categorical trait (P = 0.66). We also found no significant deviation from the null hypothesis of no association when our cohort was divided by BMI cutoffs at 25, 30, and 35 kg/m² (data not shown).

We then explored whether these consistent associations might be driven by other polymorphisms in the region. Table 4 displays the ENPP1 SNPs for which we detected nominally significant associations with any of the glycemic traits under study. Figure 1A (glucose-related traits) and 1B (insulin-related traits) display the P values for rs1044498 (K121Q) in comparison with those obtained for other SNPs across the genomic segment. Although K121Q was the SNP that showed the strongest association with fasting insulin and HOMA-IR (Fig. 1B), two other SNPs (rs7775386 and rs7773477, located in intron 1 and at the exon 2-intron 2 junction, respectively) achieved similar P values for associations with FPG and A1C; these two SNPs were also nominally associated with HOMA-IR (Table 4). To determine whether the effects of rs7775386 and rs7773477 were independent from K121Q or produced an association signal simply because they are in tight linkage disequilibrium with K121Q, we calculated the r² among those SNPs (K121Q), using three different genetic models: additive, dominant, and recessive. Table 3 presents mean trait levels for each genotypic group, with P values for association with K121Q before and after adjustment for BMI. Several associations with insulin resistance traits reached nominal levels of significance: specifically, under the additive model, the Q allele was associated with higher FPG (P = 0.01), A1C (P = 0.006), fasting insulin (P = 0.006), and HOMA-IR (P = 0.006); all of these associations remained significant after adjusting for BMI. Similar P values were obtained under the dominant genetic model. In this population sample, there were no differences in mean trait value across genotypic groups at ENPP1 K121Q for BMI (P = 0.32) or waist circumference (P = 0.64), both tested as continuous traits. When we compared the distribution of genotypes at this locus across individuals who were of normal weight (BMI <25 kg/m²), overweight (BMI 25–30 kg/m²), or obese (BMI >30 kg/m²), we found no association of K121Q with obesity as a categorical trait (P = 0.66). We also found no significant deviation from the null hypothesis of no association when our cohort was divided by BMI cutoffs at 25, 30, and 35 kg/m² (data not shown).

We then explored whether these consistent associations might be driven by other polymorphisms in the region. Table 4 displays the ENPP1 SNPs for which we detected nominally significant associations with any of the glycemic traits under study. Figure 1A (glucose-related traits) and 1B (insulin-related traits) display the P values for rs1044498 (K121Q) in comparison with those obtained for other SNPs across the genomic segment. Although K121Q was the SNP that showed the strongest association with fasting insulin and HOMA-IR (Fig. 1B), two other SNPs (rs7775386 and rs7773477, located in intron 1 and at the exon 2-intron 2 junction, respectively) achieved similar P values for associations with FPG and A1C; these two SNPs were also nominally associated with HOMA-IR (Table 4). To determine whether the effects of rs7775386 and rs7773477 were independent from K121Q or produced an association signal simply because they are in tight linkage disequilibrium with K121Q, we calculated the r² among those SNPs (K121Q), using three different genetic models: additive, dominant, and recessive. Table 3 presents mean trait levels for each genotypic group, with P values for association with K121Q before and after adjustment for BMI. Several associations with insulin resistance traits reached nominal levels of significance: specifically, under the additive model, the Q allele was associated with higher FPG (P = 0.01), A1C (P = 0.006), fasting insulin (P = 0.006), and HOMA-IR (P = 0.006); all of these associations remained significant after adjusting for BMI. Similar P values were obtained under the dominant genetic model. In this population sample, there were no differences in mean trait value across genotypic groups at ENPP1 K121Q for BMI (P = 0.32) or waist circumference (P = 0.64), both tested as continuous traits. When we compared the distribution of genotypes at this locus across individuals who were of normal weight (BMI <25 kg/m²), overweight (BMI 25–30 kg/m²), or obese (BMI >30 kg/m²), we found no association of K121Q with obesity as a categorical trait (P = 0.66). We also found no significant deviation from the null hypothesis of no association when our cohort was divided by BMI cutoffs at 25, 30, and 35 kg/m² (data not shown).

We then explored whether these consistent associations might be driven by other polymorphisms in the region. Table 4 displays the ENPP1 SNPs for which we detected nominally significant associations with any of the glycemic traits under study. Figure 1A (glucose-related traits) and 1B (insulin-related traits) display the P values for rs1044498 (K121Q) in comparison with those obtained for other SNPs across the genomic segment. Although K121Q was the SNP that showed the strongest association with fasting insulin and HOMA-IR (Fig. 1B), two other SNPs (rs7775386 and rs7773477, located in intron 1 and at the exon 2-intron 2 junction, respectively) achieved similar P values for associations with FPG and A1C; these two SNPs were also nominally associated with HOMA-IR (Table 4). To determine whether the effects of rs7775386 and rs7773477 were independent from K121Q or produced an association signal simply because they are in tight linkage disequilibrium with K121Q, we calculated the r² among those SNPs.
in a subset of unrelated FHS participants. The $r^2$ between K121Q and rs7775386 was 0.664, and between K121Q and rs7773477 was 0.285 (Table 4); however, given the high degree of relatedness within FHS pedigrees, these linkage disequilibrium measures obtained from unrelated participants may be an underestimation of the true correlation between SNPs within the analytic dataset, where, among related people, large chromosomal regions are expected to be identical by descent. We therefore also used LME models to examine simultaneously the effects of the three strongest association signals (K121Q, rs7773477, and rs7775386). Although the association of rs7773477 with FPG remained nominally significant after addition of rs7775386 to the model, statistical significance disappeared after inclusion of K121Q to models that contained either rs7773477 or rs7775386. This indicates that the associations of these SNPs with diabetes-related quantitative traits are likely accounted for by their linkage disequilibrium with K121Q and that the latter is giving the strongest common variant association signal in the region.

Supplemental Table 2 lists all the data for each polymorphism and quantitative traits examined. There was no significant association between any variant in ENPP1 and BMI or waist circumference. There was no significant relationship between the non-HapMap SNP rs1799774 (which codes for a T/del change) and any insulin resistance trait. Other than a nominal $P$ value of 0.02 for rs7775386, there was no significant association between K121Q or any other variant and risk for incident diabetes.

Finally, we also examined the interaction between K121Q and BMI because of preliminary evidence suggesting that the effect of the Q allele may be modified by an increase in adiposity. There was a nominally significant interaction between genotype at ENPP1 K121Q and BMI for the associations of the SNP with FPG (interaction $P$ value $= 0.008$, $\beta$ estimate $= 0.017$) and HOMA-IR (interaction $P$ value $= 0.003$, $\beta$ estimate $= 0.007$).
action $P$ value = 0.014, $\beta$ estimate = 0.016). This indicates a stronger genetic association of the Q allele with insulin resistance traits among people who have a higher BMI.

**DISCUSSION**

The association of the K121Q polymorphism in ENPP1 with insulin resistance and type 2 diabetes has been controversial. Regarding insulin resistance, Pizzuti et al. (2) found that nonobese, nondiabetic Q allele carriers were more insulin resistant than KK homozygotes, as defined both by OGTT and the euglycemic clamp. Subsequent studies reported both positive (20) and negative evidence (10) of association with insulin resistance. For type 2 diabetes, an initial positive result of association by Pizzuti et al. (2) was followed by replication in several independent studies (3–5) and confirmation in partial meta-analyses (10,21); however, three very large association studies (including one by members of our group) had gone extensive phenotypic characterization in a longitudinal fashion; and 4) its family component reduces the risk of population stratification.

In this study, we captured most of the common genetic variation in ENPP1 and studied its putative association with glycemic traits in a comprehensive manner. Using the FHS population to characterize the common variation around a quantitative glycemic or obesity trait; and 4) its family component reduces the risk of population stratification.

Given the new evidence suggesting a real association of this polymorphism with type 2 diabetes and functional reports implicating ENPP1 and its polymorphism K121Q in mechanisms of insulin resistance (1,8,27–29), we decided to examine its association with insulin resistance traits in a homogeneous population cohort previously unexamined for this variant. In addition, we aimed to determine whether any association, if present, stemmed from ENPP1 K121Q or another polymorphism in the region, and we 5) it is an ethnically homogeneous sample; 3) it has undergone extensive phenotypic characterization in a longitudinal fashion; and 4) its family component reduces the risk of population stratification.

In this study, we captured most of the common genetic variation in ENPP1 and studied its putative association with glycemic traits in a comprehensive manner. Using the FHS population to characterize the common variation across the ENPP1 locus, we confirmed the association of the Q allele in ENPP1 K121Q with hyperglycemia, as demonstrated by elevated FPG and A1C, under both the

FIG. 1. A: Negative log base 10 of the $P$ value for genetic associations with FPG (circles) and glycated hemoglobin (diamonds) under the additive model (left $y$-axis) graphed versus SNPs in the ENPP1 region arranged by chromosomal position (x-axis). The continuous line marked by the right $y$-axis indicates the recombination rate. The ENPP1 gene is shown by the horizontal arrow at the bottom of the plot. Open symbols indicate traits adjusted for sex and age; closed symbols indicate additional adjustment for BMI. B: Negative log base 10 of the $P$ value for genetic associations with fasting insulin (circles) and insulin resistance by HOMA (diamonds) under the additive model (left $y$-axis) graphed versus SNPs in the ENPP1 region arranged by chromosomal position (x-axis). The continuous line marked by the right $y$-axis indicates the recombination rate. The ENPP1 gene is shown by the horizontal arrow at the bottom of the plot. Open symbols indicate traits adjusted for sex and age; closed symbols indicate additional adjustment for BMI.
additive and dominant models. This supports the hypothesis that only one copy of the Q allele is necessary to cause an effect on quantitative phenotypes. We also demonstrated that the effect of K121Q on hyperglycemia is likely mediated via insulin resistance, because the Q allele is also associated with elevated fasting insulin and HOMA-IR. The lack of an association with the Gutt insulin sensitivity index may reflect differences in insulin resistance at the tissue level (basal hepatic insulin resistance vs. peripheral glucose disposal after an oral load) or indicate an imperfect correlation of these surrogate measures with true insulin resistance. Although other polymorphisms in the region showed similar associations with glycemic traits, our regression analysis demonstrated that the effect of two other significant SNPs was removed when incorporating K121Q in the models, suggesting that K121Q is the variant with the actual effect on glycemic and insulin resistance traits, as might be predicted by its impact on amino acid sequence. Because these associations represent confirmation of previous findings and other variants in the region were genotyped as a fine-mapping exercise, we do not believe statistical correction for the multiple variants analyzed is warranted.

Having established the association between ENPP1 K121Q and hyperglycemia, we explored the interaction between K121Q and BMI because of preliminary evidence that the effect of the Q allele on glycemic traits is mediated by an increase in adiposity (3,5,7,13,30) and suggestions that this variant may also contribute to obesity traits (5,6,31–33). Although we observed no association of ENPP1 K121Q with BMI or waist circumference, our interaction analysis supports the observation that a higher BMI strengthens the association of this particular polymorphism with elevated insulin resistance and glucose levels. This finding is consistent with the hypothesis that the net effect of the ENPP1 Q121 variant in modulating the risk of insulin resistance and related clinical outcomes is barely detectable in lean individuals while becoming more evident in the context of an “obesogenic” background, where the deleterious effect of the Q121 variant on the glucose disposal of skeletal muscle may be superimposed on that exerted by high BMI itself (30). This model is consistent with our previous meta-analysis in which we noted that the Q121 variant confers a modest risk of type 2 diabetes in whites with a greater effect as BMI increases. Such BMI × genotype interactions may be particularly evident with regard to genes that cause hyperglycemia by augmenting insulin resistance rather than in those that contribute to diabetes risk by diminishing insulin secretion. Because of the relationship between obesity and insulin resistance, there is more likely to be a correlation between increased adiposity and the effects of genes that modify insulin action.

In summary, our study adds further evidence in support of a potential causative role of the ENPP1 gene in the inheritance and pathophysiology of type 2 diabetes. We found that the Q allele of K121Q in ENPP1 appears to be the common variant most strongly associated with diabetes-related traits in whites, confirmed that K121Q is associated with hyperglycemia and a greater degree of insulin resistance, and found an adiposity SNP interaction, with a greater strength of association of K121Q with diabetes-related quantitative traits in people with obesity.

ACKNOWLEDGMENTS

E.S.S. is supported by National Institutes of Health (NIH) Training Grant T32 GM07748 Training Grant in Genetics. J.B.M. has received an American Diabetes Association Career Development Award, National Institute of Diabetes and Digestive and Kidney Diseases Grant K24-DK-080140, and research grants from GlaxoSmithKline and sanofi-aventis. J.C.F. has received NIH Research Career Award K23-DK-65978–04. This work has been supported by the National Heart, Lung, and Blood Institute’s FHS (Contract N01-HC-25195) and by the Boston University Linus Cluster for Genetic Analysis (LinGA) funded by NIH National Center for Research Resources Shared Instrumentation Grant 1S10-RR-163736-01A1.

REFERENCES

1. Maddux B, Goldfine I: Membrane glycoprotein PC-1 inhibition of insulin receptor function occurs via direct interaction with the receptor α-subunit. Diabetes 49:13–19, 2000