Title
High frequency of the PNPLA3 rs738409 [G] single-nucleotide polymorphism in Hmong individuals as a potential basis for a predisposition to chronic liver disease.

Permalink
https://escholarship.org/uc/item/79r231fn

Journal
Cancer, 124 Suppl 7

ISSN
0008-543X

Authors
Tepper, Clifford G
Dang, Julie HT
Stewart, Susan L
et al.

Publication Date
2018-04-01

DOI
10.1002/cncr.31122

Peer reviewed
High Frequency of the \( \text{PNPLA3} \) rs738409 [G] Single-Nucleotide Polymorphism in Hmong Individuals as a Potential Basis for a Predisposition to Chronic Liver Disease

Clifford G. Tepper, PhD; Julie H. T. Dang, MPH; Susan L. Stewart, PhD; Dao M. Fang, DHA; Kimberly A. Wong, MD; Stephenie Y. Liu, BA; Ryan R. Davis, BS; Doan Y. Dao, MD; Jeffrey P. Gregg, MD; Natalie J. Török, MD, MSc; and Moon S. Chen Jr., PhD, MPH

BACKGROUND: An exploratory study was performed to determine the prevalence of the patatin-like phospholipase domain-containing protein 3 (\( \text{PNPLA3} \)) rs738409 [G] allele among the Hmong as a risk factor for nonalcoholic fatty liver disease (NAFLD). NAFLD/nonalcoholic steatohepatitis is the world's most common chronic liver disease and is expected to replace viral hepatitis as the leading cause of cirrhosis and potential precursor to hepatocellular carcinoma (HCC). Of all populations in California, the Hmong experience the highest risk of death from HCC and the highest prevalence of metabolic syndrome risk factors among Asians that predispose them to NAFLD. Here a genetic explanation was sought for the high rates of chronic liver disease among the Hmong. The literature pointed to the \( \text{PNPLA3} \) rs738409 [G] allele as a potential genetic culprit. METHODS: Cell-free DNA was isolated from 26 serum samples previously collected in community settings. Quantitative polymerase chain reaction-based single-nucleotide polymorphism (SNP) genotyping was performed with a validated TaqMan SNP genotyping assay, and results were analyzed with TaqMan Genotyper software. RESULTS: The \( \text{PNPLA3} \) rs738409 [C>G] variant occurred at a frequency of 0.46 (12 of 26; 95% confidence interval, 0.27-0.67). This carrier rate would rank the Hmong as the third highest population in the 1000 Genomes Project. CONCLUSIONS: Although this small sample size limits the generalizability, the high frequency rates of this allele along with the presence of metabolic syndrome risk factors warrant further studies into the etiology of NAFLD among the Hmong. Cancer 2018;124:1583-9. © 2018 American Cancer Society.

KEYWORDS: carrier rate, hepatocellular carcinoma, Hmong, nonalcoholic steatohepatitis (NAFLD), patatin-like phospholipase domain-containing protein 3 (\( \text{PNPLA3} \)) rs738409.

INTRODUCTION

Chronic liver diseases, specifically nonalcoholic fatty liver disease (NAFLD) and its pathologically more advanced form, nonalcoholic steatohepatitis (NASH), along with hepatocellular carcinoma (HCC) are at epidemic proportions both worldwide and in the United States. The prevalence of NAFLD is estimated to be at 30% in the United States and up to 45% in Asia. Because of the rapid increases in fatty liver disease, NAFLD/NASH is expected to replace viral hepatitis as the leading cause of cirrhosis; NAFLD is already the most common chronic liver disease worldwide. HCC can occur as a sequel to NASH or on account of chronic infection with hepatitis B virus (HBV), hepatitis C virus, or a combination of the two. In the United States, HCC is responsible for the highest annual percentage increases in mortality rates—2.8% for males and 2.1% for...
females—in comparison with all other cancer sites, whose rates have decreased by 1.8% for males and 1.4% for females. Concurrent NASH increases the risk of HCC among patients with chronic HBV. On a global scale, HCC has become the world’s second deadliest cancer in numerical terms (after lung cancer).

According to an analysis of 33,270 cases of HCC diagnosed from 1988 to 2012 and reported to the California Cancer Registry, Laotian/Hmong experienced the highest risk of cause-specific mortality (hazard ratio, 1.50; 95% confidence interval, 1.29-1.73) among all 15 racial/ethnic groups. Multiple risk factors, including viral hepatitis, have been identified for HCC pathogenesis. There are different risk factors for NAFLD. Among Asians in Sacramento County, California, the Hmong experience the highest prevalence of metabolic risk factors for HCC, such as diabetes, a large waist circumference, and a high body mass index, as well as chronic HBV infections. However, the biological basis for the disparity is ill-defined and understudied. Hence, the purpose of our study was to ascertain the potential for ethnic-specific variations as mechanisms mediating chronic liver disease and possibly a genetic factor contributing to this disparity in the Hmong.

In the context of NAFLD, single nucleotide variants in the patatin-like phospholipase domain-containing protein 3 (PNPLA3) gene represent an important genetic mechanism. The most prominent variant is PNPLA3 rs738409 [G] (where G is guanine), which is a nonsynonymous substitution of cytosine (C) to guanine (G) [C>G] that changes codon 148 from encoding isoleucine (I) to methionine (M; I>M, 1148M). This allele was identified in a genome-wide analysis of nonsynonymous variations (ie, those likely to affect protein function) and was found to be strongly associated with hepatic fat/triglyceride content (P = 5.9 × 10^{-10}) and hepatic inflammation (P = 3.7 × 10^{-4}), and its frequency was concordant with the relative prevalence of NAFLD in European Americans, African Americans, and Hispanics. In a cohort of Chinese patients, the PNPLA3 rs738409 [G] allele was associated with susceptibility to NAFLD (odds ratio, 1.94; 95% confidence interval, 1.12-3.37; P = .018) and was found at a frequency of 0.34 (HapMap).

The PNPLA3 protein is a triacylglycerol lipase with hydrolytic activity toward triglycerides in hepatocytes and retinyl esters in hepatic stellate cells. The I148M amino acid change occurs in the patatin-like phospholipase domain and leads to a loss of function promoting triacylglycerol accumulation in hepatocytes as well as a gain of function, including elevated lysophosphatidic acid acyltransferase and thioesterase activity. Taken together, heterozygous (CG) or homozygous (GG) PNPLA3 rs738409 [G] genotypes can increase the susceptibility to the development of NAFLD, including fibrosis risk and progression. Thus, we hypothesized that the PNPLA3 rs738409 [G] variant could be a genetic mechanism that might partially explain the health disparity of increased rates of chronic liver disease among the Hmong.

MATERIALS AND METHODS

Research Participants
Twenty-six Hmong adults who participated in a community screening for viral hepatitis and cancer research in Sacramento County, California, each donated 5 mL of blood. Participants were recruited via partnering with community-based organizations and through in-language flyers and radio public service announcements. Participants completed an intake form with the assistance of lay bilingual community health workers because all participants preferred responding in Hmong rather than English. The intake included the following: 1) questions regarding the country of birth, sex, and age; 2) research staff measuring and documenting each participant’s waist circumference, height, and weight; and 3) a self-reported history of high blood pressure, high cholesterol levels, and smoking. No information on alcohol intake was collected, nor were any lipid profiles conducted. Our limited budget meant confining testing to only PNPLA3 even though we recognize that other mutations could have been studied. Participants also provided 2 additional samples for HBV and hemoglobin A1c testing. All laboratory tests were completed by the Department of Pathology and Laboratory Medicine at the University of California Davis.

Serum Samples
Whole blood samples were obtained from volunteers by venipuncture, were drawn into 6.0-mL red-top Vacutainer tubes (Becton-Dickinson), and were transferred to the Biorepository Shared Resource (University of California Davis Comprehensive Cancer Center). Serum was obtained when the blood was allowed to clot (15-30 minutes at room temperature), and this was followed by centrifugation at 1000 to 2000g for 10 minutes at 4°C to remove the clot. The serum was then aliquoted and stored at –80°C.

Serum Cell-Free DNA Isolation
Serum samples were submitted to the Genomics Shared Resource (University of California Davis Comprehensive Cancer Center) for DNA isolation and PNPLA3 single nucleotide polymorphism (SNP) analysis. Circulating cell-free DNA was isolated from serum samples (0.5 mL) with the QIAamp circulating nucleic acid kit (Qiagen).
was quantified with a NanoDrop spectrophotometer and a Qubit 2.0 fluorometer (double-stranded DNA high-sensitivity assay kit; Thermo Fisher Scientific), and was sized with an Agilent 2100 bioanalyzer.

**PNPLA3 SNP Genotyping**

We performed PNPLA3 SNP genotyping for the 26 Hmong participants. SNP genotyping analysis was performed for each DNA sample (5 ng of DNA, >0.2 ng/μL concentration, and 1342 genome equivalents). A pre-designed and validated TaqMan SNP genotyping assay (C______7241_10; catalog no. 4351379; Thermo Fisher Scientific) was used and performed according to the manufacturer's standard protocols for the reaction setup and thermal cycling with endpoint readings on a StepOnePlus real-time polymerase chain reaction system (Applied Biosystems). For the PNPLA3 rs738409 [G] SNP, the context sequence used for the assay design was AGGCCTTGGTATGTTCTGCTTCAT[C/G]CCCTTCTACAGTGGCCTTATCCCTC. Triplicate assays were performed for each sample as well as a no-template control. The data analysis was performed for genotype calling (ie, C/C, C/G, and G/G) with TaqMan Genotyper Software (Applied Biosystems), and allelic discrimination plots were generated to visualize the genotypes of the entire cohort.

**Statistical Analysis**

The distributions of participant demographics (country of birth, sex, and age) and health characteristics (waist circumference; body mass index; self-reported high blood pressure, high cholesterol levels, and smoking status; and diabetes and HBV status determined by laboratory tests) were summarized for those who had the PNPLA3 rs738409 [G] allele and those who did not and for the entire cohort. Notably, we used the Asian cut points for the waist circumference and the body mass index. An exact 95% confidence interval was computed for the proportion of participants with the PNPLA3 rs738409 [G] allele.

**Ethical Approval for Research Involving Human Participants**

All followed procedures were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975 (as revised in 2000). Informed consent was obtained from all patients for inclusion in the study. This study was performed in accordance with institutional review board protocol 128204, which was approved by the University of California Davis for the “University of California Davis Pathology Biorepository: Tissue, Blood, Urine, and Other Biological Materials”.

**RESULTS**

Twenty-six Hmong adults from community screenings for viral hepatitis donated 1 extra tube of blood (5 mL) for cancer research in Sacramento County, California. Demographic and HCC risk factors are displayed in Table 1. In general, participants were foreign-born (50%), were female (76.9%), had a waist circumference above the Asian cut point (75%), had a body mass index of 23 kg/m² (Asian cut point), or higher (90.5%), did not have high blood pressure (76.9%), did not have a high cholesterol level (80.8%), did not smoke (73.1%), and were diabetic or prediabetic (46.1%); 7.7% were chronically infected with HBV.
As noted in Figure 1, our results demonstrate that the \textit{PNPLA3} rs738409 [C>G] variant occurs at a high frequency of 0.46 among the Hmong (12 of 26 samples; 95% confidence interval, 0.27-0.67), with 10 samples being heterozygous (C/G) and 2 samples being homozygous (G/G) for the allele (Fig. 1A, B and Table 2). These results are quite compelling because a similarly high frequency was reported for Hispanics (0.49), who had the highest prevalence of NASH in the Dallas Heart Study and the only statistically significant association between the \textit{PNPLA3} rs738409 [G] allele and elevated serum alanine aminotransferase levels in comparison with European Americans (frequency = 0.23) and African Americans (frequency = 0.17).\textsuperscript{11}

If the Hmong prevalence of 0.46 were considered part of the 1000 Genomes Project, the Hmong would rank third highest (Fig. 2). All other Asian ancestral populations in the 1000 Genomes Project exhibited comparable but lower \textit{PNPLA3} rs738409 [G] allele frequencies (Fig. 2); they included Han Chinese in Beijing (frequency = 0.383), Southern Han Chinese (frequency = 0.390), Kinh in Ho Chi Minh City, Vietnam (frequency = 0.308), and Japanese in Tokyo, Japan (frequency = 0.423).\textsuperscript{19}

**DISCUSSION**

The purpose of this study was to evaluate whether the \textit{PNPLA3} rs738409 [C>G] variant could be a genetic factor contributing to the disproportionately high rates of chronic liver disease, specifically NAFLD, among California’s Hmong. With a sample size of 26, we recognize that this study should be considered exploratory and that the findings should be interpreted with caution because a larger cohort would be required to be more conclusive. For instance, the sample sizes for all populations in the 1000 Genomes Project were \textgeq{}100. Because of the small sample size, our 95% confidence interval for the prevalence of this gene variant is very wide (0.27-0.67), and this

**Figure 1.** \textit{PNPLA3} rs738409 [G] genotypes among the Hmong in California. TaqMan single nucleotide polymorphism genotyping assays were performed in triplicate on serum cell-free DNA samples obtained from 26 Hmong individuals. Genotype calls (ie, C/C, C/G, and G/G) were made with Taq-Man Genotyper software (Applied Biosystems). The genotype frequencies were summarized according to (A) the 3 possible genotypes (C/C, C/G, and G/G) and (B) samples with the G allele (C/G or G/G) or without it (C/C). (C) The genotypes of the entire cohort of samples (in triplicate) are visualized as an allelic discrimination plot. NTC indicates no-template control; \textit{PNPLA3}, patatin-like phospholipase domain-containing protein 3.
### TABLE 2. Results of the PNPLA3 rs738409 [C>G] SNP Assays for Individual Samples

<table>
<thead>
<tr>
<th>CCB Study Identifier</th>
<th>Specimen Type</th>
<th>Sex</th>
<th>Age, y</th>
<th>PNPLA3 rs738409 [C&gt;G] SNP Calls</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>96</td>
<td>Serum</td>
<td>Male</td>
<td>43</td>
<td>Heterozygous C/G</td>
<td>C/G</td>
</tr>
<tr>
<td>95</td>
<td>Serum</td>
<td>Female</td>
<td>60</td>
<td>Heterozygous C/G</td>
<td>C/G</td>
</tr>
<tr>
<td>88</td>
<td>Serum</td>
<td>Female</td>
<td>51</td>
<td>Heterozygous C/G</td>
<td>C/G</td>
</tr>
<tr>
<td>104</td>
<td>Serum</td>
<td>Female</td>
<td>56</td>
<td>Heterozygous C/G</td>
<td>C/G</td>
</tr>
<tr>
<td>90</td>
<td>Serum</td>
<td>Female</td>
<td>63</td>
<td>Heterozygous C/G</td>
<td>C/G</td>
</tr>
<tr>
<td>138</td>
<td>Serum</td>
<td>Female</td>
<td>25</td>
<td>Heterozygous C/G</td>
<td>C/G</td>
</tr>
<tr>
<td>110</td>
<td>Serum</td>
<td>Unknown</td>
<td>45</td>
<td>Heterozygous C/G</td>
<td>C/G</td>
</tr>
<tr>
<td>107</td>
<td>Serum</td>
<td>Unknown</td>
<td>47</td>
<td>Heterozygous C/G</td>
<td>C/G</td>
</tr>
<tr>
<td>97</td>
<td>Serum</td>
<td>Female</td>
<td>30</td>
<td>Heterozygous C/G</td>
<td>C/G</td>
</tr>
<tr>
<td>150</td>
<td>Serum</td>
<td>Female</td>
<td>24</td>
<td>Heterozygous C/G</td>
<td>C/G</td>
</tr>
<tr>
<td>93</td>
<td>Serum</td>
<td>Female</td>
<td>42</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>143</td>
<td>Serum</td>
<td>Female</td>
<td>26</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>140</td>
<td>Serum</td>
<td>Male</td>
<td>50</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>91</td>
<td>Serum</td>
<td>Female</td>
<td>70</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>109</td>
<td>Serum</td>
<td>Female</td>
<td>59</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>94</td>
<td>Serum</td>
<td>Male</td>
<td>45</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>89</td>
<td>Serum</td>
<td>Female</td>
<td>57</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>141</td>
<td>Serum</td>
<td>Female</td>
<td>41</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>139</td>
<td>Serum</td>
<td>Male</td>
<td>28</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>103</td>
<td>Serum</td>
<td>Female</td>
<td>48</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>92</td>
<td>Serum</td>
<td>Female</td>
<td>42</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>108</td>
<td>Serum</td>
<td>Female</td>
<td>50</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>111</td>
<td>Serum</td>
<td>Female</td>
<td>50</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>124</td>
<td>Serum</td>
<td>Male</td>
<td>44</td>
<td>Homozygous C/C</td>
<td>C/C</td>
</tr>
<tr>
<td>105</td>
<td>Serum</td>
<td>Unknown</td>
<td>Not reported</td>
<td>Heterozygous G/G</td>
<td>G/G</td>
</tr>
</tbody>
</table>

Abbreviations: CCB, cancer center biorepository; PNPLA3, patatin-like phospholipase domain-containing protein 3; SNP, single-nucleotide polymorphism.

**Figure 2.** PNPLA3 rs738409 [G] variant frequency among the Hmong versus other populations worldwide. The mean global frequency of the PNPLA3 rs738409 [G] variant and its frequencies in 26 populations worldwide were sourced from the 1000 Genomes Project (National Center for Biotechnology Information 1000 Genomes Browser). A comparison of the frequency defined for the Hmong/CA and those for the other populations is depicted in the bar graph. ACB indicates African Carribbeans in Barbados; ASW, African Ancestry in Southwest United States; BEB, Bengali from Bangladesh; CDX, Chinese Dai in Xishuangbanna, China; CEU, Utah residents (CEPH) with Northern and Western European ancestry; CHB, Han Chinese in Beijing, China; CHS, Southern Han Chinese; CLM, Colombians in Medellin, Colombia; ESN, Esan in Nigeria; FIN, Finnish in Finland; GBR, British in England and Scotland; GIH, Gujarati Indians from Houston, Texas; GWD, Gambians in western divisions in the Gambia; Hmong/CA, Hmong in California; IBS, Iberian population in Spain; ITU, Indian Telugu from the United Kingdom; JPT, Japanese in Tokyo, Japan; KHV, Kinh in Ho Chi Minh City, Vietnam; LWK, Luhya in Webuye, Kenya; MSL, Mende in Sierra Leone; MXL, Mexican ancestry in Los Angeles, California; PEL, Peruvians from Lima, Peru; PJL, Punjabi from Lahore, Pakistan; PNPLA3, patatin-like phospholipase domain-containing protein 3; PUR, Puerto Ricans from Puerto Rico; STU, Sri Lankan Tamil from the United Kingdom; TSI, Toscani in Italia; YRI, Yoruba in Ibadan, Nigeria.
indicates that the estimated prevalence of 0.46 is quite imprecise and must be interpreted with caution. Nevertheless, our findings are evidence that this gene variant occurs in the Hmong at an above-average frequency. The sample size was too small to provide adequate power for a statistical comparison of participants with and without the [G] allele, which would be necessary to interpret any differences. The sample size was also too small to assess a correlation between the prevalence of the mutation and sex; in particular, only 4 participants were identified as male. Other limitations included the fact that our sample of Hmong was derived from a community screening rather than a clinical setting. Those who participated may have been motivated by their interest in liver health and hence may not have been representative of the Hmong at large. On the other hand, a community-derived sample could also mean that the true prevalence is underestimated, whereas a clinically derived sample could mean an overestimation of the true prevalence.

Despite these known constraints, this study has several strengths. Because our original study included a collection of primary data intended to characterize HCC risk factors, we were able to report findings characterizing diabetes with Asian cutoff values for each participant. Demographic characteristics included the country of birth, sex, and age. HCC risk factors included the waist circumference, body mass index, blood pressure, cholesterol level, smoking status, diabetes status, and HBV status. Although the Hmong are a relatively small Asian group, their distinctiveness as the population experiencing the lowest liver cancer survival rates among all Californians8 means that adding a genetic dimension to our prior data on their considerable behavioral risk factors9 could help in estimating the genetic predisposition. Associating the genetic predisposition with known nongenetic and modifiable risk factors (eg, a healthier diet and more exercise) could be used to identify those with higher risks and could be the basis for precise population-based or clinical interventions to mitigate disease onset or progression.

In conclusion, the high frequency of PNPLA3 rs738409 [G] among the Hmong suggests that a genetic predisposition to NAFLD is plausible though not proven. Although this is consistent with our hypothesis that this variant is an etiologic factor in the disparity of increased NAFLD/NASH among the Hmong, a recent study suggests that it is not the sole determinant; instead, the pathogenesis is influenced through the interplay of genetic variants with environmental or physiological factors such as individual adiposity.24 Taken together, these findings provide support for pursuing research into genetic and lifestyle determinants for the chronic liver disease disparities affecting the Hmong.

FUNDING SUPPORT
This work was funded by the National Cancer Institute (U54CA153499) and by the Francis Yee Fund for Cancer Disparities Research. The Biorepository Shared Resource and the Genomics Shared Resource are supported by Cancer Center Support Grant P30CA093373. Julie H. T. Dang and Moon S. Chen Jr., report grants from the Asian American Network for Cancer Awareness and Training. Natalie Torok is funded by grant R01DK083283. However, the views expressed are those of the authors and do not necessarily reflect the views of the funders.

CONFLICT OF INTEREST DISCLOSURES
Doan Y. Dao reports consulting fees from the Francis Yee Fund for Cancer Disparities Research.
AUTHOR CONTRIBUTIONS

Clifford G. Tepper: Study design, interpretation of the data, drafting or critical review of the manuscript, final approval of the submitted manuscript, and accountability for all aspects of the work.

Julie H. T. Dang: Study design, interpretation of the data, drafting or critical review of the manuscript, final approval of the submitted manuscript, and accountability for all aspects of the work. Susan L. Stewart: Study design, interpretation of the data, drafting or critical review of the manuscript, final approval of the submitted manuscript, and accountability for all aspects of the work. Dao M. Fang: Study design, interpretation of the data, drafting or critical review of the manuscript, final approval of the submitted manuscript, and accountability for all aspects of the work. Kimberly A. Wong: Study design, interpretation of the data, drafting or critical review of the manuscript, final approval of the submitted manuscript, and accountability for all aspects of the work. Stephenie Y. Liu: Study design, interpretation of the data, drafting or critical review of the manuscript, final approval of the submitted manuscript, and accountability for all aspects of the work. Ryan R. Davis: Study design, interpretation of the data, drafting or critical review of the manuscript, final approval of the submitted manuscript, and accountability for all aspects of the work. Jeffrey P. Gregg: Study design, interpretation of the data, drafting or critical review of the manuscript, final approval of the submitted manuscript, and accountability for all aspects of the work. Natalie J. Török: Study design, interpretation of the data, drafting or critical review of the manuscript, final approval of the submitted manuscript, and accountability for all aspects of the work. Moon S. Chen Jr.: Study design, interpretation of the data, drafting or critical review of the manuscript, final approval of the submitted manuscript, and accountability for all aspects of the work.

REFERENCES