**CPT1A Missense Mutation Associated With Fatty Acid Metabolism and Reduced Height in Greenlanders**

Line Skotte, PhD; Anders Koch, MD, PhD; Victor Yakimov, MSc; Sirui Zhou, MSc; Bolette Søborg, MD, PhD; Mikael Andersson, MSc; Sascha W. Michelsen, MD, PhD; Johan E. Navne, MD, PhD; Jacqueline M. Mistry, MS; Patrick A. Dion, PhD; Michael L. Pedersen, MD, PhD; Malene L. Børresen, MD, PhD; Guy A. Rouleau, MD, PhD; Frank Geller, MSc; Mads Melbye, MD, DMSc*; Bjarke Feenstra, PhD*

**Background**—Inuit have lived for thousands of years in an extremely cold environment on a diet dominated by marine-derived fat. To investigate how this selective pressure has affected the genetic regulation of fatty acid metabolism, we assessed 233 serum metabolic phenotypes in a population-based sample of 1570 Greenlanders.

**Methods and Results**—Using array-based and targeted genotyping, we found that rs80356779, a p.Pro479Leu variant in CPT1A, was strongly associated with markers of n-3 fatty acid metabolism, including degree of unsaturation ($P=1.16\times10^{-15}$), levels of polyunsaturated fatty acids, n-3 fatty acids, and docosahexaenoic acid relative to total fatty acid levels ($P=2.35\times10^{-15}$, $P=4.02\times10^{-15}$, and $P=7.92\times10^{-13}$). The derived allele (L479) occurred at a frequency of 76.2% in our sample while being absent in most other populations, and we found strong signatures of positive selection at the locus. Furthermore, we found that each copy of L479 reduced height by an average of 2.1 cm ($P=1.04\times10^{-9}$). In exome sequencing data from a sister population, the Nunavik Inuit, we found no other likely causal candidate variant than rs80356779.

**Conclusion**—Our study shows that a common CPT1A missense mutation is strongly associated with a range of metabolic phenotypes and reduced height in Greenlanders. These findings are important from a public health perspective and highlight the usefulness of complex trait genetic studies in isolated populations. (Circ Cardiovasc Genet. 2017;10:e001618. DOI: 10.1161/CIRCGENETICS.116.001618.)

**Key Words:** fatty acid, unsaturated ■ Genome-Wide Association Study ■ Inuit ■ metabolomics ■ mutation, missense
Individuals were identified through the Greenlandic Civil Registration System and received a letter inviting them to participate. Written and informed consent was given by all participants and by parents for participants <18 years. The study was approved by the Commission for Scientific Research in Greenland (approval no. 2013-17) and by the Danish Data Protection Agency.

On the basis of power calculations done in R, we found that a sample size of \( n = 1500 \) is large enough to have adequate power (>80%) to detect common variants (minor allele frequency > 0.05%) with modest to high effect sizes (>0.6 SD) at genome-wide significance level (\( \alpha = 5 \times 10^{-8} \)) corresponding to Bonferroni correction for \( \approx 100 \) independent traits at conventional single-trait genome-wide significance (\( \alpha = 5 \times 10^{-8} \)). We randomly selected 1630 of the recruited individuals for genotyping.

### Array Genotyping and Sample Quality Control

DNA was extracted from buffy coats using a Chemagic STAR DNA Buffy Coat 200 kit (PerkinElmer). All samples were genotyped on the OmniExpressExome chip (HumanOmniExpressExome-8v1-2_A; Illumina), which is a genotyping array of 964,043 variants, including \( \approx 700,000 \) tag variants optimized for genome-wide association studies (GWAS) and \( > 250,000 \) functional exonic markers. Genotyping was performed using the HiScan System (Illumina), and genotypes were called using the GenCall module of the GenomStudio Software (Illumina) using default cluster data. Before analysis, we used the software PLINK\(^6\) to identify and remove samples with sex discordance, high rates of missing data, duplicates. This led to the removal of 60 samples, leaving 1570 genotyped samples for analysis.

### HapMap Genotypes

To investigate population structure, estimate admixture proportions, and calculate population branch statistics, we included a set of 209 individuals from the HapMap Consortium\(^10\) for joint analysis with the Greenlandic data. The HapMap samples had been genotyped by Illumina using the same genotyping array (HumanOmniExpressExome-8v1-2_A; Illumina) as the Greenlandic samples. The set of HapMap samples used in our analyses came from 44 unrelated Han Chinese individuals from Beijing, China (CHB), 45 unrelated Japanese individuals from Tokyo, Japan (JPT), 60 unrelated Utah residents with ancestry from northern and western Europe (CEU), and 60 unrelated Yoruba individuals from Ibadan, Nigeria (YRI).

### Serum Metabolomics and Quality Control

All samples were analyzed using a high-throughput serum NMR metabolomics platform.\(^8\) This methodology provides a standard output of 233 direct and derived serum measures,\(^12\) including quantitative molecular data on 14 lipoprotein subclasses, their lipid concentrations and composition, apolipoprotein A-I and B, multiple cholesterol and triglyceride measures, various fatty acids, as well as on numerous low-molecular-weight metabolites, including amino acids, glycolysis-related measures, and ketone bodies (a complete list is provided in Table I in the Data Supplement). Samples with low protein content, low glutamine/high glutamate, high lactate, high ethanol, or presence of polysaccharides were flagged during NMR analysis. For flagged samples, metabolite measures were set to missing as recommended by the laboratory, for example, glutamine and pyruvate values were set to missing for samples flagged as having low glutamine/high glutamate concentrations. In addition, we tested all other metabolite measure for correlation with any of the flags. For correlated metabolite measures (\( P < 10^{-5} \)), the value was set to missing in samples with the corresponding flag.

### Genome-Wide Association Analysis

To avoid increased type I error rates and decreased statistical power caused by population structure,\(^13\) we used the linear mixed effects model implemented in Genome-Wide Efficient Mixed Model Association\(^1\) to test for association. For each variant and each metabolite measure, we modeled the effects of the genotype and additional covariates as fixed effects and admixture and relatedness as random effects. The random effects were assumed to follow a multivariate normal distribution with variance proportional to a relatedness matrix estimated from standardized genotypes from all autosomal variants with an minor allele frequency >5% and ≤1% missing genotypes.

Before testing for association, we filtered out all variants with minor allele frequencies ≤1% or >5% missing genotypes and imputed remaining missing genotypes within the sample using Eagle.\(^11\) All traits were quantile-transformed, within sexes, to a standard normal distribution. Analysis were then performed using an additive genetic model, that is, with genotypes coded as 0, 1, and 2 denoting the number of minor alleles. The association tests were corrected for sex and age by inclusion of these variables as additional covariates in the mixed-effects model. We performed conditional analysis by further including the genotypes of the variants, which we conditioned on, as additional covariates in this linear mixed model.

To obtain a threshold for significant association, taking into account the number of metabolite measures tested, as well as their intercorrelations, the threshold value for single-trait genome-wide association was divided by the number of independent tests in the NMR data as estimated using the method of Li and Ji.\(^16\) We adopted the conventional \( P \) value of 5.0\( \times 10^{-8} \) as the threshold for single-trait genome-wide association. The effective number of phenotypes was estimated to be equal to 83, resulting in a \( P \) value threshold for genome-wide significance in the present study of 5.0\( \times 10^{-8}/83=6.02\times10^{-10} \).

### Genotyping of Candidate Causal Variant

To follow-up on the main association signal, we performed targeted genotyping of the candidate causal variant rs80356779 using competitive allele-specific polymerase chain reaction chemistry at LGC Genomics (Hoddesdon, United Kingdom).

### Anthropometric/Physiological Measurements and Clinical Biochemistry

During the consultation at the sampling location, we recorded the height and weight of the participants, as well as their systolic and diastolic blood pressure. The blood pressure measurements were taken 3 times per participant, and we used the median value in the analysis. We calculated the body mass index (BMI) of the participants based on their measured height and weight. We further performed clinically relevant routine laboratory measurements of 39 additional traits. The complete list of the anthropometric/physiological and clinical traits is shown in Table II in the Data Supplement. We performed

---

\(^{1}\)T. Skotte et al. Fatty Acid Associated CPT1A Missense Mutation

---

\(^{2}\)Figure 1. Sample locations and admixture. A, Sample locations in Greenland. B, Estimated admixture proportions of Inuit and European ancestry. Estimates for the 1570 Greenlandic samples are shown as narrow vertical bars to the left of the black line and estimates for the 60 European ancestry samples (CEU) are shown to the right. For each individual, the blue part of a vertical bar indicates the proportion of Inuit ancestry, whereas the red part indicates the proportion of European ancestry.
Admixture and Relatedness Estimation
We estimated admixture proportions for the genotyped samples using the software ADMIXTURE\(^1\) with \(K=2\), that is, we assume that the ancestry of all individuals can be explained by 2 distinct populations, the ancestral Inuit population and the European population. We included 60 unrelated CEU samples described above as a population reference. We ran the ADMIXTURE program 30 times with different random starting points to ensure that convergence was reached. The estimated admixture proportions were used as input for RelateAdmix,\(^1\) a program estimating relatedness between pairs of individuals while taking admixture into account.

Selection Analysis
On the basis of the estimated admixture proportions and relatedness described above, we identified all individuals with \(<2\%\) European ancestry and chose for selection analyses 157 of these such that none were closely related. That is, no pair had a relatedness coefficient \(r_{12}\) larger than 0.125, where \(k_1(k_2)\) is the proportion of the genome where 1(2) chromosome(s) is identical by descent between the 2 individuals. We combined the genotyping data for the 157 minimally related Greenlandic Inuit (GI) individuals with data from 60 unrelated CEU and 44 unrelated CHB samples described above. We then computed per-variant values of \(P_r\) for all pairwise combinations of populations (GI, CEU, and CHB) taking into account the sample sizes. On the basis of these estimates, we calculated variant-specific, the population branch statistics (originally proposed by Yi et al\(^2\)) for GI using CHB as sister population and CEU as an out-group.

Exome Sequencing Data
We combined exome sequencing data\(^2\) from 104 unadmixed Nunavik Inuit (NUI) individuals from 10 villages of Nunavik (Northern Quebec, Canada) with whole-genome sequencing data from 99 Europeans and 103 Chinese Han individuals from phase III of the 1000 Genomes Project (1KGP). We extracted all high confidence exome targets of 1KGP samples) in a 15-Mb region on chromosome 11 (58–74 Mb) that were polymorphic in the merged data. We combined the genotyping data for the 157 minimally related Greenlandic Inuit (GI) individuals with data from 60 unrelated CEU and 44 unrelated CHB samples described above. We then computed per-variant values of \(F_{st}\) for all pairwise combinations of populations (GI, CEU, and CHB) taking into account the sample sizes. On the basis of these estimates, we calculated variant-specific, the population branch statistics (originally proposed by Yi et al\(^2\)) for GI using CHB as sister population and CEU as an out-group.

GWAS Associations With Serum NMR Fatty Acid Measures
The genome-wide association scans identified an extended region on chromosome 11q12.2 to 13.3 strongly associated with several measures of fatty acid metabolism. The most significant association results were seen for fatty acid degree of unsaturation (Figure 2A and 2B).

Results
We recruited study participants from 7 regions in Greenland (Figure 1A) and successfully conducted array-based genotyping of 1570 individuals with the Illumina Human OmniExpressExome chip. Demographics of the genotyped participants can be found in Table II in the Data Supplement. To prioritize individuals of Inuit ancestry, we primarily invited individuals born in Greenland to parents also born in Greenland to participate. Even so, we generally observed a high degree of European admixture with a sample average of 27% European ancestry, although all participants had at least 20% Inuit ancestry and many participants had no European ancestry (Figure 1B). To identify genetic variants regulating metabolism in Greenlanders, we conducted genome-wide association scans for 233 serum NMR metabolic measures (summarized in Table I in the Data Supplement) using a linear mixed effects model to account for relatedness and admixture.\(^14\)

![Figure 2](image_url)
remained genome-wide significant when conditioning on the other 3 top variants (rs174570, rs3741395, and rs377432; Figure IA–IC in the Data Supplement; Table V in the Data Supplement), conditioning on rs1017640 left no variant in the region associated ($P<10^{-6}$; Figure ID in the Data Supplement; Table V in the Data Supplement).

Strong associations were also seen between rs1017640 and several additional measures related to fatty acid metabolism (Table IV in the Data Supplement; Figure II in the Data Supplement). Again, multiple additional association peaks were seen across a ≈10-Mb region, but could be accounted for by rs1017640 (Table V in the Data Supplement). Thus, for the examined phenotypes, we saw little evidence for multiple independent signals in this region.

Genomic inflation factors ranged from 0.993 to 1.087 and showed the highest inflation for the most strongly associated metabolic measures (Figure 2B; Figure II in the Data Supplement). Recalculating the genomic inflation factors while omitting the variants in the region harboring the association signal (chr11:58 Mb–74 Mb) showed no inflation (range, 0.987–1.012 for the most strongly associated metabolic measures; Figure 2B; Figure II in the Data Supplement). Thus, the extended LD in the region could account for the inflation observed initially.

Candidate Causal Variant rs80356779

The lead variant, rs1017640, is intronic in CPT1A, which encodes the liver isoform of carnitine palmitoyltransferase I, a key regulator of mitochondrial long-chained fatty acid oxidation. Previous studies have found that rs80356779, a p.Pro479Leu variant in CPT1A, occurs at derived allele frequencies of >70% in indigenous arctic populations, including Southwest Alaska Yup’ik, Inuit from Greenland, and Canadian Nunavut Inuit. An exome sequencing study of 100 NUI individuals found that the derived allele of rs80356779 (L479) occurred at a frequency of 95.5% and identified only 2 other coding variants in CPT1A, both of which were synonymous. In other populations, L479 is essentially absent; for example, it is seen only in 2 heterozygotes (1 Latino and 1 non-Finnish European) among the 60,706 individuals in the Exome Aggregation Consortium data.

Because rs80356779 is not present on the OmniExpressExome array, we performed targeted genotyping of this variant in our study sample. We found that L479 was strongly associated with a lower degree of unsaturation (P=1.16×10^{-34}; Figure 3A), a lower concentration of n-3 fatty acids (P=3.37×10^{-14}), a lower concentration of docosahexaenoic acid (DHA) (P=9.22×10^{-21}), lower ratios of polyunsaturated fatty acids, n-3 fatty acids, and DHA to total fatty acids (P=2.35×10^{-15}, P=4.02×10^{-19}, and P=7.92×10^{-27}, respectively), and a higher ratio of monounsaturated fatty acids to total fatty acids (P=9.02×10^{-16}; Table 1). Other fatty acid concentrations and ratios showed less significance or no association (Figure III in the Data Supplement; Table VI in the Data Supplement), and no other NMR measures were associated with genome-wide significance (Table VI in the Data Supplement).

Conditional analyses showed that rs80356779 was still significantly associated with degree of unsaturation when conditioning on the top variants from the 3 other peaks in the region (P<10^{-17}) and strongly associated (P=4.8×10^{-9}) when conditioning on the GWAS lead variant rs1017640 (Table V in the Data Supplement). Conversely, the GWAS lead variant
We found that L479 was strongly associated with decreased height ($P=1.0 \times 10^{-9}$, 8.7×10^{-8}, and 2.8×10^{-7}, respectively; Table 2). No other anthropometric/physiological and clinical traits were associated with trait-wide significance, $P<0.0011$ (Table VII in the Data Supplement). Also for these traits, the GWAS lead variant and the 3 other top variants were not associated when conditioning on rs80356779 (Table VIII in the Data Supplement).

To verify that the identified associations were not artifacts because of population stratification, we divided our study population into different strata of Inuit ancestry according to the estimated admixture proportions. Within each stratum, we then assessed the effect of the variant on the trait. We performed the stratified association analysis for degree of unsaturation and for height. Besides demonstrating that increasing proportion of Inuit ancestry is associated with lower height, the association analysis stratified by Inuit admixture proportion demonstrated that the effect of rs80356779 on height is observed across all strata (Figure 4A). Similarly, the effect of the variant was present across the different strata of samples for degree of unsaturation (Figure 4B).

**Signatures of Positive Selection**

Whole-genome sequencing of 25 individuals from indigenous Northeast Siberian coastal populations recently identified rs80356779 as the likely driver of a strong signal of positive selection. To investigate whether the observed association signal was in concordance with the causative variant being under strong selection, we calculated population branch statistics for the GWAS data. We identified a subset of 157 minimally related Inuit individuals with no European ancestry, whose genotypes we analyzed in combination with genotypes from 60 unrelated individuals of European ancestry and 44 unrelated Chinese Han. We found highly elevated levels of population branch statistics for CPTIA variants, supporting the notion of a functional variant in the region under strong positive selection.

**Exome Sequencing Data**

It is conceivable that an unknown neighboring variant with a similar frequency pattern as rs80356779 in the Inuit and European populations might be able to cause the observed association signal, mediated by linkage disequilibrium. To address this possibility, we consulted exome sequencing data from 104 NUI (northern Quebec, Canada) described in the Methods section. The NUI population is assumed to be a sister population to the GI, although some genetic differentiation is expected because of the small population sizes and geographical challenges to migration in the region. As reported previously, although some genetic differentiation is expected because of the small population sizes and geographical challenges to migration in the region. As reported previously, the frequency of L479 was 95.7% in the NUI and L479 was not found in the 1KGP CEU and CHB individuals. Among all 1403 coding variants in the 58 to 74 Mb region of

and the top variants from the 3 other peaks were not associated with degree of unsaturation when conditioning on rs80356779 ($P>$0.02; Figure 3A and 3B; Table V in the Data Supplement). Indeed, no variant in the 13-Mb surrounding region was associated ($P<10^{-4}$) when conditioning on rs80356779 (Figure 3B). We found similar results for conditional analysis for degree of unsaturation and for height. Besides demonstrating that increasing proportion of Inuit ancestry is associated with lower height, the association analysis stratified by Inuit admixture proportion demonstrated that the effect of rs80356779 on height is observed across all strata (Figure 4A). Similarly, the effect of the variant was present across the different strata of samples for degree of unsaturation (Figure 4B).

**Additional Associated Traits**

Next, we tested for association between rs80356779 and 44 additional anthropometric/physiological and clinical traits that were available for the study participants (summarized in Table II in the Data Supplement). We found that L479 was strongly associated with decreased height ($P=1.0 \times 10^{-9}$), and each copy of L479 corresponded to a 2.1-cm decrease in height (Table 2). We further found that L479 was associated with decreased levels of thyroid-stimulating hormone and with higher levels of alkaline phosphatase and transferrin ($P=1.0 \times 10^{-9}$, 8.7×10^{-8}, and 2.8×10^{-7}, respectively; Table 2). No other anthropometric/physiological and clinical traits were associated with trait-wide significance, $P<0.0011$ (Table VII in the Data Supplement). Also for these traits, the GWAS lead variant and the 3 other top variants were not associated when conditioning on rs80356779 (Table VIII in the Data Supplement).

**Table 1. Significant Associations of rs80356779 With Serum Nuclear Magnetic Resonance Spectroscopy Metabolic Measures**

<table>
<thead>
<tr>
<th>NMR Measure</th>
<th>Beta*</th>
<th>SE*</th>
<th>CI, Lower*</th>
<th>CI, Upper*</th>
<th>PValue</th>
<th>Beta†</th>
</tr>
</thead>
<tbody>
<tr>
<td>UnSat</td>
<td>−0.599</td>
<td>0.046</td>
<td>−0.690</td>
<td>−0.508</td>
<td>1.16E-34</td>
<td>−0.06</td>
</tr>
<tr>
<td>DHA/FA</td>
<td>−0.513</td>
<td>0.047</td>
<td>−0.605</td>
<td>−0.422</td>
<td>7.92E-27</td>
<td>−0.28</td>
</tr>
<tr>
<td>DHA</td>
<td>−0.427</td>
<td>0.045</td>
<td>−0.514</td>
<td>−0.339</td>
<td>9.22E-21</td>
<td>−0.03</td>
</tr>
<tr>
<td>FAw3/FA</td>
<td>−0.426</td>
<td>0.047</td>
<td>−0.518</td>
<td>−0.334</td>
<td>4.02E-19</td>
<td>−0.62</td>
</tr>
<tr>
<td>MUFa/FA</td>
<td>0.393</td>
<td>0.046</td>
<td>0.303</td>
<td>0.483</td>
<td>9.02E-16</td>
<td>1.51</td>
</tr>
<tr>
<td>PUFa/FA</td>
<td>−0.363</td>
<td>0.044</td>
<td>−0.449</td>
<td>−0.277</td>
<td>2.35E-15</td>
<td>−1.41</td>
</tr>
<tr>
<td>FAw3</td>
<td>−0.344</td>
<td>0.045</td>
<td>−0.432</td>
<td>−0.256</td>
<td>3.37E-14</td>
<td>−0.07</td>
</tr>
</tbody>
</table>

DHA indicates docosahexanoic acid; FA, fatty acids; FAw3, n-3 fatty acids; MUFa, mono-unsaturated fatty acids; NMR, nuclear magnetic resonance spectroscopy; MUFa, mono-unsaturated fatty acids; and UnSat, degree of fatty acid unsaturation.

*L479 effect size (in SD units), SE, and 95% confidence interval (CI) based on quantile transformed NMR measure values.

†L479 effect size based on untransformed NMR measure values.

**Table 2. Significant Associations of rs80356779 With Anthropometric/Physiological Traits and Clinical Measurements**

<table>
<thead>
<tr>
<th>Look-Up Trait</th>
<th>Beta*</th>
<th>SE*</th>
<th>CI, Lower*</th>
<th>CI, Upper*</th>
<th>PValue</th>
<th>Beta†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height</td>
<td>−0.315</td>
<td>0.049</td>
<td>−0.412</td>
<td>−0.219</td>
<td>1.04E-09</td>
<td>−2.08</td>
</tr>
<tr>
<td>Alkaline phosphatase</td>
<td>0.277</td>
<td>0.051</td>
<td>0.177</td>
<td>0.378</td>
<td>8.67E-08</td>
<td>5.39</td>
</tr>
<tr>
<td>TSH</td>
<td>−0.322</td>
<td>0.051</td>
<td>−0.421</td>
<td>−0.223</td>
<td>1.05E-09</td>
<td>−0.15</td>
</tr>
<tr>
<td>Transferin</td>
<td>0.273</td>
<td>0.053</td>
<td>0.170</td>
<td>0.376</td>
<td>2.76E-07</td>
<td>1.55</td>
</tr>
</tbody>
</table>

TSH indicates thyroid-stimulating hormone.

*L479 effect size (in SD units), SE, and 95% confidence interval (CI) based on quantile transformed trait values.

†L479 effect size based on untransformed trait values.
chromosome 11, we found none that were as frequent in the NUI and at the same time as rare in the CEU as rs80356779. In Table X in the Data Supplement, we list all coding variants with European minor allele frequency <35% that occur at 65% or higher frequency in the NUI.

**Discussion**

Circulating metabolites have key roles in many disease-related biological pathways, and understanding the genetic basis of their regulation may translate into improved prevention and clinical care. This study found that rs80356779, a p.Pro479Leu variant in CPT1A, was highly significantly associated with a range of fatty acid metabolism measures in a population-based sample from Greenland. We found strong signatures of positive selection at the locus, most extremely for rs80356779, where the derived allele was fixed in individuals of pure Inuit ancestry within our sample, and absent in the CEU and CHB HapMap populations. Furthermore, we found that each copy of L479 was associated with a 2.1-cm reduction in average height.

In exome sequencing data from a sister population, the NUI, we found no other likely causal candidate variant than rs80356779. Conditional analyses indicated that the associations observed for other variants in the region (including rs174570 in the FADS2 gene) could be accounted for by rs80356779 for all reported significant NMR measures.

![Figure 4](http://circgenetics.ahajournals.org/)

**Figure 4.** Association of rs80356779 stratified by admixture proportions. The estimated means are corrected for sex and age, and the error bars are standard error of the mean obtained using a standard linear model. **A**, Height. **B**, Degree of fatty acid unsaturation.

![Figure 5](http://circgenetics.ahajournals.org/)

**Figure 5.** Population branch statistics (PBS). **A**, The distribution of the PBS for all genome-wide association study (GWAS) variants. A red line marks the PBS value for the GWAS lead variant rs1017640. **B**, PBS for GWAS variants in a 0.5-Mb region surrounding the GWAS lead variant rs1017640. Each circle represents a variant, colored according to LD with the GWAS lead variant in the whole sample. The **bottom panel** shows the position of protein-coding genes in the region.
and anthropometric/physiological and clinical traits. These findings support rs80356779 as the likely causative variant, although it remains possible that the association could be caused by a noncoding variant in high LD not captured by the exome sequencing.

The effect of rs80356779 on various traits in Arctic populations has been addressed before,²²,²³ but with a study design that did not allow correction for population structure. Studies in Inuit from Greenland and Yu’pik Eskimos have reported that L479 was associated with elevated fasting high-density lipoprotein-cholesterol and ApoA1 levels²²,²³ and with lowered BMI and other obesity measures.²¹ We were not able to replicate these results. BMI showed no association, and although we found nominally significant associations between rs80356779 and high-density lipoprotein-cholesterol and ApoA1 (Table VI in the Data Supplement; Table VII in the Data Supplement), L479 corresponded to lower rather than higher levels of these metabolites. However, high-density lipoprotein-cholesterol and ApoA1 showed strong positive correlation with overall Inuit ancestry in our sample (P=7.92×10⁻¹⁷ and P=1.01×10⁻¹³, respectively), whereas BMI showed negative correlation (P=2.39×10⁻⁷). Because the L479 is a proxy for Inuit ancestry, the estimates of the variant’s effects can be highly confounded if the statistical model does not correct for population structure. Thus, uncorrected analyses of our data showed that L479 corresponded to higher levels of high-density lipoprotein and ApoA1 and lower levels of BMI (Table XI in the Data Supplement). Such examples of confounding underline the importance of proper modeling in studies of admixed populations.

Interestingly, in the final stages of the preparation of this article, a GWAS on erythrocyte membrane fatty acid composition among Greenlanders was published.²⁷ The study found that the CPT1A variant rs80356779 was associated with the levels of a range of fatty acids in the phospholipid fraction of erythrocyte membranes. These findings complement our serum metabolomics results in understanding the role of the p.Pro479Leu variant in CPT1A in regulating fatty acid levels in different blood components. Similar to our results, the authors report that L479 was associated with measures of smaller body size.

Our study further highlights the usefulness of complex trait genetic studies in historically isolated, but recently admixed populations. The demographic history of the Greenlandic population makes it possible to investigate the effects of a variant like rs80356779, which reached fixation in the ancestral Inuit population. Not only are the observed effects not an artifact of population structure it would also not have been possible to detect them in a purely Inuit population because all individuals would be homozygous for L479.

CPT1A is located in the mitochondrial outer membrane and is the rate-limiting enzyme responsible for importing long-chain fatty acids into the mitochondria of liver cells.²⁸ During fasting, CPT1A activity increases, allowing fatty acids to enter the mitochondria for energy production by β-oxidation. As fasting continues, the acetyl-CoA resulting from β-oxidation is increasingly used to synthesize ketone bodies (ketogenesis) that are circulated to provide energy for the brain and other tissues.²⁹ In the fed state, CPT1A activity is inhibited by the presence of malonyl-CoA, an intermediate product in fatty acid synthesis. Allosteric binding of malonyl-CoA to the regulatory domain of CPT1A prevents fatty acid entry to mitochondria and thereby effectively inhibits liver β-oxidation and ketogenesis.²⁰ Given the central importance of CPT1A for fatty acid metabolism, deficiencies in CPT1A can have severe consequences.²⁰ It is therefore intriguing that a p.Pro479Leu missense mutation in CPT1A exists in high frequencies in our study sample and in other indigenous Arctic populations. This suggests that the effects of the mutation have historically provided a selective advantage for individuals in these populations.

The variant affects the activity of CPT1A in 2 different ways, as illustrated in Figure IV in the Data Supplement. First, in vitro studies of cultured fibroblasts from homozygous carriers of L479 found that the activity of the enzyme was greatly reduced compared with control cells.²⁴,³¹ Thus, under low malonyl-CoA concentrations (in the fasting state), the enzymatic activity of CPT1A is reduced, although overall β-oxidation seems to be only moderately decreased.²⁴ Second, however, the mutation also markedly decreases the inhibitory effect of malonyl-CoA on liver β-oxidation,³¹ apparently by affecting a predicted binding site for malonyl-CoA, formed by Pro-479 together with neighboring amino acids.³²,³³ Thus, under high malonyl-CoA concentrations (in the fed state), the residual CPT1A activity has been shown to be 3 to 4 times higher in homozygous L479 fibroblasts compared with wild-type cells.³¹

This ability to rely on fatty acids and ketone bodies for energy even in the nonfasting state is predicted to be highly advantageous in individuals living on a traditional Inuit diet.³⁴ In the keto-adapted state, glucose is spared for cells with no or few mitochondria (eg, erythrocytes), whereas the brain is fueled by ketone bodies and most other tissues by fatty acids.³ It is thought that carriers of L479 have had a distinct survival advantage by being able to rapidly attain the keto-adapted state and remain in that state also in periods of lower dietary fat and higher protein intake, where ketosis would be switched off in noncarriers.³⁴ It has also been hypothesized that the lower activity of the mutant enzyme might convey a selective advantage by preventing overproduction of ketone bodies.³⁴ Whatever the specific fitness benefits might be, the selection analysis results presented here add to earlier evidence from Northeast Siberian coastal populations⁸ that L479 has been the likely target of strong positive selection in indigenous peoples of the Arctic region.

At a time where diet and lifestyles become increasingly westernized in arctic communities, genetic variants that were advantageous historically may now be neutral or even associated with adverse health outcomes. Understanding their effects therefore represents an important public health question. Some reports have found that L479 is associated with increased childhood mortality, but whether this represents a causal effect remains to be investigated.³⁴

Strengths of our study include the comprehensive analysis of a wide range of metabolic, anthropometric/physiological, and clinical traits in a GWAS setting, allowing for careful control of potential population structure biases. Finding an independent arctic replication population with genome-wide genotypes, as well as metabolite measurements, available
has not been possible and is a limitation of the study. Also, although our results point to rs80356779 as the likely causal variant, further studies involving, for example, cell line experiments would be needed to establish causality at the locus and to illuminate the molecular mechanisms underlying the observed associations.

In conclusion, we found that the p.Pro479Leu mutation was associated with lower degree of unsaturation, lower levels of DHA and n-3 fatty acids in general, lower ratios of polyunsaturated fatty acids, n-3 fatty acids, and DHA to total fatty acids, and a higher ratio of monounsaturated fatty acids to total fatty acids. It was also strongly associated with lower attained height, possibly relating to the effects of fatty acid metabolism on growth hormone secretion. To illuminate the biological mechanisms behind the associations reported here, further functional studies are required. In particular, studies including complete characterization of fatty acid and lipid profiles and direct measurements of CPT1A enzymatic activity together with genotypic data might be revealing. Our findings illustrate how carefully designed studies of populations adapted to extreme dietary or environmental conditions can provide important knowledge about basic human physiology.

Acknowledgments

We thank all volunteers for their generous participation in this study. We are grateful to Henrik Hjalgrim, Karen Bjorn-Mortensen, Emilie Birch, Anne Christine Nordholm, Marie Mila Broby Johansen, Helle Møller, and Lisbeth Carstensen for assistance with the collection and handling of data and samples. We are also thankful to Dr Allan Meldgaard Lund for helpful discussion. Web Resources: R software, version 3.2.0, http://www.r-project.org/. Exome Aggregation Consortium Browser, http://exac.broadinstitute.org/; accessed April 21, 2016.

Sources of Funding

The study was supported by grants from the Danish Medical Research Council, The Greenlandic Ministry of Education, Church, Culture and Gender Equality, the Maersk Foundation (Fonden til Lægevidenskabens Fremme), and the Aase and Ejnar Danielsens Foundation. The Danish National Biobank is supported by the Novo Nordisk Foundation. Dr Feenstra is supported by an Oak Foundation. The Danish National Biobank is supported by the Novo Nordisk Foundation. We are grateful to Henrik Hjalgrim, Karen Bjørn-Mortensen, Emilie Birch, Anne Christine Nordholm, Marie Mila Broby Johansen, Helle Møller, and Lisbeth Carstensen for assistance with the collection and handling of data and samples. We are also thankful to Dr Allan Meldgaard Lund for helpful discussion. Web Resources: R software, version 3.2.0, http://www.r-project.org/. Exome Aggregation Consortium Browser, http://exac.broadinstitute.org/; accessed April 21, 2016.

Disclosures

None.

References

29. Wang SP, Yang H, Wu JW, Gauthier N, Fukao T, Mitchell GA. Metabolism as a tool for understanding human brain evolution: lipid energy...


**CLINICAL PERSPECTIVE**

Variation in fatty acid levels is strongly associated with cardiovascular and metabolic health outcomes. An improved understanding of the combined effects of diet and genetics on fatty acid levels is therefore of considerable medical and physiological interest. Indigenous people of the Arctic represent an intriguing study population, having lived for thousands of years in an extremely cold environment relying on a diet highly enriched for marine-derived fat and depleted of carbohydrates. Here, we address the question of how this selective pressure is reflected in the genetic regulation of metabolism in present-day Greenlanders, by combining detailed metabolomics and genomics data in a population-based sample. Using the genome-wide association study approach, we identified a CPT1A missense mutation (rs80356779, p.Pro479Leu) strongly associated with a range of measures related to fatty acid metabolism, including lower degree of fatty acid unsaturation and lower concentrations of docosahexaenoic acid and n-3 fatty acids. CPT1A encodes the liver isoform of carnitine palmitoyltransferase I, a key regulator of mitochondrial long-chained fatty acid oxidation. The p.Pro479Leu variant is common in Arctic populations but virtually absent elsewhere, and it constitutes one of the most pronounced examples of human genetic adaptation. As diet and lifestyle become increasingly westernized in Arctic communities, genetic variants that were historically advantageous may now be associated with adverse health outcomes. Therefore, understanding the effects of p.Pro479Leu represents an important public health question. Furthermore, deeper insights into the molecular-level effects of the variant form of the CPT1A enzyme may provide leads for novel therapeutic approaches for cardiovascular disease.
CPTIA Missense Mutation Associated With Fatty Acid Metabolism and Reduced Height in Greenlanders

Line Skotte, Anders Koch, Victor Yakimov, Sirui Zhou, Bolette Søborg, Mikael Andersson, Sascha W. Michelsen, Johan E. Navne, Jacqueline M. Mistry, Patrick A. Dion, Michael L. Pedersen, Malene L. Børresen, Guy A. Rouleau, Frank Geller, Mads Melbye and Bjarke Feenstra

Circ Cardiovasc Genet. 2017;10:
doi: 10.1161/CIRCGENETICS.116.001618

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circgenetics.ahajournals.org/content/10/3/e001618

Data Supplement (unedited) at:
http://circgenetics.ahajournals.org/content/suppl/2017/06/13/CIRCGENETICS.116.001618.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Genetics can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Genetics is online at:
http://circgenetics.ahajournals.org/subscriptions/
SUPPLEMENTAL MATERIAL

SUPPLEMENTAL TABLES

Table I: Overview of NMR serum metabolite measurements used for the GWAS, (see separate Excel file).

Table II: Overview of anthropometric/physiological and clinical measurements used for look-up, (see separate Excel file).

Table III: Demography of study participants, (see separate Excel file).

Table IV: Association results for the GWAS lead variant rs1017640, (see separate Excel file).

Table V: P-values for conditional analyses of genome-wide significant NMR measures, (see separate Excel file).

Table VI: Association results for genotyped candidate variant rs80356779, (see separate Excel file).

Table VII: Association results for rs80356779 for anthropometric/physiological and clinical measurements, (see separate Excel file).

Table VIII: P-values for conditional analyses of trait-wide significant anthropometric/physiological and clinical measurements, (see separate Excel file).

Table IX: Top 30 population branch statistics (PBS) of all GWAS variants on the OmniExpressExome chip, (see separate Excel file).

Table X: All SNVs and 1bp INDELs in chr11:58Mb-74Mb with CEU minor allele frequency less than 35% and a frequency above 65% in Nunavik Inuit, (see separate Excel file).

Table XI: Illustrative association tests for ApoA1, HDL and BMI for rs80356779, (see separate Excel file).
Figure I: Conditional association results for degree of unsaturation with all tested Onmi Express Exome variants in a 15-Mb region containing the GWAS lead variant rs1017640. A) Association test p-values conditional on rs3741395. B) Association test p-values conditional on rs174570 (FADS2). C) Association test p-values conditional on rs377432. D) Association test p-values conditional on lead variant rs1017640.

Each variant is represented by a circle; the colour indicates the $r^2$ value between this variant and the variant with the lowest p-value in the region (rs1017640 for panel A-C and rs7936185 for panel D). The position along the left y-axis shows the negative logarithm of the p-value when testing for association using an additive genetic model with the conditioning variant included as covariate in the mixed effects regression model. The solid blue line illustrates the recombination rates from the Chinese HapMap (CHB) panel (right y-axis).
Figure II: Genome-wide association scan overviews for the OmniExpressExome variants and the significantly associated NMR measures (plot for degree of unsaturation is in the main article). Results are based on quantile transformed phenotypes and corrected for sex and age. For each associated trait, a Manhattan plot showing the $-\log_{10}(p)$-values for all variants on the chip and a quantile-quantile (QQ-) plot of the observed $-\log_{10}(p)$-values (y axis) against the corresponding expected values under the null hypothesis of no association (x axis) are included. The red line on the Manhattan plots indicates the genome-wide significance threshold of $6.02 \times 10^{-10}$. On the QQ-plots, the light blue dots are all variants on the chip, while variants in chr11:58MB-74MB are excluded when plotting the dark blue dots. The $\lambda$-values are the genomic inflation factors with/without chr11:58MB-74MB, respectively.
Figure III: Association results for rs80356779 with fatty acid related NMR traits. A: Each dot represents an NMR trait, the position along x-axis is the effect size estimated based on the quantile transformed trait values and the color corresponds to the \( p \)-value for the association test. B: Spearman correlation matrix for the fatty acid related NMR traits, trait name is colored according to \( p \)-value for association with rs80356779.
**Figure IV.** Illustration of changes in CPT1A activity for fasting and fed states. *No data available for P479/L479 heterozygous cells/carriers.

<table>
<thead>
<tr>
<th>Fasting state</th>
<th>Fed state</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low plasma glucose and insulin concentrations</td>
<td>High plasma glucose and insulin concentrations</td>
</tr>
<tr>
<td>Low lipogenesis</td>
<td>Glucose oxidation, glycogen storage and lipogenesis</td>
</tr>
<tr>
<td>Low Malonyl-CoA levels</td>
<td>High Malonyl-CoA levels</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CPT1A active</th>
<th>Malonyl-CoA inhibition of CPT1A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta-oxidation of long-chain fatty acids</td>
<td>Little CPT1A activity</td>
</tr>
<tr>
<td>Production of gluconeogenesis cofactors and ketone bodies</td>
<td>Long-chain fatty acids esterified to TAGs and secreted in VLDLs</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CPT1A less active</th>
<th>Incomplete malonyl-CoA inhibition</th>
</tr>
</thead>
<tbody>
<tr>
<td>On traditional diet, possibly compensated by increased expression due to omega-3</td>
<td>Moderate CPT1A activity</td>
</tr>
<tr>
<td>Production of gluconeogenesis cofactors and ketone bodies</td>
<td>Continued production of ketone bodies</td>
</tr>
</tbody>
</table>

**SUPPLEMENTAL REFERENCES**