The Allele Frequency of the HFE gene mutation H63D (rs1799945) and Its Relationship to a Hereditary Hemochromatosis Diagnosis in Metabolic Nutrition Students at Virginia Tech

Tyler R. Ferqueron, Angela S. Anderson, & Deborah J. Good*
Department of Human Nutrition, Foods, and Exercise at Virginia Polytechnic Institute and State University, Blacksburg, VA 24061

https://doi.org/10.33697/ajur.2020.006

Student: tylerf4@vt.edu
Mentors: goodd@vt.edu*; asphay@vt.edu

ABSTRACT
Hereditary hemochromatosis (HH) is a disease that causes excess iron absorption from the diet. This excess iron can be stored in the liver, skin, heart, pancreas, and joints, and then can lead to other health conditions, as the human body has no way of actively excreting iron. The human hemochromatosis protein (HFE protein) is encoded by the HFE gene, and mutations in this gene can lead to a dysfunction of the protein resulting in HH or iron overload later in adulthood. The objective of this study was to analyze the mutant allele frequency and the penetrance of the H63D mutation (SNP rs1799945) of the HFE gene in a cohort of Virginia Tech students. This study had a total of 69 participants. Fifty-two participants provided saliva samples, genomic data from 23andMe®, and surveys with phenotypic information. Of these, 6 were genotyped using the RFLP technique and served as controls for genotype confirmation. An additional 17 participants provided saliva samples, but did not provide 23andMe® data; genomic DNA from these participants were genotyped using the RFLP technique. Our results showed that although none of the participants had been diagnosed with HH, the mutant allele frequency of this population was 13.04%. In conclusion, as HH is usually diagnosed in older adults, we could not identify any students with a phenotype of HH, even though we could detect the mutant allele. This data suggests that affordable and accessible genetic ancestry and health kits such as the 23andMe® kit, could provide an efficient way to identify, prevent, and manage HH and other genetic diseases before symptoms arise.

KEYWORDS
Hereditary Hemochromatosis; Iron Absorption; Single Nucleotide Polymorphism; Restriction Fragment Length Polymorphism; College Student Population; 23andMe®; Population Analysis; Survey Results

INTRODUCTION
Hemochromatosis (HH) is a condition that causes excess absorption of the mineral iron. Iron works in the body as an electron donor and acceptor. Therefore, it is not surprising that the symptoms of HH, due to iron toxicity, are metabolic dysfunctions and imbalances from the production of reactive oxygen species (ROS).1,2 Excess iron produces ROS through the Fenton reaction, forming hydroxyl radicals from superoxide or hydrogen peroxide. Lipid peroxidative products which can form more toxic radicals are also formed in iron overload. Radicals formed from both the fenton reaction and from lipid peroxidative products can react with carbohydrates, proteins, and most importantly nucleic acids.1 Not only does excess iron lead to reactive oxygen species, it can exacerbate other conditions that cause the dysregulation of iron distribution among the tissues and organs of the body. The liver is one of the first organs to be affected by excess iron absorption. The liver stores the largest amount of iron in the body, and there are no pathways for the excretion of iron.3 Too much iron in the liver can potentially lead to cirrhosis, fibrosis, inflammation, cancer, and more minor to severe liver conditions.1,2,4,5 Excess iron can also lead to cardiac disease including congestive heart failure, arrhythmias, hypertrophy, and endocarditis.1

Currently, HH is associated with 5 different genes including HFE (hemochromatosis), TfR2 (transferrin receptor 2), HJV (hemojuvelin), HAMP (hepcidin), and SLC40A1 (solute carrier family 40 member 1).3 This study focused on the HFE gene as mutations in this gene account for the most common form of HH, affecting nearly 1 million people.6 The most severe HFE gene variant that is associated with HH is a mutation at position 282 in which the amino acid cysteine is replaced by tyrosine, written as Cys282Tyr or C282Y (rs1800562).3 However, the most common HFE gene variant is due to the single nucleotide polymorphism (SNP), rs1799945, that results in a mutation at amino acid position 63 of the HFE protein, where the amino acid histidine is replaced with an aspartate, written as Hisp63Asp or H63D. In the gene, the wild-type allele is a cytosine (C) and the mutant allele is a guanine (G) (Figure 1).3 The rs1799945 variant has been linked with a mild form of HH when the homozygous genotype (GG) is present. According to the National Center for Biotechnology Information (NCBI) SNP database, Northern European populations (locations include the United Kingdom, Ireland, Guernsey, Finland, Germany, Austria, Belgium, France,
Luxembourg, Netherlands, Switzerland, Denmark, Faroe Islands, Iceland, Norway, and Sweden according to 23andMe® have a higher H63D mutant allele frequency than other populations.⁵ Reports suggest that the frequency of the H63D mutant allele among the American population is from 11.5 to 13.5%.⁵,⁷ For this study, we decided to focus on the most common variant, rs1799945 (Figure 1).

![Image](image.png)

**Figure 1.** The location of rs1799945 within the HFE gene on chromosome six. As shown, rs1799945 is located within exon two. The HFE gene contains six total exons.

Transferrin is the transport protein of iron in the blood. The HFE protein normally interacts with the transferrin receptor (TfR) lowering its affinity for transferrin, and ultimately lowering the total amount of iron absorbed.⁸ When mutated, transferrin levels continue to signal to the intestinal cells to export iron into the blood stream. The C282Y and H63D variants are thought to disrupt the normal interaction of the HFE protein with TfR, leading to the disruption of the regulation of iron absorption and resulting in an excessive increase of iron absorbed from the diet.⁹ This leads to the condition known as type 1 hemochromatosis.⁶,⁸ Type 1 hemochromatosis typically affects men more than women, and first show symptoms in those around 40 to 60 years of age.⁹ The HFE protein also regulates the production of hepcidin, which is produced by the liver and decreases the amount of iron absorption, recycling, and storage.⁹

The objective of this study was to evaluate the allele frequency of the H63D mutation and its relationship to a hereditary hemochromatosis diagnosis and population ancestry in nutrition students of a junior/senior level metabolic nutrition course, using raw data from the 23andMe® ancestry kit and a Restriction Fragment Length Polymorphism (RFLP) technique.

**METHODS AND PROCEDURES**

**Participant enrollment**
A total of sixty-nine students in an upper level Metabolic Nutrition course were enrolled in this study. The participants ranged from 20 to 23 years of age. The study protocol was approved by the Institutional Review Board at Virginia Tech. Student saliva samples were deidentified and considered biospecimens. Saliva data was linked to survey and 23andMe® raw data using unique identifiers that were not linked to any identifying information.

Each participant was given a 23andMe® ancestry kit (Mountain View, CA) and completed two surveys. In the first survey, participants were given a variety of questions, one of which was “Have you or a family member been diagnosed with any of the following genetic conditions? (check all).” One of the possible responses was hereditary hemochromatosis. In the second survey, among other questions included in the study, participants were asked to rank their ancestry from their 23andMe® ancestry results. The genotypes of the participants were compared to their previous diagnoses and ancestry. The frequency of the alleles was also calculated.

**DNA isolation**
The collected saliva samples were mixed with a DNA stabilization buffer and stored at 4 °C. The DNA was isolated and purified from 2.5 mL samples of saliva donated by each participant using a published protocol.¹⁰ Briefly, cell lysis solution was added to degrade the cells and expose the DNA. RNA was removed using and RNaseA solution (Thermo-Fisher Scientific, Waltham, MA), while proteinase K solution (Thermo-Fisher Scientific, Waltham, MA) was used to digest the proteins. A protein precipitation solution was used to help separate the proteins from the genomic DNA. Centrifugation separated the proteins and lipids from the DNA. Addition of isopropanol and a glycogen solution followed by an ethanol wash was used to isolate and purify the genomic (g)DNA. The gDNA was rehydrated in Tris-EDTA. The concentration of DNA (μg/μL)(OD 260 x 50 ng/μL) yielded from the isolation was evaluated using a NanoDrop spectrophotometer (Thermo-Fisher Scientific, Waltham, MA), and DNA quality analyzed by agarose gel electrophoresis.

**Primer preparation and Restriction Fragment Length Polymorphism (RFLP) design**
Primers for the polymerase chain reactions were created using NCBI's Primer-BLAST program¹¹ (Figure 2A). The product target size was set from 200 bp to 500 bp. From here, a set of primers was returned that put the variant allele right in the middle. The 3’ primer was reversed so it was converted to the reverse complement. Standard primers were ordered (Eurofins MWG Operon, Luxembourg). A primer stock solution was made by diluting with milli-Q (autoclaved water) to make a 100 μM solution. The forward primer (34.4 nmol) was diluted with 344 μL of pure water, and the reverse primer (39.0 nmol) was diluted with 390 μL of pure water. Twenty μL of the diluted forward and twenty μL of reverse primer were added into a small tube with 60 μL of pure water to create a 10 μM stock solution. The stocks were stored at -20 °C.
The RFLP genotyping technique is an older, but still a useful technique. RFLP viability was determined for the H63D mutation using the sequence amplified by the primers developed above and the Restriction Mapper program (http://www.restrictionmapper.org/). The program determined an RFLP after inputting 10 base pairs of the sequence on both sides of the mutation. The program showed that the BclI restriction enzyme would cut the wild-type allele, but not the variant allele (Figure 2B).

**Figure 2.** Primers and restriction fragment polymorphism design. **A.** The forward and reverse primers surrounding the location of the rs1799945 C/G variant (indicated by the IUPAC symbol “S”, shown in bold, gray highlight) are indicated by bold, underlined font. **B.** The consensus BclI sequence is shown, with the normal location of cut sites indicated by the arrowheads. The sequence surrounding the variant in the mutant and wildtype alleles are shown. Only the wildtype (normal) allele is cut by BclI, as the variant sequence no longer matches the consensus.

**Polymerase Chain Reaction (PCR)**

The polymerase chain reaction was used to amplify the HFE region surrounding rs1799945 using the primers shown in **Figure 2A.** Each PCR tube contained 10.0 μL of 2X PCR master mix (GoTaq, Promega, Madison, WI), 0.5 μL of primer stock, 2.0 μL of template genomic DNA from individual volunteers and 7.5 μL of nuclease-free water. Following mixing, the tubes were centrifuged for a few seconds to insure all reagents were at the bottom of the PCR tube. The samples were then run in an Eppendorf (Hauppauge, NY) PCR machine with an initial denaturation cycle of 95 °C for 10 minutes, followed by 30 cycles of a denaturation cycle of 95 °C for 30 seconds, an annealing cycle of 55 °C for one minute, an extension cycle of 72 °C for one minute. After a final extension of 72 °C for 10 minutes, the samples were stored at a temperature of 4 °C.

**Agarose gel electrophoresis**

After the PCR amplification, the samples were examined for a correctly-sized PCR amplification using a 1.5% agarose TAE gel containing ethidium bromide. Six (6) μL of each sample was mixed with 2 μL of loading dye and dispensed into their respective wells. The gel electrophoresis was carried out at 83 V for around 25 minutes. Once the bands had appeared to separate, the gels were viewed under a UV light using the BioRad Chemidoc System (Hercules, CA). Visualization of a 380bp band indicated a positive PCR amplification.

**RFLP genotyping**

Amplified samples from the PCR reaction were digested by the BclI enzyme (Thermo-Fisher Scientific, Waltham, MA) by adding 10 μL of the DNA, 18 μL of nuclease-free water, 2 μL of 10x Buffer G, and 1.5 μL of the BclI enzyme to a 0.5 mL centrifuge tube. The mixtures were mixed gently and incubated for approximately 1.5 hours at 37 °C. After incubation, digested samples were analyzed using the gel electrophoresis protocol previously described. The homozygous wild-type variants showed two bands of 244 bp and 136 bp respectively, while the homozygous mutant variants showed a single band of 380 bp. Heterozygous samples showed all three bands. Visualization of the bands using the BioRad Chemidoc System (Hercules, CA), allowed for the determination of the genotypes. This method of genotyping was used for the samples of the 17 participants that did not provide 23andMe® data or the second survey. Six samples from the 52 participants that provided 23andMe® data and the second survey were also genotyped using this method as controls to confirm the RFLP results.
Survey data analysis
Data on the participants’ gender, ancestries, HH diagnoses, and genotypes were collected from the survey results. This information was broken down into the following groups: “Diagnosis with HH,” “Northern European Ancestry,” “Non-Northern European ancestry,” “Unknown ancestry,” “Male,” and “Female.” These groups were compared to the wild-type, heterozygous, and mutant genotypes.

Statistical Analysis
The significance of the genotypes relation to a HH diagnosis was assessed using an online genotype analysis program (http://www.physics.csbsju.edu/cgi-bin/stats/exact_form.sh?nrow=2&ncolumn=3), which used Fisher’s Exact Test. Significance was set a priori at $P < 0.05$.

RESULTS

PCR Results
Figure 3A shows representative PCR results that were obtained for 23 participants who originally provided saliva for genomic DNA isolation. Seventeen were from volunteers who provided saliva samples, but who did not provide genotype data from 23andMe®. An additional 6 samples served as study controls as they were from volunteers who provided both saliva samples, and genotype data from 23andMe®.

RFLP genotyping
Twenty three samples that successfully amplified were genotyped with representative samples shown in Figure 3B. Of the 23 samples, 17 were genotyped as WT and contained only the normal alleles (CC), 4 were genotyped as heterozygous (HET) containing one normal allele and one variant allele (CG), and 2 were genotyped as mutant (MUT), or homozygous for the mutant alleles (GG). These genotypes are presented in Table 1.

In silico analysis and survey results
In silico analysis included a total of 69 participants. Fifty-two participants provided saliva samples, genomic data from 23andMe®, and surveys with phenotypic information. Seventeen participants provided saliva samples, but did not provide 23andMe® data; genomic DNA from these participants were genotyped using the RFLP technique. Genotype results for males and females are shown in Table 2. None of the volunteers indicated that they had ever been diagnosed with HH, and therefore no significant relationship between the mutant allele and an HH diagnosis was found ($P=1.000$). However, for those volunteers that responded to the second survey, a significant relationship was found in that the mutant allele was significantly related to those of Northern European descent ($P=0.0066$). For those individuals, the mutant allele frequency among this population was found to be 13.04% (Table 3).
Table 1. DNA samples and associated genotypes determined by RFLP analysis.
*These samples were controls, and were genotyped by 23andMe® and the RFLP technique.

<table>
<thead>
<tr>
<th>Sample Code</th>
<th>Genotype</th>
<th>Sample Code</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>C004</td>
<td>CC</td>
<td>C056*</td>
<td>CG</td>
</tr>
<tr>
<td>C006</td>
<td>CC</td>
<td>C057*</td>
<td>CC</td>
</tr>
<tr>
<td>C014</td>
<td>CC</td>
<td>C061</td>
<td>CC</td>
</tr>
<tr>
<td>C020</td>
<td>CC</td>
<td>C069</td>
<td>CC</td>
</tr>
<tr>
<td>C023*</td>
<td>CC</td>
<td>C075*</td>
<td>GG</td>
</tr>
<tr>
<td>C027</td>
<td>CC</td>
<td>C076</td>
<td>CG</td>
</tr>
<tr>
<td>C029</td>
<td>CC</td>
<td>C077</td>
<td>CC</td>
</tr>
<tr>
<td>C033</td>
<td>CC</td>
<td>C084</td>
<td>CC</td>
</tr>
<tr>
<td>C043*</td>
<td>CG</td>
<td>C089*</td>
<td>GG</td>
</tr>
<tr>
<td>C044</td>
<td>CC</td>
<td>C091</td>
<td>CC</td>
</tr>
<tr>
<td>C046</td>
<td>CG</td>
<td>C095</td>
<td>CC</td>
</tr>
<tr>
<td>C047</td>
<td>CC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Genotypes from both 23andMe® and RFLP genotyping were combined to quantify the genotype frequency of the H63D mutation in males versus females.

<table>
<thead>
<tr>
<th>TOTALs per category (percent)</th>
<th>C:C Homozygous WT</th>
<th>C:G Heterozygous HET</th>
<th>G:G Homozygous MUT</th>
<th>Fisher's Exact Test (P-value)</th>
<th>Mutant Allele Frequency (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (Total)</td>
<td>69 (100%)</td>
<td>53</td>
<td>14</td>
<td>2</td>
<td>N.A.</td>
</tr>
<tr>
<td>% of Population</td>
<td>76.8</td>
<td>20.3</td>
<td>2.9</td>
<td>N.A.</td>
<td>13.04</td>
</tr>
<tr>
<td>Diagnosis with HH</td>
<td>0 (0 %)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>P = 1.0 (Not significant)</td>
</tr>
<tr>
<td>Females</td>
<td>54 (78.2%)</td>
<td>41</td>
<td>11</td>
<td>2</td>
<td>N.A.</td>
</tr>
<tr>
<td>Males</td>
<td>15 (21.8%)</td>
<td>12</td>
<td>3</td>
<td>0</td>
<td>N.A.</td>
</tr>
</tbody>
</table>

Table 3. Genotypes from the both 23andMe® and RFLP genotyping were combined to quantify the genotype frequency of the H63D mutation. Ancestry analysis was obtained by self-report of the volunteers using 23andMe generated data.

<table>
<thead>
<tr>
<th>TOTALs per category (percent)</th>
<th>C:C Homozygous WT</th>
<th>C:G Heterozygous HET</th>
<th>G:G Homozygous MUT</th>
<th>Fisher's Exact Test (P-value)</th>
<th>Mutant Allele Frequency (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (Total)</td>
<td>69 (100%)</td>
<td>53</td>
<td>14</td>
<td>2</td>
<td>N.A.</td>
</tr>
<tr>
<td>% of Population</td>
<td>76.8</td>
<td>20.3</td>
<td>2.9</td>
<td>N.A.</td>
<td>13.04</td>
</tr>
<tr>
<td>Diagnosis with HH</td>
<td>0 (0 %)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>P = 1.0 (Not significant)</td>
</tr>
<tr>
<td>Northern European Ancestry</td>
<td>47 (68%)</td>
<td>33</td>
<td>12</td>
<td>2</td>
<td>P = 0.0066 (highly significant)</td>
</tr>
<tr>
<td>(responded to 2nd survey)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Northern European ancestry</td>
<td>5 (7.2%)</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>N.A.</td>
</tr>
<tr>
<td>(responded to 2nd survey)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown ancestry (did not respond to 2nd survey)</td>
<td>17 (N.A.)</td>
<td>15</td>
<td>2</td>
<td>0</td>
<td>N.A.</td>
</tr>
</tbody>
</table>
DISCUSSION
Type I hemochromatosis is one of the most common genetic conditions in the United States and worldwide. In our results, 20.3% of individuals in a college nutrition course were mutant allele carriers, and 2.9% of the individuals were homozygous for the allele. These results are similar to published reports of the H63D mutant allele frequency among Americans (11.5 to 13.5%), as well as a recent study of a Greek population for rs1799945, which found that up to 25% of individuals were carriers of the mutant allele. We found no individuals of non-Northern European descent that were carriers. This low carrier rate among those of non-Northern European descent is consistent with previously published reports. For example, as shown by a study examining rs1799945 in Russian and Asian groups, the frequency is only 3-8%. Our data also indicates a significant correlation with Northern European ancestry, as determined by the Fisher’s Exact Test. Overall the study verifies the validity of studying single gene variants within small populations using both laboratory-based and in-silico based analyses.

Among all the participants, none of them had ever been diagnosed with hemochromatosis. As expected, there was no association between the mutant (G) allele and a hemochromatosis diagnosis. However, there may be reasons for why neither of the two individuals who were homozygous for the mutant allele have ever been diagnosed with HH. First, both these participants were female. This is relevant, because most females menstruate every month. With blood loss comes iron loss, so this release of blood from the body might be enough to reduce the overall levels of iron to prevent a diagnosis of HH. Also, this population was relatively young (20-23 years of age). Typically, signs and symptoms of HH do not appear until around 40 to 60 years of age, meaning that this population might be too young to have severe enough symptoms of HH to warrant testing for a diagnosis.

The results of this study shows that determining participants genotypes using the 23andMe data can provide genetic risk information prior to onset of HH symptoms. While current treatments for HH are effective in lowering the iron levels within acceptable ranges, diseases caused by HH such as hypogonadism, cirrhosis, arthritis, and insulin-dependent diabetes are often irreversible. By simply using an direct-to-consumer DNA kit, people of this age could predict their predisposition to HH and other genetic disorders before the symptoms arise at a relatively low cost. This would help individuals with HH or other genetic disorders, begin to manage and limit the possible harmful effects of genetic disorders.

There are a few improvements that could be made upon this study. First, not all the samples that underwent PCR were successfully genotyped due to technical errors in isolation, storage, or analysis. Due to the biospecimen nature of our study, volunteers could not be re-contacted to provide saliva samples for re-analysis. A new direction for this research would be to compare female populations before and after menopause, as previous findings suggest that males present greater phenotypic effects correlating with the mutant genotype than females. Also, future studies should further explore direct-to-consumer DNA kits as a means to predict HH and other harmful genetic disorders. Both of these suggestions could be done using a different recruiting strategy, such as recruiting from the entire campus including both students and faculty/staff, rather than from a single class. Finally, measuring blood serum levels of iron and transferrin may help strengthen the findings and show a better relationship between H63D and iron overload, even without the diagnosis of HH.

ACKNOWLEDGEMENTS
The authors thank the Office of Undergraduate Research at Virginia Tech for a grant that ASA and DJG secured, to fund our course-based undergraduate research program.

REFERENCES


ABOUT STUDENT AUTHOR
Tyler Ferqueron graduated in May 2019 with a B.S. in Human Nutrition, Foods, and Exercise, science option, from Virginia Tech. He plans to pursue a M.D. or M.D./Ph.D.

PRESS SUMMARY
Hereditary hemochromatosis (HH) results in a faulty iron sensing mechanism and causes the human body to absorb too much iron from the diet. Stored excess iron in the body can lead to several health conditions. Certain genetic mutations have been shown to lead to this disease. One of these mutations (referred to as H63D) within the HFE gene was studied in students aged 20-23 years old in an upper-level metabolic nutrition course to determine how frequently it occurred, with what population ancestry it occurred in, and whether carriers of the variant gene had received an HH diagnosis. It was found that there was no relationship between having this mutation and having been diagnosed with HH in this population, likely due to the young age of the population. However, a significant correlation was found between those of European decent and being a H63D variant carrier, consistent with other published findings. Direct-to-consumer DNA kits similar to the one used in this study may be helpful in identifying HH before symptoms arise.