NATIONAL CENTER FOR CASE STUDY TEACHING IN SCIENCE

Precision Over Presumption:
Connecting Pharmacogenomics and Genetic Ancestry

by
Erica L. Gerace[1,2], Alexa Wnorowski[1], and Sarah Wojiski[1]
[1]Genomic Education, The Jackson Laboratory for Genomic Medicine, Farmington, CT
[2]Department of Genetics, Yale School of Medicine, New Haven, CT

Part I – Splitting Headaches

Sonya described her headache as something she could feel in her stomach. "It starts as a throbbing pain in my head that is unbearable and then I start to feel incredibly nauseous."

The neurologist, Dr. Elena Carter, nodded at her patient. "I see. How long does the headache last?"

Sonya replied, "It can last anywhere from a few hours to a few days."

"Does anything help with pain relief? What have you tried so far?" asked Dr. Carter.

"Over-the-counter painkillers can't touch it. When the throbbing starts, I get in bed and lie in the dark. Light makes the throbbing and nausea worse. I just lie there with my eyes closed until it subsides," Sonya detailed.

"This sounds like a migraine. These headaches are a bit of a mystery but have the symptoms you are describing, among others. It would be useful if you could identify what might be triggering the migraine for you. For some people it's stress or hormones, for others it's lack of sleep or certain foods. Maintaining a journal and looking for a common pattern in the days leading up to your migraine could be helpful in pinpointing and hopefully decreasing or eliminating the triggers. But I can also prescribe you a medication you can take if you feel a migraine coming on."

Questions

1. What are the symptoms of a migraine headache?

2. What are the common triggers of a migraine headache?

3. What type of doctor is a neurologist?

Part II – Pharmacogenomics

Later in her office, Dr. Carter looked over a pamphlet on her desk. Migraine medications were "finicky," meaning their effectiveness varied by person, and finding an effective one for an individual could take some time. A new medication, AxonMax, had come out a few months ago and seemed to be more effective at treating migraines than some existing treatments. Dr. Carter recalled the conference she had recently attended and how her colleagues had raved about the success they had seen with this new medication, including for patients with stubborn migraines. *It's worth giving this one a try,* she thought to herself.

After searching for the medication in the database, Dr. Carter paused as she read the dosing guide on her computer screen. Dosing varied, and the guide suggested that the patient's reported race and ethnicity would determine the starting dose. *Hmm...,* she thought to herself. *This raises so many questions.*

Dr. Carter called down to the pharmacy to inquire. She asked the pharmacist, "Can you explain the dosing for AxonMax, the new migraine medication?"

The pharmacist responded, "Let me see . . . yeah, this is one of those medications that is affected by metabolism. Some people break down the drug faster than others and if it's metabolized too quickly, it doesn't work for a patient."

"Is this genetic? The dosing guide mentioned something about race and ethnicity," asked Dr. Carter.

"Yes, metabolism of this drug involves the *CYP2C19* gene. This gene comes up a lot in pharmacogenomics and a patient's response to a medication. Also, there are variants of this gene that are common to certain populations," the pharmacist replied.

"I see. I'll investigate this more. Thanks for your help!" replied Dr. Carter.

Questions

1. Think about the word "pharmacogenomics" and come up with a definition for this word.

2. What do you think "metabolize" means when referring to medicines in the body?

3. What do you predict happens if a drug is metabolized too quickly?

4. What do you predict happens if:
   a. A person metabolizes a drug too quickly; would they need a larger or smaller dose of the drug?

   b. A person receives a large dose of a medication but metabolizes this drug slowly?

Part III – Cytochrome P450

Later that day, Dr. Carter continued her investigation and pulled up some information on the gene *CYP2C19*. The field of pharmacogenomics, or any genomics for that matter, was only just emerging when she went to medical school and so she needed to learn more. Here is what she found:

CYP2C19

The cytochrome P450 (CYP) families of genes encode a series of enzymes, many of which are involved with metabolism in the liver (Figure 1). The gene CYP, family 2, subfamily C, polypeptide 19 (*CYP2C19*) encodes an enzyme that catalyzes many reactions including those involved in making cholesterol and other lipids, and importantly, reactions required for metabolizing drugs and medicines.

Variants in *CYP2C19* can affect how fast the enzyme works. If the enzyme works slower, it is going to take longer to metabolize medicines or break them down, and the drug will stay in the body for a longer time. If the enzyme works faster, medicines will be metabolized faster, and the drug may not be in the body long enough to work. Along with other biological and environmental factors, this variation in metabolic speeds can influence drug effectiveness or the number of side effects.

A single nucleotide substitution in a specific position in the DNA sequence of *CYP2C19* changes the function of the enzyme encoded by this gene (See Figure 2). When there is a G nucleotide present (called allele G) the encoded enzyme functions efficiently and there is effective drug metabolism in the liver. When there is an A nucleotide present (called allele A), the mRNA created from the gene is not correctly processed (spliced) resulting in an enzyme that is not functional, which impairs drug metabolism in the liver.

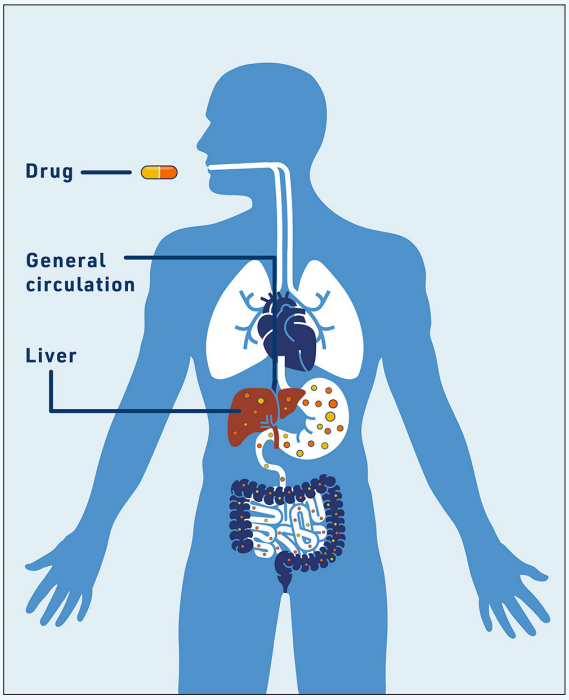

Figure 1. Illustration of drug metabolism. Once a drug is ingested orally, it is absorbed through the digestive track and enters the bloodstream. The drug is then carried to the liver where it is metabolized. (Image by The Jackson Laboratory © 2023.)

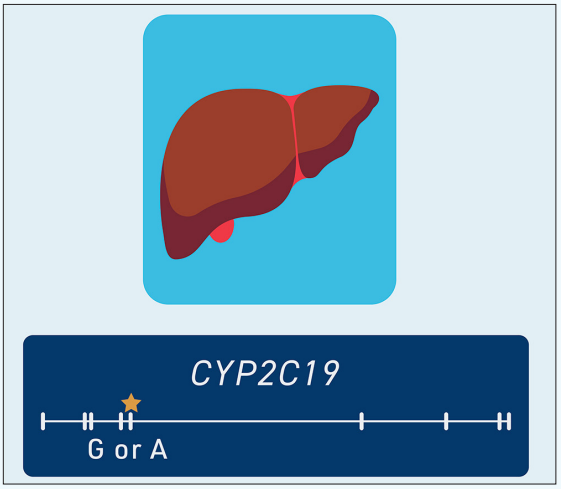

Figure 2. *CYP2C19* genomic variant. A single nucleotide substitution (G or A) creates two variants in the *CYP2C19* gene that impact the function of the enzyme this gene encodes. This image of the *CYP2C19* gene shows exons as small boxes on a line representing intron sequences. The G/A variant is in the fifth exon. (Image by The Jackson Laboratory © 2023.)

Questions

1. Which variant of *CYP2C19* is associated with slow drug metabolism? Which variant is associated with fast drug metabolism?

2. Recall that humans have two alleles for most genes, including *CYP2C19*. For the following genotypes, hypothesize about the rate of drug metabolism (slow, fast, or intermediate) and provide an explanation for your answer.
   a. GG

   b. GA

   c. AA

3. Based on your hypotheses above:
   a. Rank the genotypes GG, GA, and AA from slowest to fastest metabolism. Do you think any of the genotypes have the same metabolic rate or will they all be different?

   b. Rank the genotypes GG, GA, and AA from lowest to highest likelihood of having side effects from a medication.

   c. What do you notice about the relationship between metabolic rate and the likelihood of side effects? Use your answers to (a) and (b) above to describe the relationship.

Part IV – How Common are *CYP2C19* Variants?

Dr. Carter now understood the role of *CYP2C19* in the metabolism of medicines like AxonMax and how the amount of drug a patient took could lead to the drug staying in the body for too long or not long enough. Considering this was a drug to treat migraines, which have lots of symptoms, the wrong dosing could cause unnecessary side effects or make the drug ineffective. Dr. Carter decided to look deeper into the racial and ethnic dosing guidelines mentioned in the database and by the pharmacist on the phone.

Dr. Carter typed "CYP2C19 variants population ethnicity" into her browser window. Among the search results describing population frequencies, she was drawn to one link that led her to a study called *All of Us*. She read the following:

> *The National Institutes of Health created the All of Us Research Program to study the connection between genetics, environment, and lifestyle and the impact of all three on human health. Individuals from across the United States have volunteered to provide genetic and other biological and health data for scientists and clinicians to study. The goal is to compile data from one million different people.*

"Wow, what a cool program!" Dr. Carter said aloud to herself. Then she noticed there was a way to investigate specific genes and gene variants on the *All of Us* Research Hub (https://www.researchallofus.org/) and that data on *CYP2C19* was publicly available. She typed "CYP2C19" into the search bar, scanned the results, and located the information about the G and A variants she had read about previously. She focused on the data detailing the allele frequency, which is a measure of the number of a specific allele (or gene variant) relative to the total number of alleles present in a population. (For example, if there are 50 people in a given population, then there are 100 total alleles for a specific gene because each person has two alleles. If the allele frequency for a specific variant is 0.5, that means half or 50 of the 100 total alleles are that variant.)

The data Dr. Carter found on *All of Us* about the *CYP2C19* gene in the US population is shown in Table 1 below.

Table 1. Frequencies of *CYP2C19* alleles G and A. *CYP2C19* alleles in the US population based on a sample of individuals providing DNA for sequence analysis (*All of Us* data collected 09/26/2025).

Variant	*Allele Count*	*Total Allele Number*	*Allele Frequency*
CYP2C19 allele G	702,473	829,554	0.8468
CYP2C19 allele A	127,081	829,554	0.1532

Help Dr. Carter make sense of this table by answering the questions below.

Questions

1. What is the total number of alleles counted in this sample population?

2. How many alleles are G?

3. How many alleles are A?

4. Which allele is more common (frequent) in the US population?

5. If humans each have two alleles for each gene in their DNA, how many people were included in this sample population?

6. In your own words, define "allele frequency."

Challenge Question: Using the allele frequencies in Table 1, the Hardy-Weinberg Equilibrium equation, and your answer to Question 5 in Part IV, calculate the number of people in this data set who are homozygous for allele A.

Part V – *CYP2C19* Variants by Ancestry

Based on the overall population frequency, Dr. Carter realized it was more likely that her patient was a fast drug metabolizer. With that thought in mind, Dr. Carter referred again to the dosing guide in the drug information online, which described how to provide a lower dose for patients who identify as Asian. She pulled up Sonya's medical record to see if race or ethnicity was self-reported. The box for White/Caucasian was checked. Before putting in the prescription information to have the drug sent to the pharmacy, Dr. Carter decided to look more closely at the actual population frequencies in the *All of Us* database.

Dr. Carter noted that the population data included in the *All of Us* Research Hub was considered *genetic ancestry*, which involved using genomic analysis and the identification of specific variants in the genome to determine a person's ancestry. She knew that genetic ancestry could be very different from *reported race/ethnicity*, which is the population(s) that an individual identifies with.

Here is the data she found:

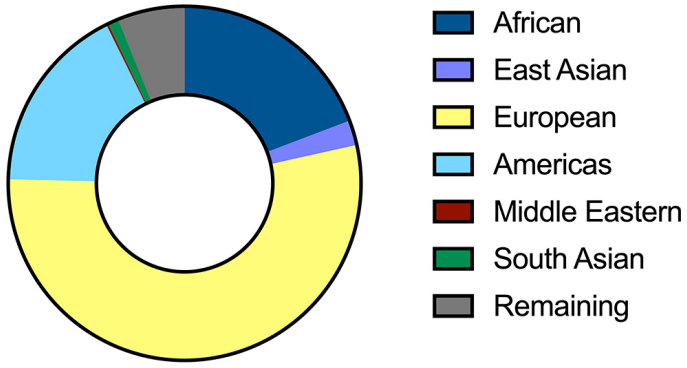

Figure 3. Proportions of *CYP2C19* allele A in the *All of Us* database by genetic ancestry. The percentage of *CYP2C19* A alleles in the US population sample categorized by ancestry: African (dark blue); East Asian (purple); European (yellow); Latin American (light blue); Middle Eastern (red); South Asian (green); and any remaining individuals not aligning with one of the named ancestries (gray). Ancestry was determined by short-read whole genomic sequencing. Data collected from *All of Us* SNV/Indel Variant data browser for rs4244285 on September 26, 2025.

Questions

1. Use Figure 3 to answer the following:
   a. Which ancestry group has the greatest representation in this sample?

   b. Estimate the percentage of A alleles from this ancestry group.

   c. Estimate the percentage of A alleles from individuals with:
      i. African ancestry

      ii. East Asian ancestry

d. Based on your perception of the racial/ethnic make-up in the United States, do you think this figure accurately represents the US population? Explain.

The next set of questions are focused on the data table (Table 2) below and you will use the allele counts in this table to answer the remaining questions in this section.

Table 2. Allele A counts stratified by genetic ancestry. The number of A alleles in the US population categorized by genetic ancestry based on a sample of individuals providing DNA for sequence analysis (*All of Us* data, 09/26/2025). The allele count represents the number of A alleles counted and the allele number is the total number of *CYP2C19* alleles that were counted for each ancestry.

Ancestry	*Allele A Count*	*Total Allele Number*	*Allele Frequency*
African	28,393	159,636	0.1779
East Asian	5,826	18,878	
European	65,107	446,634	
Latin American	16,851	143,690	
Middle Eastern	158	1,620	
South Asian	2,781	8,092	
Remaining	7,965	51,004	
Total	127,081	829,554	0.1532

2. Analyze Table 2 to answer the following questions:
   a. Which ancestry group had the greatest number of alleles counted in the population?

      i. Is this answer the same or different from your answer in Part V, Question 1a?

   b. How many individuals in this study have East Asian ancestry? (Hint: remember that humans have two alleles for *CYP2C19*.)

   c. Use the allele count and allele numbers to calculate the allele frequencies for each ancestry group and fill in the blank cells in Table 2.

   d. Which ancestry group has the highest allele frequency for allele A? Which one has the lowest frequency for allele A?

   e. What is the frequency of allele G for the East Asian ancestry group?

Challenge Questions: Use the allele frequencies you calculated and added to Table 2 and the Hardy-Weinberg equation to answer the following questions:

a. What is the frequency of the AA genotype in population with East Asian ancestry?

b. Assuming this frequency represents the entire population of people with East Asian ancestry in the US, what percentage of the East Asian population is a slow metabolizer?

c. What is the frequency of the AA genotype in population with European ancestry?

d. Assuming this frequency represents the entire population of people with European ancestry in the US, what percentage of the European population is a slow metabolizer?

e. Given the frequencies you calculated in parts *a* and *c* of this question set, describe the population overall. Are most people fast or slow metabolizers?

f. Compare the frequencies you calculated in parts *a* and *c* of this question set. How many times more likely is a person with East Asian ancestry to be a slow metabolizer than a person with European ancestry?

Part VI – Reported Identity and Genotype

Dr. Carter realized that given the allele frequency data included in the *All of Us* database, she was not comfortable prescribing a medication based on someone's reported racial and ethnic identity. If Sonya was White/Caucasian as indicated on her chart, she would likely have European ancestry, and Dr. Carter could assume that the chance of her being a slow metabolizer was quite small and could prescribe the higher dose. Yet, the consequences could be significant if she was wrong with this assumption.

"There's really only one way to find out . . . genetic testing!" Dr Carter exclaimed to herself. She knew that pharmacogenomics was a growing field, and that it was possible to test gene variants like those in *CYP2C19* to gauge how a patient would respond to a specific drug. The availability of these tests was increasing, and the costs were decreasing, making this a great option to confirm the patient's genotype.

A few weeks later, Sonya agreed to receive genetic testing after Dr. Carter explained to her the purpose and why it would be important for correct dosing for the medication that could treat her migraines. Here are the results that were sent to Dr. Carter:

Test Result: *CYP2C19* rs4244285
Allele 1: A
Allele 2: A

Questions

1. Is Dr. Carter's patient a fast or slow metabolizer of AxonMax?

2. Why is this a surprising result?

3. What would have been the result if Dr. Carter had given her the recommended dosage for people of European descent?

4. What are the dangers of assuming someone's genotype based on their reported racial/ethnic identity?