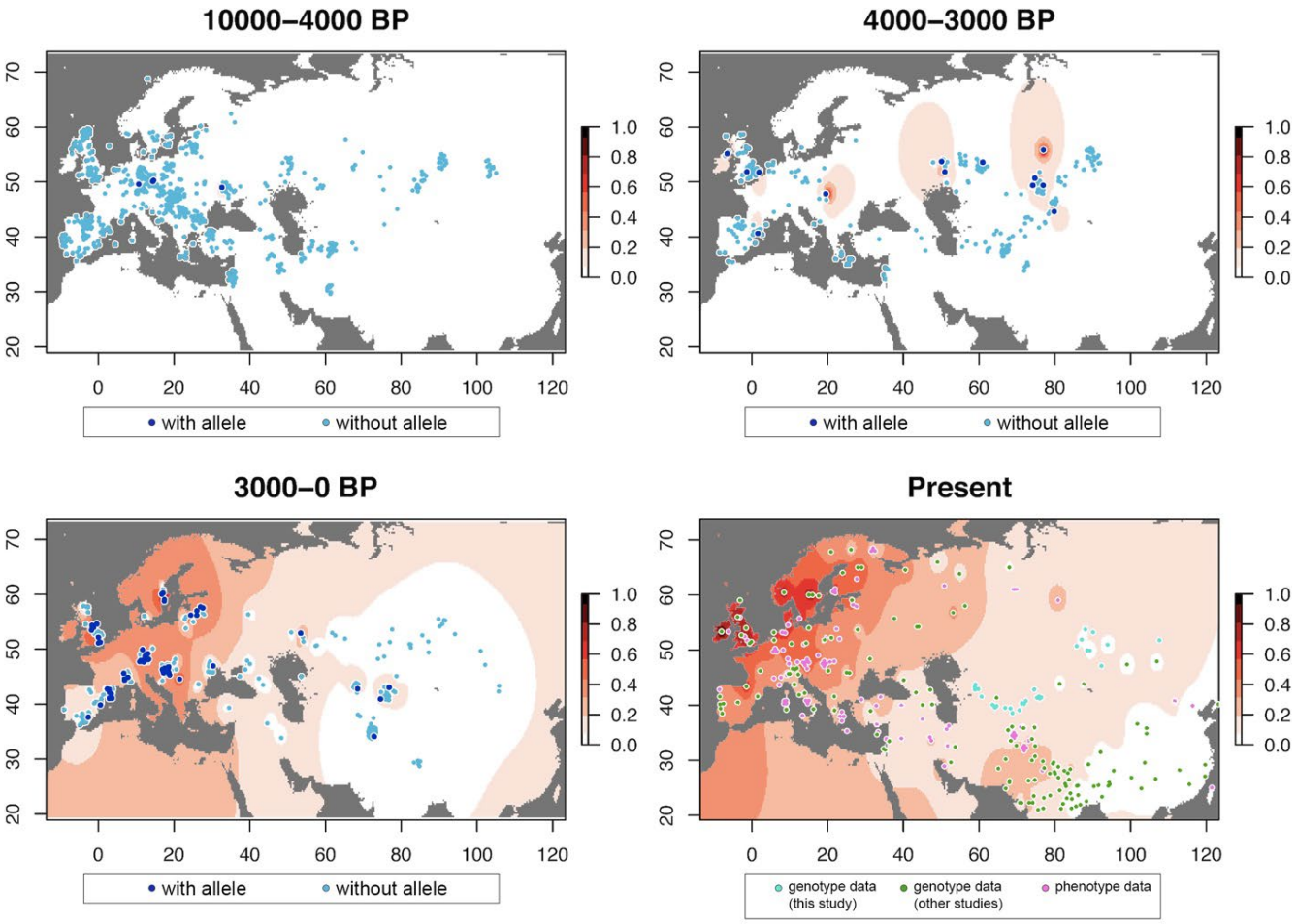

Caption: *Maps showing how a lactase-persistence allele spread through human populations over the last 10,000 years. Dots represent specific individuals or populations sampled in the study. The red/orange shading shows the estimated frequency of the allele among people in different geographic regions, based on the samples.*

Three maps are for "ancient" time periods before present (BP): "10000–4000 BP," "4000–3000 BP," and "3000–0 BP." The frequency of the lactase-persistence allele (red/orange shading) was estimated based on the distribution of individuals with the allele (dark-blue dots) and without the allele (light-blue dots).

The "Present" map is for the present day. The frequency of the lactase-persistence allele (red/orange shading) was estimated based on data from specific populations (dots). Populations are colored based on the type of data that was used: genotype data from this study (cyan dots), genotype data from other studies (green dots), or phenotype data (magenta dots).

BACKGROUND INFORMATION

As babies, all mammals (including humans) produce an enzyme called **lactase**. Lactase breaks down **lactose**, the main sugar in milk, into other sugars that the body can use to get energy.

Around the age of 5, most humans stop producing lactase. People with this trait, called **lactase nonpersistence**, are **lactose intolerant**, meaning they cannot break down lactose as adults. They may experience pain, gas, bloating, and diarrhea after drinking milk.

However, some people keep producing lactase even as adults. This trait, called **lactase persistence**, allows them to break down lactose (and thus get more energy from milk) throughout their lives.

People with lactase persistence have been shown to have specific variations in their DNA called **lactase-persistence alleles**. Today, lactase-persistence alleles are most common in people with certain ancestry living in African, Arab, or European regions. Populations in these regions were traditionally **pastoral**, meaning that they raised animals for food (including milk) since about 10,000 years ago.

One hypothesis suggests that lactase persistence was strongly selected for in pastoral populations, because people who could break down lactose could get more energy from their food. To investigate this hypothesis, scientists traced the spread of a common lactase-persistence allele through Asia and Europe over the last 10,000 years. Their study included populations in Central Asia (such as Mongolia and Kazakhstan), which were also traditionally pastoral.

To determine where and how common the lactase-persistence allele was in the *past*, the scientists analyzed DNA data from ancient human remains. They determined the presence or absence of the allele in many ancient individuals, whom they grouped into three time periods: 10000–4000 BP (10,000 to 4,000 years before present), 4000–3000 BP (4,000 to 3,000 years before present), and 3000–0 BP (3,000 to 0 years before present).

To determine where and how common the lactase-persistence allele is in the *present*, the scientists analyzed data from present-day populations, specifically using individuals who had ancestors living in the same region over multiple generations. The scientists considered **genotype data**, which is obtained using DNA sequencing, and **phenotype data**, which is obtained using tests that measure a person's ability to digest lactose.

Based on these data, the scientists estimated how common the allele was among people in different regions over time. Their estimates of the allele's frequency are shown by red/orange shading on the maps.

INTERPRETING THE FIGURE

Lactase persistence has evolved multiple times in different human populations since the domestication of dairy animals. It is associated with multiple mutations in the regulatory region ("genetic switch") of the lactase gene, each of which can cause lactase production to remain permanently "on." Different populations may have different mutations that cause lactase persistence. In other cases, the same mutation spread across populations or arose independently.

This study (Ségurel et al. 2020) focuses on a specific lactase-persistence allele called the **–13.910*T allele**, which is due to a single-nucleotide change (C to T) in the lactase switch. The number "–13.910" indicates that the mutation is 13,910 nucleotides upstream from the lactase gene. This allele acts in a mostly dominant Mendelian fashion, meaning that if a person has at least one copy of the allele, they will have lactase persistence.

The −13.910*T allele is first observed around 6,000 BP in Ukraine and spread through human migration and natural selection. Lactase persistence in Eurasia is typically associated with this allele. (In other parts of the world, lactase persistence may be associated with other alleles. For example, multiple alleles are associated

with lactase persistence in North Africa and East Africa, which makes it more difficult to study. This is why the authors did not include populations from those regions in this study.)

The maps in the figure show how the −13.910*T allele spread across Eurasia over the last 10,000 years, based on evidence from ancient DNA and present-day data. Some takeaways from the maps are as follows:

- During the first time period, **10000–4000 BP**, most of the individuals analyzed did not have the lactase-persistence allele. The allele is found in only three individuals (all in Europe) out of hundreds sampled, and the estimated population frequency of the allele is relatively low throughout Eurasia.
- By the second time period, **4000–3000 BP**, the allele became more common. Its estimated frequency is higher in certain areas — not only in Europe, but also in Central Asia.
- By the third time period, **3000–0 BP**, the allele's estimated frequency increased dramatically in places like Northern Europe but stayed relatively low in Central Asia.
- In the **present**, the allele's estimated frequency is still high in places like Northern Europe compared to Central Asia. The allele has also become more common in certain parts of the Arabian Peninsula, Pakistan, and Russia.

These results suggest that lactase persistence was *not* strongly selected for in all pastoral populations — particularly those in Central Asia. These populations traditionally raise many animals for milk, including horses, goats, yaks, and/or camels. They are also some of the only populations to traditionally consume horse's milk, which has even more lactose than cow's milk. However, the frequency of the lactase-persistence allele still stayed relatively low in Central Asia, suggesting that lactase persistence did not provide a strong selective advantage there.

The authors speculate that one reason there was less selection for lactase persistence in Central Asia is because the populations there traditionally don't drink milk raw. Instead, they ferment it to make other food products, such as yogurt and kefir. These products are easier to digest than raw milk, because the bacteria used for fermentation produce their own lactase, which decreases the amount of lactose in the food. So, people can consume these products even if they don't produce lactase themselves.

The authors suggest that this is an example of a "cultural adaptation" (i.e., developing cultural practices for consuming milk that circumvent needing lactase), in contrast to a "genetic adaptation" (i.e., natural selection for lactase-persistence alleles).

TEACHING TIPS

Prompt students to explain the parts of the figure as applicable:

- Graph type: Geographical maps, heat maps, dot distribution maps
- *X*-axis: Longitude (degrees)
- *Y*-axis: Latitude (degrees)
- Maps before present (BP): Distributions of sampled "ancient" individuals, and overall −13.910*T lactase-persistence allele frequencies estimated from these samples, over three past time periods. Scientists analyzed DNA from each individual's remains. Ages of the remains were determined using methods such as radiocarbon dating.
    - Blue Dots: Each dot represents a different individual who was sampled. Dark blue indicates the individual had the −13.910*T allele. Light blue indicates the individual did *not* have the allele.

  - Red/Orange Shading: Population frequency of −13.910*T allele (i.e., the estimated proportion of individuals in the population with the allele), estimated based on the distribution of individuals with and without the allele. Lighter areas indicate lower frequencies, and darker areas indicate higher frequencies.
  - BP: An abbreviation for "before present." The year used for the "present" is often 1950. For example, 10,000 BP means 10,000 years before 1950. (Because the remains are so ancient, the difference between 1950 and more recent years is typically negligible.)
  - 10000–4000 BP: 10,000 to 4,000 years before present
  - 4000–3000 BP: 4,000 to 3,000 years before present
  - 3000–0 BP: 3,000 to 0 years before present
- Present map: Distributions of sampled present-day populations, and overall −13.910*T lactase-persistence allele frequencies estimated from these samples. In each population, the frequency of the allele was investigated using genotype or phenotype data.
  - Cyan/Green Dots: Each dot represents a population investigated using *genotype* data — i.e., DNA sequencing. *Cyan* indicates data collected by the authors (Ségurel et al.), and *green* indicates data obtained from other studies. The genotype data was used to determine the −13.910*T allele frequency in these populations.
  - Magenta Dots: Each dot represents a population investigated using *phenotype* data — i.e., using tests that measure a person's ability to digest lactose, such as a blood glucose test or a hydrogen breath test. The phenotype data was used to infer the −13.910*T allele frequency in these populations. (Though this phenotype can also be caused by other lactase-persistence alleles, it has been shown to be strongly correlated with just the −13.910*T allele in Eurasia.)
  - Red/Orange Shading: Population frequency of −13.910*T allele (i.e., the estimated proportion of individuals in the population with the allele), estimated by extrapolating from the sample data. Lighter areas indicate lower frequencies, and darker areas indicate higher frequencies.

Complement this Data Point with the following related resources:

- The short film *Got Lactase? The Co-evolution of Genes and Culture*, and its accompanying film activity, can be used to introduce the phenomenon of lactase persistence and provide more information on the genetics and evolution of this trait.
- The activity "Blood Glucose Data Analysis" discusses some of the "phenotype" tests that measure a person's ability to digest lactose.
- The planning tool *Storyline Viewer* contains two storylines (phenomena-driven lesson sequences) that explore the evolution and genetics of lactase persistence.
- The 2017 scientific review paper "On the Evolution of Lactase Persistence in Humans" provides more background on lactase persistence, including how it is studied and key research findings.

DISCUSSION QUESTIONS

Figure Interpretation

- Describe the *present-day* distribution of lactase persistence across Europe and Asia. What are some general patterns you observe? (For example, where are people more likely to have vs. not have the allele?) Why do you think these patterns are the way they are?
- Describe the distribution of the lactase-persistence allele over the last 10,000 years. What are some general patterns you observe? Why do you think these patterns are the way they are?

- In 4000–3000 BP, is the lactase-persistence allele present in Central Asia? In Europe?

Biological Concepts

- Why did the lactase-persistence allele originally appear? *(Students may have the misconception that alleles arise only if and when people need them. If so, clarify that new alleles arise randomly through mutations.)*
- Describe the selective pressure that a lactase-persistence allele might have been under in *ancient* populations, considering the following factors:
    - What might have been some advantages of lactase persistence for ancient humans? In other words, what advantages did individuals who could digest lactose as adults have over individuals who could not?
    - When would lactase persistence *not* have been a significant advantage? In other words, when would the ability to digest lactose as an adult *not* increase a person's chance of survival? *(You may want to prompt students to think about whether all populations actually drank milk as adults and/or why they might not* need *to drink milk to survive.)*
- One hypothesis suggests that lactase persistence was strongly selected for in all traditionally pastoral populations that raised animals for milk — including populations in Africa, Europe, the Arabian Peninsula, and Central Asia. By similar logic, lactase persistence would *not* be strongly selected for in populations that were generally not pastoral, such as populations in Southeast Asia.
    - What parts of the figure support this hypothesis and why?
    - What parts of the figure contradict the hypothesis and why?
- Some cultures traditionally used milk to make other types of food products, such as yogurt or kefir, through fermentation. Fermentation is a process that chemically changes food using microorganisms like bacteria. The bacteria involved in fermentation make their own lactase.
    - Would fermented milk products, like yogurt or kefir, be easier for lactase-nonpersistent people to digest? Why or why not?
    - Would lactase persistence be advantageous in populations that consume fermented milk products (which contain much less lactose) instead of raw milk? Why or why not?
    - Based on the figure, which populations may have traditionally consumed fermented milk products?

Methods

- The "Present" map uses different colors to distinguish different sources of data.
    - What might be some advantages and disadvantages of compiling data from different sources?
    - Why do you think the scientists distinguish between data they collected (cyan) and data from other studies (green)?
    - Why do you think the scientists distinguish between genotype data (cyan and green) and phenotype data (magenta)?
- Phenotype data is collected using tests that measure a person's ability to digest lactose. How do you think tests like this might work? *(The "Blood Glucose Data Analysis" activity discusses this in more detail.)*
- Why are the maps with ancient humans (10000–4000 BP, 4000–3000 BP, and 3000–0 BP) based upon genotype data and not phenotype data?

Applications

- How do you think lactase persistence is distributed across the present-day United States, which has people with ancestry from many different parts of the world? Would you expect similar frequencies of lactase persistence across different cities, states, rural/urban areas, etc.? Why or why not?
- Describe the selective pressure that a lactase-persistence allele might be under in *present-day* populations, considering the following factors:
  - Food other than milk is readily available throughout much of the developed world.
  - People with lactase nonpersistence can now use lactase-containing supplements, like Lactaid, to break down lactose.
- The scientists who did this study (Ségurel et al. 2020) suggested that populations may adapt to milk consumption through "genetic adaptation" (e.g., natural selection for lactase-persistence alleles) or "cultural adaptation" (e.g., developing cultural practices for consuming milk that avoid the need for lactase).
  - What might be some other examples of genetic adaptations in humans?
  - What might be some other examples of cultural adaptations?

SOURCE

Figure 2 from:
Ségurel, Laure, Perle Guarino-Vignon, Nina Marchi, Sophie Lafosse, Romain Laurent, Céline Bon, Alexandre Fabre, et al. "Why and when was lactase persistence selected for? Insights from Central Asian herders and ancient DNA." *PLOS Biology* 18, 6 (2020): e3000742. https://doi.org/10.1371/journal.pbio.3000742. Used under CC BY 4.0 / Legends added and some colors adjusted for accessibility.

CREDITS

Written by Jean Weaver, The Prairie School, WI

Edited by Paul Beardsley, California State Polytechnic University, Pomona; Esther Shyu, HHMI

Scientific review by Laure Ségurel, CNRS-Université Claude Bernard Lyon 1, France