PATTERNS IN THE DISTRIBUTION OF LACTASE PERSISTENCE

INTRODUCTION

At some point after weaning, most humans around the world lose the ability to digest lactose, the main sugar found in milk. Lactose is broken down by the enzyme lactase, which is produced by cells lining the small intestine. Individuals that are **lactose intolerant**, or **lactase nonpersistent**, no longer produce this enzyme. However, in some parts of the world, most of the adult population continues to digest lactose. These people are called **lactose tolerant** or **lactase persistent**, because expression of the lactase gene *persists* beyond childhood.

In people that are lactose intolerant, the lactase gene gets "turned off" sometime after breastfeeding stops, whereas in people who are lactose tolerant the lactase gene is permanently "turned on." Scientists have discovered that the gene remains "on" due to mutations that are not in the lactase gene itself but in a control region near it.

Lactose tolerance (or lactase persistence) is found in so-called pastoralist populations. About 7,500 to 9,000 years ago, certain groups of people began domesticating cattle and drinking their milk. In such cultures, the lactase-persistence trait increased in frequency over time.

You will explore the geographic distribution of lactase persistence around the world by analyzing real data collected by scientists that will help you identify patterns.

MATERIALS

- Genetic Data Table
- Phenotype Data Table
- Pie Chart Stencils or circle stickers
- World Map (supplemental handout from instructor)
- Calculator (one per student)
- Two colored pencils
- Scissors
- Glue stick or tape
- Computer/references for geography research
- *Quick Guide: Measuring an Individual's Ability to Digest Lactose* (optional)

PROCEDURE PART 1: ANALYZING THE DATA TABLES

1. View the film *Got Lactase? The Co-evolution of Genes and Culture*, paying particular attention to the distributions and frequencies of people who are lactose intolerant (lactase nonpersistent) and lactose tolerant (lactase persistent) in various populations around the world.
2. Examine the Genetic Data Table. The genetic data were collected by sequencing DNA near the lactase gene. Scientists have identified at least four mutations near the lactase gene associated with the lactase-persistence trait. These mutations resulted in the formation of new alleles. The alleles that cause lactase persistence are called lactase-persistence alleles, and they act in a dominant fashion to the nonpersistence alleles. So a person only needs one copy of a lactase-persistence allele to show the trait. The number of people who had one or two copies of a lactase-persistence allele were counted and recorded in the column "# People with a lactase-persistence allele."
3. Calculate the frequency of lactase persistence in each human population sampled. Record the frequencies (expressed as a decimal and rounded to the hundredths place) in the column labeled "Lactase-persistence frequency." *(You may be asked to work in groups or as a class.)*
4. Examine the Phenotype Data Table. The phenotype data were collected using tests that measured an individual's ability to digest lactose. Some studies used a blood glucose test, while others used a hydrogen

breath test. To learn more about these two tests, read the optional handout *"Quick Guide: Measuring an Individual's Ability to Digest Lactose"* on page 11.

5. In the Phenotype Data Table, calculate the frequency of lactase persistence in each population. Record the frequencies in the column labeled "Lactase-persistence frequency." *(You may be asked to work in groups or as a class.)*

QUESTIONS PART 1: DATA ANALYSIS

6. Look at the Somali population in Ethiopia (A) in the **genetic** data set (page 8). Answer the following questions:
   a. How many people were tested in the study?
   b. How many people have at least one allele associated with the lactase-persistence trait?
   c. How many people did not have alleles associated with the lactase-persistence trait?
   d. Calculate the frequency of people in the Somali study who **did not** have one of the four alleles associated with the lactase-persistence trait. *(Show your work.)*

   e. Assume that the calculated frequency was an accurate representation of the entire Somali population. The population of Somalia is around 10 million. Approximately how many people in total would you expect to have at least one allele associated with the lactase-persistence trait? *(Show your work.)*

7. Using the **phenotype** data table (page 9), look at the Somali population (A) in Ethiopia. Answer the following questions:
   a. How many people were tested in the phenotype study?
   b. How many people tested positive for the ability to digest lactose?
   c. How many people were unable to digest lactose?
   d. Calculate the frequency of people in the Somali study who tested **negative** for the lactase-persistence phenotype. *(Show your work.)*

e. What did the scientists measure to collect the genetic and phenotype data? In other words, in what way do the two data sets differ?

8. From the data tables, record the lactase-persistence (LP) frequencies and references for the following population:

POPULATION	DATA TYPE	LP FREQUENCY	REFERENCES
Hungary (Hungarians) (X)	**Genetic data**		
	Phenotype data		

a. Write down two differences you notice about the data.

b. Provide two reasons why the frequencies of lactase persistence might be different when comparing genetic and phenotype data in the Hungarian population (X). *(Hint: Think of both the methods used to determine these frequencies and the biological differences between genotype and phenotype frequencies.)*

9. In a genetic study, 1876 individuals were sampled in Finland, whereas 58 people were sampled in France. (Finland's total population in 2008 was 5.3 million, while France's population in 2004 was 60.4 million.)

a. Which data set probably represents the country more accurately?

b. What two questions would you like answered about the samples to help you feel comfortable that they accurately represent the population of a particular country?

10. Comparing the genetic data to the phenotype data, identify a pattern you observe in the two data sets for the three populations from Sudan (F, G, H).

11. Using the genetic information provided, what do the lactase-persistence frequencies from the populations from Senegal (Wolof - D), Sudan (Dinka - G), and Japan (Japanese - N) have in common?

12. The lactase-persistence **phenotype** frequency data from these three populations range from 25% to 51%. Based on your knowledge of molecular genetics (DNA, genes, gene expression), what might be a reason for this difference?

13. Based on genetic **and** phenotype data, develop a hypothesis about the cultural practices regarding domesticating livestock and consuming milk of the populations sampled in Uganda (I), China (P), and Papua New Guinea (R).

14. The Dinka people of Sudan (G) are known as *agropastoralists*. They depend on agriculture during the rainy season and livestock herding and domestication during the dry season. According to the study, the Dinka people sampled had a 0% frequency of lactase persistence when looking at the genetic data, which is not consistent with their agropastoralist culture. What might be a limitation to this particular research study?

PROCEDURE PART 2: MAPPING THE ALLELE FREQUENCIES

To map the frequencies for lactase persistence and nonpersistence, create a mini pie chart for each population. Notice that the first column of the **Genetic** Data Table has letters denoting each population. The pie chart stencils on page 10 also include letters in the center of each circle. Each of the pie chart stencils corresponds to the lettered population in the **Genetic** Data Table.

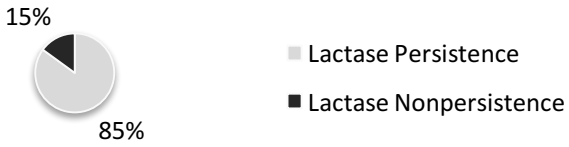
Note: The exact procedure will vary from class to class. For example, your instructor may have you work on only a few of the pie charts or all of them, working in small groups or as a class.

15. Choose two different color pencils. Designate one color for lactase persistence and the other for lactase nonpersistence.

Frequency	Color
Lactase persistence	
Lactase nonpersistence	

16. Using the pie chart stencils and colored pencils, color the appropriate pie chart to represent the lactase persistence and nonpersistence frequency for each population from the **Genetic** Data Table. *(Be sure to match the letter of each population to the letter on the pie chart stencil.)*

Example (Ex) Pie Chart



17. Cut out all the colored pie charts. *(They don't need to look pretty.)*
18. Use a reliable reference (i.e., a website, social studies textbook, or a reference provided by your teacher) to place the pie chart(s) in the appropriate locations on the provided world map. *(Do **NOT** tape or glue.)*
19. Once all the pie charts are placed on the map, space them so that you can see each one including the letters. Neatly tape or glue them to the world map. Be sure to add a title and key to your map.

QUESTIONS PART 2: MAP ANALYSIS

20. Looking at the global distribution of pie charts on your map, write three claims that you could make about the worldwide distribution of lactase persistence. Note the evidence that supports your claim and alternative explanations or data that do not support your claim. (For example, in Europe, you could make a claim about how the pattern of lactase-persistence/nonpersistence frequencies relates to latitude.)

    Claim 1:

    Claim 2:

    Claim 3:

21. What do you notice about the lactase-persistence frequencies of the Maasai people in Kenya (B) and the Sengwer people, also in Kenya (C)?

    a. Based on the information presented in the film, what could have accounted for this difference? Explain your answer.

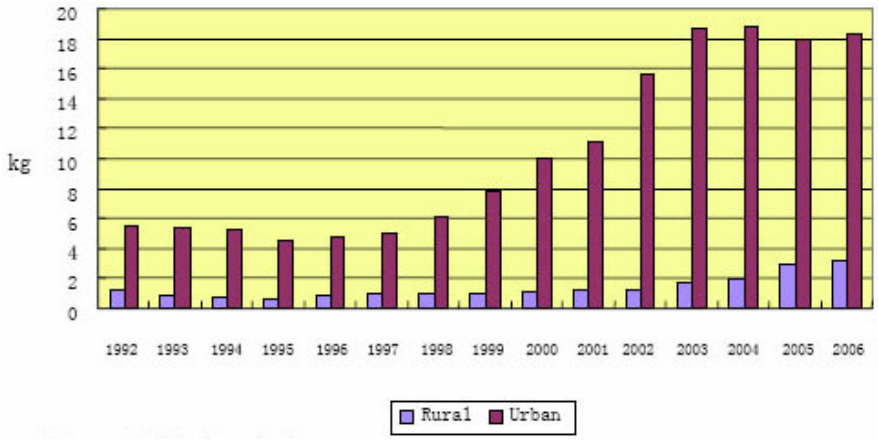
22. Look at the genetic data for the two populations from England (S, T) and the two populations from Italy (Y, Z). What do you notice when you compare S to T and Y to Z within each country? What question or questions does this comparison raise for you?

23. What do you notice when analyzing the genetic data for Northern Europe and those for East Asia/Australasia? What reason could there be for the difference you observe?

24. Find the allele frequency pie chart in China for the Han people (P). The Han people make up 92% of mainland China's population and roughly one-fifth of the world's population. Of the 200 people sampled, what percentage was predicted to be lactose intolerant?

A study published in 2009 shows an increase in milk consumption among the Chinese people. See the graph:

Per capita consumption of dairy products in rural and urban households



Source: Chinese statistical yearbook

25. Using the graph on the previous page, predict how the genetic data for lactase-persistence frequency might change if the same population is sampled again in 1000 years. Explain your prediction, noting what (if any) selective pressures would be present.

26. The main food sources of the Bantu people of Uganda (I) are matoke (a fruit in the banana family), eggs, fish, beans, nuts, beef, chicken, goats, and various fruits and vegetables. Explain how the data presented in this activity either support or do not support what you now know about their cultural diet.

27. In a few sentences, support the claim that lactase persistence is a good example of gene/culture co-evolution. Your explanation should include several pieces of evidence.

Genetic Data

Data indicate whether a person in the sample has at least one copy of the lactase-persistence allele.

	Continent/ Region	Country	Population	# People sampled (*n*)	# People with lactase-persistence allele	Lactase-persistence frequency	Reference
A	Africa	Ethiopia	Somali	74	22		Ingram *et al.* (2009) *Hum. Gen.* **124**, 579.
B	Africa	Kenya	Maasai	64	54		Tishkoff *et al.* (2007) *Nat. Genet.* **39**, 31.
C	Africa	Kenya	Sengwer	32	4		Tishkoff *et al.* (2007) *Nat. Genet.* **39**, 31.
D	Africa	Senegal	Wolof	118	0		Ingram *et al.* (2009) *Hum. Gen.* **124**, 579.
E	Africa	South Africa	Xhosa	109	27		Torniainen *et al.* (2009) *BMC Genet.* **10**, 31.
F	Africa	Sudan	Beni Amer	162	73		Ingram *et al.* (2009) *Hum. Gen.* **124**, 579.
G	Africa	Sudan	Dinka	18	0		Tishkoff *et al.* (2007) *Nat. Genet.* **39**, 31.
H	Africa	Sudan	Jaali	172	46		Ingram *et al.* (2009) *Hum. Gen.* **124**, 579.
I	Africa	Uganda	Bantu	44	0		Mulcare *et al.* (2004) *Am. J. Hum. Genet.* **74**, 1102.
J	Africa	Tanzania	Burunge	36	22		Tishkoff *et al.* (2007) *Nat. Genet.* **39**, 31.
K	Africa	Tanzania	Maasai	38	26		Tishkoff *et al.* (2007) *Nat. Genet.* **39**, 31.
L	Asia	Afghanistan	Tadjik	98	19		Mulcare (2006) London: University of London PhD.
M	Asia	India	Indian	68	17		Mulcare (2006) London: University of London PhD.
N	Asia	Japan	Japanese	62	0		Bersaglieri *et al.* (2004) *Am. J. Hum. Genet.* **74**, 1111.
O	Asia	Russia	Udmurt	60	33		Enattah *et al.* (2008) *Am. J. Hum. Genet.* **82**, 57.
P	Asia	China	Han	200	0		Enattah *et al.* (2008) *Am. J. Hum. Genet.* **82**, 57.
Q	Asia	China	Mongol	82	8		Sun *et al.* (2007) *Asia Pac. J. Clin. Nutr.* **16**, 4.
R	Australasia	Papua New Guinea	Papuan	34	0		Bersaglieri *et al.* (2004) *Am. J. Hum. Genet.* **74**, 1111.
S	Europe	England	English, northern	1168	1098		Davey Smith *et al.* (2009) *Eur. J. Human Gen.*, **17**, 357-367.
T	Europe	England	English, southeastern	947	862		Davey Smith *et al.* (2009) *Eur. J. Human Gen.*, **17**, 357-367.
U	Europe	Finland	Finn	1876	1538		Enattah *et al.* (2008) *Am. J. Hum. Genet.* **82**, 57.
V	Europe	France	French	58	39		Bersaglieri *et al.* (2004) *Am. J. Hum. Genet.* **74**, 1111.
W	Europe	Greece	Greek	100	17		Anagnostou *et al.* (2009) *Am. J. Hum. Biol.* **21**, 217.
X	Europe	Hungary	Hungarian	110	95		Nagy *et al.* (2009) *Eur. J. Clin. Nutr.* **63**, 909.
Y	Europe	Italy	Northern Italian	28	17		Bersaglieri *et al.* (2004) *Am. J. Hum. Genet.* **74**, 1111.
Z	Europe	Italy	Sardinian	153	21		Anagnostou *et al.* (2009) *Am. J. Hum. Biol.* **21**, 217.
AA	Near/Middle East	Jordan	Jordanian	112	22		Enattah *et al.* (2008) *Am. J. Hum. Genet.* **82**, 57.
BB	Near/Middle East	Turkey	Anatolian Turk	98	6		Mulcare (2006) London: University of London PhD.
CC	Near/Middle East	Saudi Arabia	Bedouin	94	69		Ingram *et al.* (2009) *Hum. Gen.* **124**, 579.
DD	Near/Middle East	Saudi Arabia	Arab	248	206		Enattah *et al.* (2008) *Am. J. Hum. Genet.* **82**, 57.

Source: Global Lactase Persistence Association Database, http://www.ucl.ac.uk/mace-lab/resources/glad

Phenotype Data

	Continent/ Region	Country	Population	# people sampled (*n*)	# people tested positive for lactase persistence	Lactase-persistence frequency	Reference
A	Africa	Ethiopia	Somali	90	22		Ingram *et al*. (2009) *Hum. Gen.* ***124,*** 579.
B	Africa	Kenya	Maasai	26	23		Tishkoff *et al*. (2007) *Nat. Genet*. **39**, 31.
C	Africa	Kenya	Sengwer	12	2		Tishkoff *et al*. (2007) *Nat. Genet*. **39**, 31.
D	Africa	Senegal	Wolof	53	27		Arnold *et al.* (1980) *C. R. Seances Soc. Biol. Fil*. **174**, 983.
E	Africa	South Africa	Xhosa	17	3		Segal *et al*. (1983) *Am. J. Clin. Nutr*. **38**, 901.
F	Africa	Sudan	Beni Amer	40	35		Bayoumi *et al.* (1982) *Am. J. Phys. Anthropol*. **58**, 173.
G	Africa	Sudan	Dinka	208	52		Bayoumi *et al.* (1982) *Am. J. Phys. Anthropol*. **58**, 173.
H	Africa	Sudan	Jaali	113	60		Bayoumi *et al.* (1981) *Hum. Genet*. **57**, 279.
I	Africa	Uganda	Bantu	17	1		Cook *et al.* (1966) *Lancet* **1**, 725.
J	Africa	Tanzania	Burunge	16	6		Tishkoff *et al*. (2007) *Nat. Genet.* **39**, 31.
K	Africa	Tanzania	Maasai	15	10		Tishkoff *et al*. (2007) *Nat. Genet.* **39**, 31.
L	Asia	Afghanistan	Tadjik	79	14		Rahimi *et al*. (1976) *Hum. Genet*. **34**, 57.
M	Asia	India	Indian	100	36		Desai *et al.* (1970) *Indian J. Med. Sci*. **24**, 729.
N	Asia	Japan	Japanese	40	11		Yoshida *et al*. (1975) *Gastroenterol. Jpn*. **10**, 29.
O	Asia	Russia	Udmurt	30	18		Kozlov (1998) *Int. J. Circumpolar Health* **57**, 18.
P	Asia	China	Han	248	20		Yongfa *et al*. (1984) *Hum. Genet*. **67**, 103.
Q	Asia	China	Mongol	198	24		Yongfa *et al*. (1984) *Hum. Genet*. **67**, 103.
R	Australasia	Papua New Guinea	Papuan	30	3		Jenkins *et al*. (1981) *Ann. Hum. Biol*. **8**, 447.
*	Europe	England	British	150	143		Ferguson *et al.* (1984) *Gut* **25**, 163.
U	Europe	Finland	Finn	638	530		Jussila (1969) *Ann. Clin. Res*. **1**, 199.
V	Europe	France	French	102	78		Cloarec *et al.* (1991) *Gastroenterol. Clin. Biol*. **15**, 588.
W	Europe	Greece	Greek	600	330		Kanaghinis *et al*. (1974) *Am. J. Dig. Dis*. **19**, 1021.
X	Europe	Hungary	Hungarian	535	337		Czeizel *et al.* (1983) *Hum. Genet*. **64**, 398.
Y	Europe	Italy	Northern Italian	208	102		Burgio *et al.* (1984) *Am. J. Clin. Nutr*. **39**, 100.
Z	Europe	Italy	Sardinian	53	6		Meloni *et al*. (1998) *Ital. J. Gastroenterol. Hepatol.* **30**, 490.
AA	Near/Middle East	Jordan	Jordanian	148	37		Hijazi *et al*. (1983) *Trop. Geogr. Med*. **35**, 157.
BB	Near/Middle East	Turkey	Anatolian Turk	122	32		Flatz *et al.* (1986) *Am. J. Hum. Genet*. **38**, 515.
CC	Near/Middle East	Saudi Arabia	Bedouin	21	17		Dissanayake *et al.* (1990) *Ann. Saudi Med*. **10**, 598.
DD	Near/Middle East	Saudi Arabia	Arab	109	47		Dissanayake *et al.* (1990) *Ann. Saudi Med*. **10**, 598.

Source: Global Lactase Persistence Association Database, http://www.ucl.ac.uk/mace-lab/resources/glad

Pie Chart Stencils

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y

Z

AA

BB

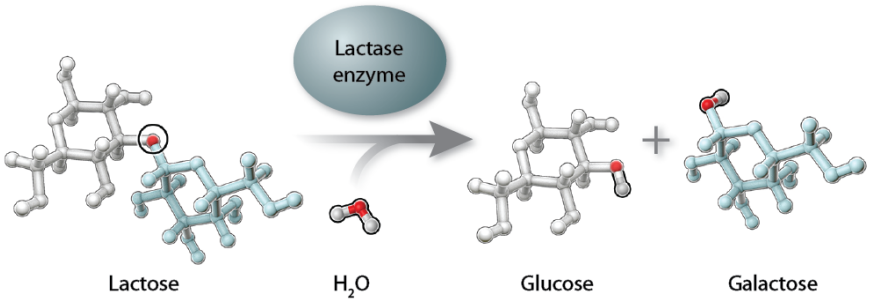
CC

DD

QUICK GUIDE: MEASURING AN INDIVIDUAL'S ABILITY TO DIGEST LACTOSE

"Mother's milk" is packed with the proteins, fats, and carbohydrates that support the growth, development, and survival of baby mammals. The sugar lactose is the main carbohydrate in milk. Lactose can be cleaved into two simpler sugars, glucose and galactose, by lactase, an enzyme produced in the small intestine. The two smaller sugars are readily absorbed though the intestinal wall into the bloodstream for delivery to the cells of the body, where they are used for energy.

After infant mammals are weaned from their mother's milk, lactase production shuts down, presumably because it is no longer needed. This condition is called lactase nonpersistence—meaning that production of the lactase enzyme does not *persist* into adulthood. The general condition for mammals is not to consume milk after weaning and to be lactase nonpersistent. Some populations of humans are unusual in that adults continue to consume milk from other mammals, such as cows.



If a person who is lactase nonpersistent drinks milk, undigested lactose passes from the small intestine to the large intestine, where it is fermented by bacteria. Fermentation produces various gases in the large intestine, which can cause abdominal pain, bloating, flatulence, and diarrhea—all symptoms of lactose intolerance. Worldwide, most adults are lactose intolerant, although some people may not know because their symptoms are mild. Only a minority of human adults (35% of the global human population) continues to produce lactase into adulthood and can drink milk without any problems. These individuals are said to be lactase persistent or lactose tolerant.

Testing Methods

One way to test whether a person is lactase persistent is to measure their blood glucose levels. In the short film, *Got Lactase? The Co-evolution of Genes and Culture*, the narrator, Dr. Spencer Wells, takes a **blood glucose test** to deduce his lactase status. If the lactase enzyme is present, blood glucose levels increase within 20 to 60 minutes of drinking milk.

Another common test used to determine whether a person is lactase persistent is the **hydrogen breath test**. This test measures the amount of hydrogen in a person's breath. Undigested lactose is fermented by bacteria in the large intestine and produces several gases, including hydrogen. These gases exit the body through the anus; they can also be absorbed into the blood, circulated to the lungs, and eliminated through the breath. If a lactase nonpersistent person consumes lactose, the amount of hydrogen in their breath will go up whereas the amount of hydrogen in the breath of someone who is lactase persistent will stay the same.