Exploring Lactase Persistence: Interrogating the *MCM6* Gene to Identify Genetic Variants Associated with Lactase Persistence in Geographically Diverse Populations

Brian Phan, Morgan Pugh de los Reyes, Michael C. Campbell, Ph.D.

Dept of Biological Sciences, Human and Evolutionary Biology section, University of Southern California, Los Angeles, CA, USA

Bridge **U**nder**G**rad **S**cience **(BUGS)** Summer Research Program

USC University of Southern California

Background

Lactase persistence—the ability of adults to digest the lactose in milk—is a genetically-determined trait that varies widely in frequency across the globe (**Figure 1**). In humans, the capacity to digest lactose—the sugar present in milk—declines after weaning. However, in a number of ethnic groups from Europe, the Middle East, and Africa, the ability to digest lactose persists into adulthood; this condition is known as lactase-persistence (LP)(Ranciaro et al. 2014).

To date, several common polymorphisms associated with LP have been identified in human populations. For example, the T_{-13910} allele in intron 13 of the *MCM6* gene contributes to LP in northern European populations with a tradition of fresh milk production and dairy consumption. Although studies have shown that the T_{-13910} is a key variant regulating LP in Europe, this allele is absent in most African and Middle Eastern populations that practice pastoralism and consume high quantities of milk (Ranciaro et al. 2014).

Recently, several studies have identified additional variants (namely, C_{-14010}, G_{-14009}, G_{-13907}, and G_{-13915} in intron 13 of *MCM6*) that are associated with the LP trait in eastern African or Middle Eastern pastoralist populations (Ranciaro et al. 2014). Thus, different populations from Africa and the Middle East have their own distinct variants that contribute to LP.

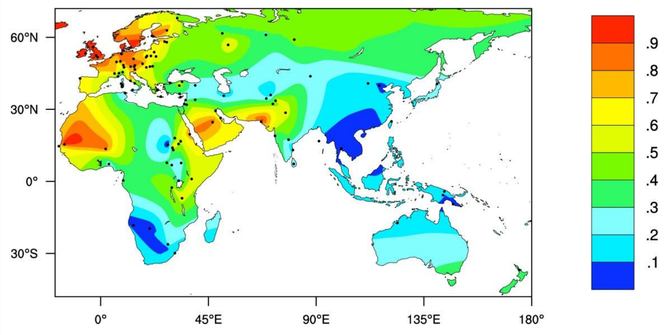


Figure 1: Distribution of lactase persistence trait across the globe. Colors and color key show the frequencies of the LP phenotype (Itan et al. 2010). LP is found moderate to high frequencies in Europeans and some African, and Middle Eastern populations.

Overview of Project

- To perform PCR of a 600–base pair region of the *MCM6* gene (containing the variants known to be associated with LP) in 29 individuals from Africa, the Middle East, and Taiwan.
- To perform gel electrophoresis of PCR products to determine if the PCR step was successful..
- To sequence 600–base pair region of the *MCM6* gene.
- To identify genetic variants associated with LP in African, Middle Eastern, and Taiwanese populations.

Materials and Methods

Figure 2: PCR and gel electrophoresis procedures. PCR was performed successfully in 28 out of 29 samples.

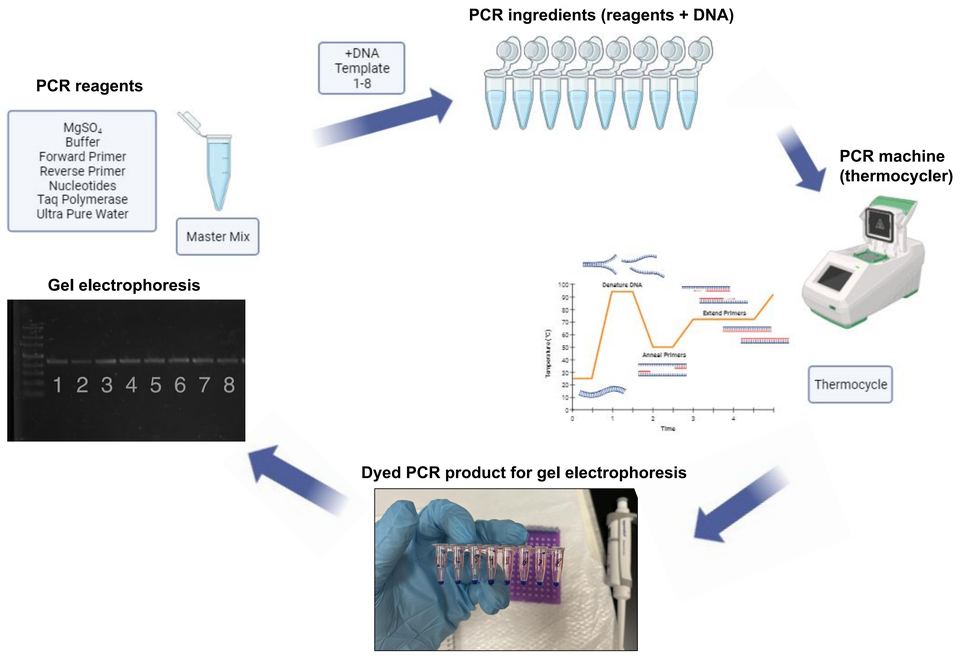


Figure 3: NCBI Downloaded Reference Sequence. Nucleotides highlighted in red are the PCR primers. Nucleotides highlighted in yellow are the sequencing primers. Individual nucleotides highlighted in grey, purple, green, red, and blue are the variants associated with LP (namely, C/G_{-14010}, G/T_{-14009}, G/T_{-13915}, C/T_{-13910}, and G/C_{-13907}, respectively).

TCCAAAGAGTCAGAGGACTTCATTGTGGAGCAATATAAACATCTCCGCCAGAGAGATGGTTCTG
GAGTGACCAAGTCTTCATGGAGGATTACAGTGCGACAGCTTGAGAGCATGATTCGTCTCTCTGAA
GCTATGGCTCGGATGCACTGCTGTGATGAGGTATCAGAGTCACTTTGATATGATGAGAGCAGAG
ATAAACAGATTTGTTGCATGTTTTTAATCTTTGGTATGGGACATACTAGAATTCACTGCAAATACA
TTTTTATGTAACTGTTGAATGCTCATACGACCATGGAATTCTTCCCTTTAAAGAGCTTGGTAAGCA
TTTGAGTGTAGTTGTTAGACGGAGACGATCACGTCATAGTTTATAGAGTGCATAAAGACGTAAGT
TACCATTTAATACCTTTCATTCAGGAAAAATGTACTTAGACCCTACAATGTACTAGTAGGCCTCTG
CGCTGGCAATACAGATAAGATAATGTAGCCCTGGCCTCAAAGGAACTCTCCTCCTTAGGTTGCA
TTTGTATAATGTTTGATTTTTAGATTGTTCTTTGAGCCCTGCATTCCACGAGGATAGGTCAGTGGG
TATTAACGAGGTAAAAGGGGAGTAGTACGAAAGGGCATTCAAGCGTCCCATCTTCGCTTCAACC
AAAGCAGCCCTGCTTTTTCCTAGTTTTATTAATAGGTTTGATGTAAGGTCGTCTTTGAAAAGGGG
GTTTGGCTTTTTTTTACAGTGTGACTGAGGTATAATTTATAAAAAGGGAAATGTATGGCATGGTG
AGTTTTTTCACATACATCCTTGTGAATACCCAGCTCAAGATCCAAAACATTTCCATAATTTCAGAAA

Figure 4: Example of Sequencher chromatogram. Double peaks in chromatogram indicate two alleles present in an individual (heterozygote).

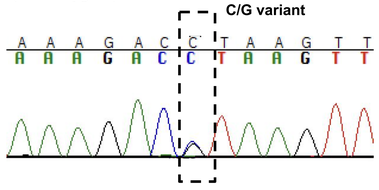


Figure 5: Example of Sequencher chromatogram. No double peaks in chromatogram indicate homozygous alleles in **individual A** and **individual B**.

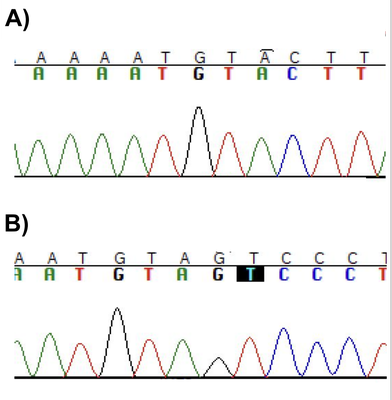
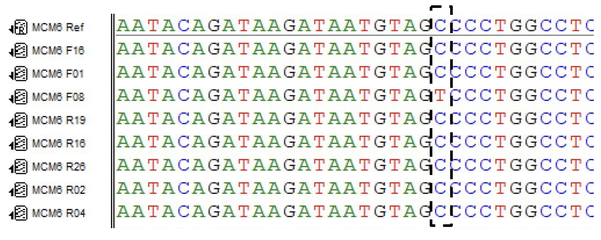
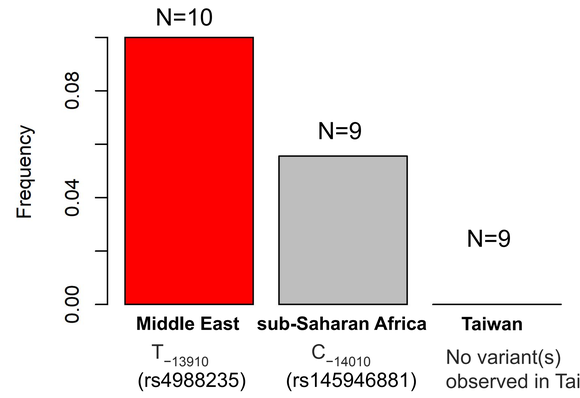


Figure 6: **Aligned DNA sequences from multiple individuals**. An example of a single nucleotide difference spotted (in the dashed box) in the 600-base pair region of the *MCM6* gene when multiple DNA sequences are compared.



Results and Conclusions

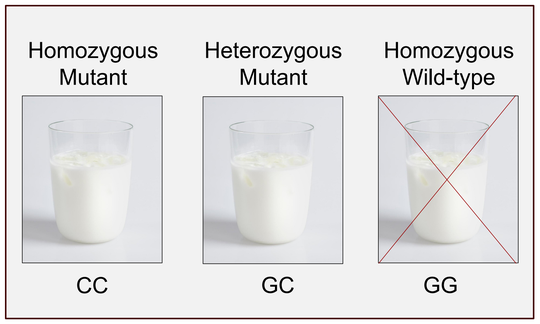
Figure 7: Frequency of LP-associated alleles. Here, we show the different frequencies of LP-associated alleles present in human populations in distinct geographic regions.



1. The T_{-13910} allele was found in a Middle Eastern individual as a homozygous mutation. The presence of the T_{-13910} allele, which is commonly found Europe, suggests there may have gene flow from Europe to the Middle East.
2. The C_{-14010} allele was found in a sub-Saharan African individual in heterozygous form (*i.e.*, both the C and G alleles were present in this individual). Interestingly, this C_{-14010} is commonly found in East Africa.
3. Aboriginal Taiwanese samples had no LP-associated mutations as expected.

Summary

Figure 8: Summary explanation of the G/C-14010 SNP. Different genotypes are associated with different phenotypic outcomes (*i.e.*, the ability to digest milk).



References

1. Itan, Y. et al. (2010) "A Worldwide Correlation of Lactase Persistence Phenotype and Genotypes." *BMC Evolutionary Biology* 10 (1): 36.
2. *Scientific Image and Illustration Software | BioRender*. https://www.biorender.com/.
3. Ranciaro, A. et al. (2014) "Genetic Origins of Lactase Persistence and the Spread of Pastoralism. *American Journal of Human Genetics* 94(4): 496-510.

Contact Us

bridge.usc.edu/bugs