Mediterranean Journal of Nutrition and Metabolism 8 (2015) 61–66
DOI:10.3233/MNM-150030
IOS Press

Screening for lactase persistence associated genotypes in cattle domestication groups of northern India

Raja A.H. Kuchay[a,b], Babu R. Thapa[c], Akhtar Mahmood[b] and Safrun Mahmood[a,*]

[a]*Department of Experimental Medicine and Biotechnology, Postgraduate Institute of Medical Education and Research, Chandigarh, India*

[b]*Department of Biochemistry, Panjab University, Chandigarh, India*

[c]*Division of Pediatric Gastroenterology, Hepatology and Nutrition, Postgraduate Institute of Medical Education and Research, Chandigarh, India*

Abstract. The ability to digest lactose is a variable genetic trait in human populations. Two distinct phenotypes are present: lactase non-persistence (LNP) and lactase persistence (LP). LP trait is likely to have conferred a selective advantage in individuals who consume appreciable amounts of milk. In European derived populations a single nucleotide polymorphism (SNP) C/T_{-13910} residing 13.9 kb upstream from the lactase gene has been shown to define lactase activity, and several other SNPs in the same region have been identified in African and Middle East populations. We evaluated the frequency of C/T_{-13910} SNP among cattle domesticating groups of northern India. It was observed that frequency of LP phenotype and that of T_{-13910} allele of LP was much higher in these groups compared to those with no history of cattle domestication. This may be to due positive selection of LP associated genotypes with spread of cattle domestication in northern India. However, the introduction of T_{-13910} allele into India seems unclear and may require further research.

Keywords: Lactase persistence, domestication, northern India

1. Introduction

Lactose in milk is digested and absorbed in the small intestine. The enzyme lactase is responsible for cleaving lactose into its constituent monosaccharides, glucose and galactose, which are transported across the epithelial cell membranes into the enterocytes [1]. Lactase activity is high and vital during infancy, when milk is the main source of nutrition. In most mammals, however, lactase activity declines after the weaning, a condition known as lactase non-persistence (LNP). Lactase persistence (LP) is an autosomal dominant trait enabling the continued production of the enzyme lactase throughout adult life. People who have LP can usually hydrolyze large amounts of lactose and can thus consume large quantities of fresh milk without complication. People with LNP have a much lower lactose digestive capacity than those with LP, and thus often, but not always, show symptoms of lactose intolerance after consumption of fresh milk. The distribution of these different lactase phenotypes in human populations is highly variable, an observation that has long been a source of interest in relation to evolutionary genetics [1].

The LP phenotype is very frequent in the northern parts of Europe, and also in nomadic populations in the arid zones of north and central Africa and Arabia [2]. There is a decline of LP toward southern parts of the European continent [1]. Frequency of T_{-13910} allele ranges from 0.080 in southern Italy to 0.237 in north-east Italy [3]. In Sardinia and Greece, the frequency of T_{-13910} allele ranges from 0.072 and 0.090 [3]. LNP is the predominant phenotype in the native populations of Australia and America and in the Pacific, east and southeast Asia and tropical Africa [2].

*Corresponding author: Dr. Safrun Mahmood, Geneticist, Department of Experimental Medicine and Biotechnology, Postgraduate Institute of Medical Education and Research, Chandigarh 160012, India. Tel.: +91 172 2755237; Fax: +91 172 2744401; E-mail: mahmoodpgi@gmail.com.

ISSN 1973-798X/15/$35.00

World-wide distribution of LP and the generally coinciding pattern of historically milk-drinking populations led to '*culture historical hypothesis*'. According to this hypothesis, LP co-evolved alongside the cultural adaptation of milk and the selective force for LP was milk dependence [4].

A single nucleotide polymorphism (SNP) C/T_{-13910} residing in intron 13 of the *MCM6* gene 14 kb upstream from the lactase gene (*LCT*) has been found to correlate with lactase activity [5]. It has been revealed that a single allele, carrying T_{-13910} variant 14 kb upstream *LCT*, fully correlates with LP in many global populations [5–7]. Functional studies have shown that the T_{-13910} variant acts as an enhancer and it regulates the *LCT* gene at the transcriptional level [8]. It has been suggested that the patterns of variation in and surrounding the lactase gene show a strong signature of recent positive selection roughly consistent with the likely time of onset of dairy farming [6]. We evaluated the frequency of T_{-13910} allele in two cattle domestication groups of northern India (from Panjab and Jammu and Kashmir regions) to determine the relationship between C/T_{-13910} SNP and the LP phenotype in these groups. We also determined the frequency of T_{-13910} allele in a mixed population of northern India without any known history of cattle domestication or pastoralism.

2. Methods

2.1. Study population

The study population included 100 adult volunteers from cattle domestication groups of Panjab and Jammu and Kashmir (J&K) (50 each). The volunteers were recruited from the population group with a history of cattle domestication and milk drinking culture. Furthermore, 150 adult volunteers were recruited from a mixed population of northern India (50 each from Panjab and J&K; 25 each from Haryana and Himachal Pradesh) with no history of cattle domestication or pastoralism. Additionally, 25 volunteers from southern India (studying at PGIMER, Chandigarh) were also recruited. At the time of recruitment, the volunteers were asked to complete a questionnaire concerned with milk consumption and possible milk related symptoms. Written informed consent to participate in the study was obtained from these volunteers. Blood samples were collected from all the volunteers who participated in this study. The study was approved by Institute Ethics Committee PGIMER, Chandigarh.

2.2. DNA isolation and genotyping

DNA was isolated from whole blood samples using REDExtract-N-Amp Blood PCR Kit (Sigma, Saint Louis, USA). C/T_{-13910} SNP was analysed through polymerase chain reaction (PCR) amplification followed by restriction enzyme digestion assay (PCR-RFLP). PCR was carried out using REDExtract-N-Amp Blood/Saliva PCR Kit (Sigma, Saint Louis, USA). Primers used to amplify the fragment spanning C/T_{-13910} variant were: sense 5′-gga tgc act gct gtg atg ag-3′ and antisense 5′-ccc act gac cta tcc tcg tg-3′. Each reaction was denatured at 95°C for 3 min, followed by 35 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 30 s. The reaction was given a final 10 min extension at 72°C. PCR products were quantified and 500 ng of DNA was digested with 1 unit of *Bsm*FI restriction enzyme (New England Biolabs, Foster City, CA) SNP in a 15μL reaction containing 1X reaction buffer and BSA. The reaction were incubated for 4-5 hours at 65ºC. A total of 7μL of reaction mixtures were electrophoresed using 2% agarose gel and stained with ethidium bromide. To confirm the PCR–RFLP results and discover other possible SNPs, PCR products were sequenced by using an automated DNA Sequencer (ABI 3730XL Genetic analyzer; Xcelris Genomics) with forward primer to read ∼400 base pairs in one direction. When necessary, the results were reconfirmed by sequencing the other strand with the reverse primer. PCR products were first purified using the PCR gel purification kit (Qiagen), as per the manufacture's procedure and later on send for sequencing (Xcelris labs, India).

3. Results

Amplification of the fragment spanning C/T_{-13910} SNP resulted in a 448 bp PCR product. For C/T_{-13910} SNP, digestion of the PCR product with restriction enzyme *Bsm*FI revealed two fragments of ≈350bp and ≈100 bp in

case of the *C/C*$_{-13910}$ genotype, and of ≈350bp, ≈250bp and ≈100/98bp for *C/T*$_{-13910}$ genotype (Fig.1). Among the volunteers from cattle domestication group of Panjab, 56% (28/50) had the *C/T*$_{-13910}$ genotype of LP and 44% (22/50) had the *C/C*$_{-13910}$ genotype of LNP. None of the individuals had *T/T*$_{-13910}$ genotype. From the cattle domestication group of J&K, 64% (32/50) had the *C/T*$_{-13910}$ genotype of LP and 36% (18/50) had the *C/C*$_{-13910}$ genotype of LNP. Among the volunteers from non-cattle domestication groups, the frequency of *C/T*$_{-13910}$ genotype in volunteers from Panjab, J&K, Haryana and Himachal Pradesh was 38% (19/50), 42% (21/50), 48% (12/25) and 36% (9/25) respectively. The frequency of *C/C*$_{-13910}$ genotype in this group was 62% (31/50), 58% (29/50), 52% (13/25) and 64% (16/25) for volunteers of Panjab, J&K, Haryana and Himachal Pradesh respectively. Among the volunteers of southern India, only 12% (3/25) had *C/T*$_{-13910}$ genotype of LP and 88% (22/25) had *C/C*$_{-13910}$ genotype of LNP (Table 1).

Thus, a gradient in the frequency of *T*$_{-13910}$ allele was observed from northern to southern India. Northern India has the highest frequency of *T*$_{-13910}$ allele and its frequency decreases as we move towards southern regions of the country. In northern India, the cattle domestication group of J&K had the highest frequency of *T*$_{-13910}$ allele (64%). This group mainly consisted of the *Gujjars* and *Bakarwals*, a prominent heading community in J&K. Highest frequency of *T*$_{-13910}$ in this group may suggest a positive selection of this allele alongwith the pastoral culture in the *Gujjar* and *Bakarwals* community. The frequency of *C/T*_13910 genotype in Panjab and J&K (cattle domestication+non-

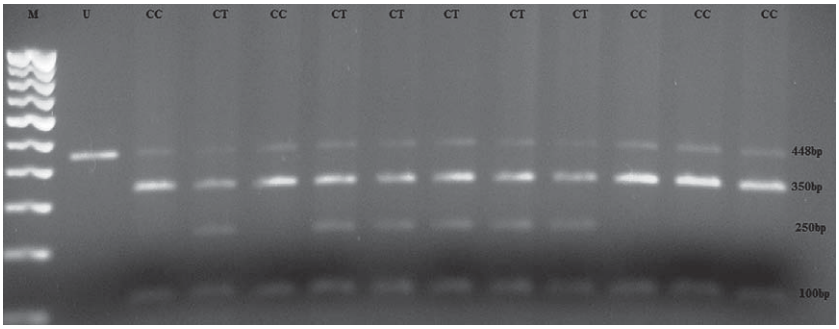

Fig. 1. Agarose gel electrophoresis for PCR-RFLP analysis of C/T -13910 SNP from blood. Genotyping of individuals for the C/T -13910 SNP was carried out by restriction digestion with of *Bsm*F1. Lane M, DNA molecular weight marker. Lane U, Undigested PCR product. C/C and C/T represent individuals with respective genotypes at -13910 positions.

Table 1
Prevalence of various genotypes of C/T -13910 SNP in volunteers from cattle domestication and non-domestication groups from different regions of north India

Group (N)	Genotype		%LP[1]
	(LNP Associated) C/C -13910	(LP Associated) C/T-13910	
Domestication			
J&K (50)	18	32	64
Panjab (50)	22	28	56
Non-Domestication			
J&K (50)	29	21	42
Panjab (50)	31	19	38
Haryana (25)	13	12	48
Himachal (25)	16	9	36
South India (25)	22	3	12

[1]Based on C/T$_{-13910}$ genotype.

Table 2
Self–reported gastrointestinal symptoms of volunteers in cattle domestication and non-domestication groups after milk consumption

Gastrointestinal symptoms[1]	Genotype		Total
	C/C_{-13910} n/N (%)	C/T_{-13910} n/N (%)	n/N (%)
Domestication			
Flatulence	13/40 (32.5)	7/60 (11.6)	20/100 (18)
Bloating	4/40 (10)	2/60 (3.3)	6/100 (6)
Diarrhea	5/40 (12.5)	4/60 (6.6)	9/100 (9)
Borborgymi	7/40 (17.5)	5/60 (8.3)	12/100 (12)
Constipation	3/40 (7.5)	4/60 (6.6)	7/100 (7)
Non-Domestication			
Flatulence	74/111 (66.6)	21/64 (32.8)	95/175 (54.2)
Bloating	35/111 (31.5)	14/64 (21.8)	49/175 (28)
Diarrhea	21/111 (18.9)	8/64 (12.5)	29/175 (16.5)
Borborgymi	29/111 (26.1)	14/64 (21.8)	43/175 (24.5)
Constipation	12/111 (10.8)	7/64 (10.9)	19/175 (10.8)

[1]Self-reported gastrointestinal symptoms after consumption of two glasses of milk per day (∼500 mL).

domestication group) was 41% and 51% respectively. Overall frequency of C/T_{-13910} genotype in northern India (combining the data from all four states and including the data from cattle domestication group) was 59.5%. This frequency is much higher than 12% as was reported for south Indian group.

We examined the correlation of C/T_{-13910} SNPs with self-reported clinical symptoms after consumption of at least two glasses of milk per day (approximately 500 mL) in all the volunteers. Milk related clinical symptoms were much less frequent in volunteers from cattle domestication group than the non-domestication group (Table 2). The symptoms were least frequent in the volunteers from cattle domestication group with C/T_{-13910} genotype. This again suggests a positive selection of T_{-13910} allele alongwith the pastoral and milk drinking culture.

4. Discussion

LP is a subject of great interest because the genomic region including *LCT* gene, which encodes the enzyme lactase, is largely considered one of the strongest examples of positive selection pressure [6]. It also represents a good example of genetic-environment co-evolution, the genetic trait seems to have increased in frequency starting from the Neolithic transitions, together with the diffusion of dairy farming and the habit of drinking milk in adulthood, that gives nutritional advantages [9]. Variation in linked microsatellites [7], an absence of the T_{-13910} allele in a sample of early Neolithic central European skeletons [10] and spatially explicit computer simulations [11] have all added support to the case for a strong selective advantage. We report a high frequency of T_{-13910} allele in pastoral groups of northern India. This may be attributed to the positive selection for LP genotypes alongwith the spread of cattle domestication and pastoralism in these groups. The data from self reported gastrointestinal symptoms after consumption of two glasses of milk also seems to indicate an adaptation to milk drinking culture in pastoral groups. The adaptive potential towards milk drinking appears to be more in volunteers with T_{-13910} allele from cattle domestication group. In the present study, the DNA sequences from PCR amplification of C/T_{-13910} SNP demonstrated no polymorphism at positions _13915, _13907, or _14010 and the sequence data conformed exactly to the reference sequence for this entire region. Thus, majority of Indian population (especially northern Indians) unlike some ethnic groups in sub-Saharan Africa shows the same genetic association of LP with C/T_{-13910} SNP as other northern European populations. It has been noted that T_{-13910} allele is extremely rare in sub-Saharan African populations, even in those populations where LP frequency had previously been reported to be high [12]. Three studies revealed several new sequence variants in very close proximity T_{-13910} allele [13–15] two of which are clearly associated with LP in different

parts of East Africa (G_{-13915} and C_{-14010}). One of these, G_{-13915}, was also shown to be associated with high lactase expression in Saudi Arabia [16]. A third SNP, G_{-13907}, showed much weaker evidence, but was found in several studies [13–17]. Romero et al. [18] reported that apart from *C/T*_13910, seven other segregating sites were found in Indian samples from pastoral populations with a combined frequency of the derived alleles 0.035. Of these 7 sites, the mutation G/C_{-13779} is the most common, with a countrywide sample frequency of 0.024. Also of interest is the C/T_{-13915} site, which was found in Muslim populations from southern India, and which Romero et al. [18] reported at a low frequency in other populations from the same region. These findings suggest that some other regulatory SNPs of LP may be present at a very low frequency in certain pastoral tribes of India.

In our study, sequence analysis of various samples from mixed northern Indian population did not reveal any new sequence variants apart from C/T_{-13910} SNP. Thus, the new sequence variants at positions G/C_{-13779} or C/T_{-13915} are very rare and may only be present at a very low frequency in certain pastoral tribes of India. Also, because of the small sample size, we may have missed some of the novel mutations associated with LP in various pastoral tribes of northern India. The study needs to be carried out with a larger sample size covering various pastoral groups of northern India. The study also needs to be supported by lactose tolerance test (LTT) or hydrogen breath test (HBT) to ascertain the LP/LNP status of an individual. Preliminary investigations suggest that T_{-13910} allele is a major determinant of LP in pastoral groups of northern Indian and may have undergone recent positive selection with the spread of cattle domestication and milk drinking culture. However, the origin of cattle domestication and the spread of T_{-13910} allele in India needs further research. It is unclear whether the T_{-13910} allele was introduced in India from outside or did the mutation arise *de-novo* in Indian population.

Acknowledgments

We thank all the volunteers for participating in this study. We acknowledge Rekha Madam, Mumtaz Anwar, Navdeep, Neha and Reyhan for helping in analysis and the collection of samples. Laboratory of Dr Safrun Mahmood was supported by Department of Biotechnology (DBT), New Delhi, India. Raja Amir was supported by Junior Research Fellowship, from University Grants Commission, New Delhi, India.

None of the authors reported any personal or financial conflict of interests.

References

[1] Swallow DM. Genetics of lactase persistence and lactose intolerance. Annu Rev Genet. 2003;37:197-219.

[2] Troelsen JT. Adult-type hypolactasia and regulation of lactase expression. Biochim Biophys Acta. 2005;1723:19-32.

[3] Anagnostou P, Battaggia C, Coia V, Capelli C, Fabbri C, Pettener D, Destro-Bisol G, Luiselli D.Tracing the distribution and evolution of lactase persistence in Southern Europe through the study of the T(-13910) variant. Am J Hum Biol. 2009;21:217-19.

[4] McCracken. Lactase deficiency: An example of dietary evolution. Current Anthropology. 1971;12:479-517.

[5] Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L, Jarvela I. Identification of a variant associated with adult-type hypolactasia. Nat Genet. 2002;30:233-7.

[6] Bersaglieri T, Sabeti PC, Patterson N, Vanderploeg T, Schaffner SF, Drake JA, Rhodes M, Reich DE, Hirschhorn JN. Genetic signatures of strong recent positive selection at the lactase gene. Am J Hum Genet. 2004;74:1111-20.

[7] Coelho M, Luiselli D, Bertorelle G, Lopes AI, Seixas S, Destro-Bisol G, Rocha J. Microsatellite variation and evolution of human lactase persistence. Hum Genet. 2005;117:329-39.

[8] Olds LC and Sibley E. Lactase persistence DNA variant enhances lactase promoter activity *in vitro*: Functional role as a *cis* regulatory element. Hum Mol Genet. 2003;12:2333-40.

[9] Holden C and Mace R. Phylogenetic analysis of the evolution of lactose digestion in adults. Hum Biol. 1997;69:605-28.

[10] Burger J, Kirchner M, Bramanti B, Haak W, Thomas MG. Absence of the lactase-persistence-associated allele in early Neolithic Europeans. Proc Natl Acad Sci USA. 2007;104:3736-41.

[11] Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG. The origins of lactase persistence in Europe. PLoS Comput Biol. 2009; 5:e1.

[12] Mulcare CA, Weale ME, Jones AL, Connell B, Zeitlyn D, Tarekegn A, Swallow DM, Bradman N, Thomas MG. The T allele of a single-nucleotide polymorphism 13.9 kb upstream of the lactase gene (LCT) (C-13.9kbT) does not predict or cause the lactase-persistence phenotype in Africans. Am J Hum Genet. 2004;74:1102-10.

[13] Ingram CJ, Mulcare CA, Itan Y, Thomas MG, Swallow DM. Lactose digestion and the evolutionary genetics of lactase persistence. Hum Genet. 2009;124:579-91.

[14] Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, Powell K, Mortensen HM, Hirbo JB, Osman M, Ibrahim M, Omar SA, Lema G, Nyambo TB, Ghori J, Bumpstead S, Pritchard JK, Wray GA, Deloukas P. Convergent adaptation of human lactase persistence in Africa and Europe. Nat Genet. 2007;39:31-40.

[15] Enattah NS, Jensen TG, Nielsen M, Lewinski R, Kuokkanen M, Rasinpera H, El-Shanti H, Seo JK, Alifrangis M, Khalil IF, Natah A, Ali A, Natah S, Comas D,Mehdi SQ, Groop L, Vestergaard EM, Imtiaz F, Rashed MS, Meyer B, Troelsen J, Peltonen L. Independent introduction of two lactase-persistence alleles into human populations reflects different history of adaptation to milk culture. Am J Hum Genet. 2008;82:57-72.

[16] Imtiaz F, Savilahti E, Sarnesto A, Trabzuni D, Al-Kahtani K, Kagevi I, Rashed MS, Meyer BF, Jarvela I. The T/G 13915 variant upstream of the lactase gene (LCT) is the founder allele of lactase persistence in an urban Saudi population. J Med Genet. 2007;44:e89.

[17] Ingram CJ, Elamin MF, Mulcare CA, Weale ME, Tarekegn A, Raga TO, Bekele E, Elamin FM, Thomas MG, Bradman N, Swallow DM. A novel polymorphism associated with lactose tolerance in Africa: Multiple causes for lactase persistence? Hum Genet. 2007;120:779-88.

[18] Romero IG, Mallick CB, Liebert A, Crivellaro F, Chaubey G, Itan Y, Metspalu M, Eaaswarkhanth M, Pitchappan R, Villems R, Reich D, Singh L,Thangaraj K, Thomas MG, Swallow DM, Mirazón Lahr M, Kivisild T. Herders of Indian and European cattle share their predominant allele for lactase persistence. Mol Biol Evol. 2012; 29:249-60.