Genotype frequencies of C/T_{-13910} and G/A_{-22018} polymorphisms in a Colombian Caribbean population do not correspond with lactase persistence prevalence reported in the region

Evelyn Mendoza, Microbiol[1], Adriana Carolina Hernández, Biol-Microbiol[2], Ricardo Wilches, MSc[3], Lourdes Luz Varela, PhD (C)[4], José Luis Villarreal, Microbiol[1], Luis Alejandro Barrera, PhD[5], Daniel Antonio Villanueva, PhD[6]

SUMMARY

Introduction: The C/T_{-13910} and G/A_{-22018} polymorphisms located upstream of the lactase gene are reliable predictors of lactase persistence in Caucasian-derived populations. Assessing the presence and distribution of these polymorphisms in other populations is central to developing genotyping assays and understanding the evolutionary mechanism behind this trait in several human populations.

Objective: Genotyping the C/T_{-13910} and G/A_{-22018} polymorphisms in a sample of Colombian Caribbean individuals.

Materials and methods: The polymorphisms were identified through Polymerase Chain Reaction/Restriction Fragment Length Polymorphism. Amplified fragments were digested using *Hinf I* and *Hha I*. Arlequin v. 3.1 was used to determine allelic and genotypic frequencies, Hardy Weinberg equilibrium, and linkage disequilibrium.

Results: Genotypic frequencies were CC (81.4%), CT (18.6%), and TT (0%) for the C/T_{-13910} polymorphism. Frequencies were AA (55.5%), GA (45.5%), and GG (0%) for the G/A_{-22018} polymorphism. No linkage disequilibrium was found between the two *loci*. Only the *locus* containing the C/T_{-13910} polymorphism was found in Hardy Weinberg equilibrium.

Conclusion: The allelic and genotypic distributions observed in this first genotyping study in a Colombian Caribbean population indicate a distribution pattern different from the one of the North European Caucasians and do not correspond to the lactase persistence prevalence reported for Caribbean populations.

Keywords: *Genetic polymorphism; Alleles; Lactase; Genotype; Frequencies; Genotyping.*

Las frecuencias genotípicas de los polimorfismos C/T_{-13910} and G/A_{-22018} en una población caribeña colombiana no corresponden con la prevalencia de lactasa persistencia que se informó en la región

RESUMEN

Introducción: Los polimorfismos C/T_{-13910} y G/A_{-22018}, que se localizan corriente arriba del gen de la lactasa son predictores confiables de la persistencia de lactasa en poblaciones derivadas de caucásicos. Conocer la presencia y distribución de esos polimorfismos en otras poblaciones es fundamental para el desarrollo de métodos de diagnóstico de lactasa persistencia y para comprender los mecanismos evolutivos de este fenotipo en seres humanos.

Objetivo: Genotipificar los polimorfismos C/T_{-13910} y G/A_{-22018} en una muestra de sujetos caribeños colombianos.

1. Research Professor, Grupo de Investigación en Bioquímica Patológica (GRUBIOPAT), Universidad Libre, Barranquilla, Colombia. e-mail: evemendoza5@hotmail.com joseluisvillarrealcamacho@hotmail.com
2. Researcher, Instituto de Errores Innatos del Metabolismo (IEIM), Pontificia Universidad Javeriana, Bogotá, Colombia. e-mail: adrianacarolina@gmail.com
3. PhD student, Gradschool for Evolution, Ecology and Systematics, University of Munich, Planegg-Martinsried, Germany. e-mail: wilches@bio.lmu.de
4. Genetics Professor, Grupo de Investigación en Bioquímica Patológica (GRUBIOPAT), Universidad Libre, Barranquilla, Colombia. e-mail: lourdesvarela@hotmail.com
5. Director, Instituto de Errores Innatos del Metabolismo (IEIM), Pontificia Universidad Javeriana, Bogotá, Colombia. e-mail: abarrera@javeriana.edu.co
6. Director, Grupo de Investigación en Bioquímica Patológica (GRUBIOPAT), Universidad Libre, Barranquilla, Colombia. e-mail: danielvillanueva@unilibrebaq.edu.co

Received for publication March 11, 2009 Accepted for publication June 3, 2010

Materiales y métodos: Los polimorfismos se identificaron mediante la digestión de productos amplificados, que se hizo con *Hinf I* y *Hha I*. Se usó el programa Arlequín versión 3.1 para determinar las frecuencias alélicas y genotípicas, el equilibrio de Hardy-Weinberg y el desequilibrio de ligamiento.

Resultados: Para el polimorfismo C/T_{-13910} las frecuencias genotípicas fueron CC (81.4%), CT (18.6%) y TT (0%), mientras que para el polimorfismo G/A_{-22018} fueron AA (55.5%), GA (45.5%) y GG (0%). No se encontró desequilibrio de ligamiento entre los *loci* que contienen los polimorfismos y sólo el polimorfismo C/T_{-13910} está en equilibrio de Hardy-Weinberg en comparación con G/A_{-22018}.

Conclusión: Las distribuciones alélicas y genotípicas observadas en este primer estudio de genotipificación en una muestra de la población caribeña colombiana muestra un patrón de distribución diferente del encontrado en poblaciones caucásicas del norte de Europa y no corresponden con la prevalencia de lactasa persistencia que se ha informado en caribeños.

Palabras clave: *Polimorfismo genético; Alelos; Lactasa; Genotipo; Frecuencias; Genotipificación.*

There are two lactase phenotypes in humans. The first is known as primary adult-type hypolactasia. It is characterized by a decrease in the activity of lactase after weaning, retaining only 5-10% of the enzymatic activity in adulthood[1]. The second is lactase persistence, due to high lactase activity maintained through adulthood. Only a minority of the world population retains lactase activity, noted mainly in northern Europeans[2,3].

In individuals with primary adult-type hypolactasia, the manifestations of the impairment of lactose digestion include flatulence, diarrhea, and abdominal pain, which are consequences of lactose metabolism by the colonic microflora. This clinical profile is known as lactose intolerance[4].

The gold standard for the diagnosis of primary adult-type hypolactasia is the evaluation of lactase activity through intestinal biopsy[5]. This method is invasive, cumbersome, and only evaluates activity in a very small area in the surface of intestinal mucosa.

In search of a more sensitive, efficient, and less invasive diagnostic method, Enattah *et al.*[6], studied a cohort of 196 non-related individuals of finnish origin, reporting a 100% correlation of the lactase persistence phenotype with the presence of allele T of a SNP located 13910 bp upstream of the lactase gene. This C/T -13910 polymorphism is located in intron 13 of the *MCM6* gene (C/T_{-13910}). Furthermore, they found a high correlation (96%) between lactase persistence and the presence of allele A in SNP G/A, located in intron 9 of the *MCM6*, 22018 bp upstream of lactase gene (G/A_{-22018}).

A high correlation between T_{-13910}, A_{-22018} and the lactase persistence phenotype has also been reported in a Brazilian population of Caucasian descent[7] and many European countries[8-10]. On the other hand, in lactose-persistant sub-Saharan African and Middle Eastern populations, such a high correlation was not found. Moreover, novel SNPs exclusive in these populations show association with lactase persistence [11-13].

The behavior of allelic frequencies of these two SNPs in the Colombian population is thus far unknown. Historically, this population has been regarded as the mixture of three main ethnicities: Africans, Amerindians, and Caucasians[14]. Learning about the prevalence of these alleles in different Colombian subpopulations will clarify whether it is useful to genotype them during clinical assessments for lactase persistence. We have aimed at genotyping the SNPs C/T_{-13910} and G/A_{-22018} in a Colombian Caribbean population. The observed frequencies for these SNPs are presented and discussed in light of the existing Colombian and international information.

MATERIALS AND METHODS

A survey was done on 367 individuals between 17 and 69 years of age (Mean: 30±1), 156 met the following inclusion criteria: they were born in the Colombian Caribbean with no family ties among them, and both their parents and grandparents were also born in the Caribbean.

All the individuals selected signed an informed consent form and the study was approved by the Ethics Committee from Universidad Libre, Barranquilla, Colombia.

DNA was obtained from venous blood using the Wizard® Genomic DNA Purification Kit (Promega, USA). Polymorphism identification was done by Polymerase Chain Reaction/Restriction Fragment Length Polymorphism (PCR/RFLP). To identify the alleles of the SNP C/T_{-13910}, a fragment of 201 bp was first amplified using the primers: 5'-GCTGGCAA TACAGATAAGATAATGGA-3' and 5'-CTGCTTT GGTTGAAGCGAAGAT-3' (11). Reaction was carried

Table 1
Genotypic frequencies observed and expected for SNPs C/T_{-13910} and G/A_{-22018}, and their relationship with phenotypes reported by Enattah *et al.*, in Finnish individuals. Genotypes TT and GG were not found in the population studied

Polymorphism	Phenotype reported by Enattah *et al.*[6]	Genotype	Genotypic frequency (n=156)	
			Observed	Expected according to Hardy Weinberg
C/T_{-13910}	Hypolactasia	CC	127 (0.814)	128.5 (0.824)
	Persistence	CT	29 (0.186)	26.5 (0.167)
	Persistence	TT	0	1.38 (0.008)
G/A_{-22018}	Persistence	AA	85 (0.545)	92.49 (0.59)
	Persistence	GA	71 (0.455)	55.22 (0.35)
	Hypolactasia	GG	0	8.25(0.06)

Table 2
Allelic frequencies for SNPs C/T_{-13910} and G/A_{-22018}

Allele	Frequency (n=312)	%
C/T_{-13910}		
C	283	90.7
T	29	9.3
G/A_{-22018}		
A	241	77.3
G	71	22.7

out in a total volume of 20 µL, with Tris HCl buffer 1X pH 9.0, $MgCl_2$ 1.5 mM, dNTPs 0.2 mM, *Taq* polymerase 1.5U and 1 µM for each primer (Invitrogen®, USA). PCR initiated with a denaturation step of 95ºC for 10 min followed by 35 cycles of 95ºC for 1 min, 59ºC for 1 min and 72ºC for 1 min, and a final elongation cycle at 72ºC for 8 min. PCR product was digested with *Hinf* I (New England Biolabs, USA). Amplification and digestion products were run in an 8% polyacrylamide gel and visualized with ethidium bromide staining. Samples that only presented a 201 bp (C) fragment or a 177 bp (T) fragment were interpreted as CC and TT genotype, respectively, while samples that presented two fragments of 201 bp and 177 bp were interpreted as the CT genotype.

The 271 bp region that contained SNP G/A_{-22018} was amplified with the primers: 5'-CTCAGTGATCC TCCCACCTC-3' and 5'-CCCCTACCCTATCAG TAAAGGC-3' (Invitrogen, USA)[15]. PCR conditions were the same as for C/T_{-13910}, but reducing the amplification cycles to 34 and using an annealing temperature of 62°C. PCR product was digested with *Hha* I (New England Biolabs, USA) and visualized as described for the other SNP. The samples that only presented a 271 bp fragment (A) or a 196 bp fragment (G) were interpreted as genotype AA and GG, respectively, while samples that displayed two fragments of 271 bp and 196 bp were interpreted as GA genotype.

Allelic and genotypic frequencies, as well as Hardy-Weinberg equilibrium and linkage disequilibrium (LD) were determined by using Arlequin version 3.1 software with 95% confidence interval[16].

RESULTS

Genotype and allele frequencies found in this study are shown in Tables 1 and 2, respectively. From the genotypes found, the following haplotypes were inferred with their respective frequencies: CA (69%), CG (22%), TA (8%), and TG (1%).

Allelic frequencies for the genotypes CC, CT, AA, and GA were used to analyze the Hardy-Weinberg equilibrium. Only SNP C/T_{-13910} was in Hardy-Weinberg equilibrium ($p=0.367$; >0.05). Likewise, LD was analyzed and independent segregation was found for C/T_{-13910} and G/A_{-22018} ($p=0.335$; >0.05).

DISCUSSION

We determined the frequencies of two SNPs associated with the lactase persistence in Europeans

from a sample of *mestizos* from the Colombian Caribbean population. Upon reviewing the genotypic distribution of SNP G/A_{-22018} in Table 1, it can be noted that the totality of the individuals have the dominant A_{-22018} allele, which is bound to lactase persistence in North European populations. If this allele were to perform the same function in the population under study, 100% of these individuals would be persistent as well. This is not coherent with the frequency data of the lactase persistence phenotype reported for Caribbean populations (60%) and also not with the data reported for Colombian Andean populations in which the influence of Caucasian European populations has been greater[17,18]. These data also differ from the frequency of the dominant and indicative of persistence in Europeans, allele T of SNP C/T_{-13910}. The contrast between the high frequency of allele A_{-22018} and the low frequency of allele T_{-13910}, both associated to lactase persistence in northern European population, could be explained by the differences in evolutionary forces acting upon each of the SNPs. In fact, only C/T_{-13910} was found in Hardy–Weinberg equilibrium. Furthermore, that difference cannot be explained as a consequence of a biased selection in the population, because the inclusion criteria were strict.

Our identified T_{-13910} allele frequency is in close agreement with that found in several African populations, where is too low or absent and cannot explain reported frequencies of lactase persistence in such populations[11]. This finding led Ingram *et al.*[12], to suspect that the cause of lactase persistence in Africans came from different SNPs. Indeed, a handful of SNPs, in the vicinity of C/T_{-13910}, found in different lactase persistent populations, including Africans and Middle East pastoralists, have been identified and functionally tested. Such novel SNPs, namely the G_{-13907}, G_{-13915}, and C_{-14010} alleles constitute putative causes of lactase persistence in non-European populations[13,19]. Surveys carried out in other populations highlight the importance of ethnic composition when looking for genotype-phenotype associations. Haiming *et al.*[20], studied SNP C/T_{-13910} in Chinese populations and did not find concordance with lactase phenotypes; they concluded that allele T, at this low frequency (1.9%), cannot be used as a predictor for lactase persistence. In contrast, Bulhoes *et al.*[7], demonstrated in Brazil concordance between polymorphisms C/T_{-13910}, G/A_{-22018}, and lactase persistence as revealed by lactose absorption in the intestine. However, 19 out of the 20 genotyped individuals were of Caucasian descent.

The fact that the demographic history of the population under study includes the admixture of three main founder ethnicities, namely Amerindians, Europeans, and Africans[14,21], is central to understanding the absence of LD among studied SNPs. Recent studies in Northern-European populations[6], where strong LD between both loci is an inferred consequence of ongoing positive selection with concomitant hitch-hiking of the allele A-22018 with the mutation T_{-13910}[22,23], can help us rule out the role of selection acting on this genomic neighborhood in our studied population. Further analyses in our cohort, including more markers in the genomic vicinity are required to test whether the observed lack of LD is correlated with the demographic history of the studied population.

In those populations which have shown high association between genotype and phenotype, genotyping is used as a diagnostic method[6,24]. An association study implies that the genotypic distribution has to be known. The genotype frequencies found show that the alleles studied would not be predictors of lactase persistence in the population studied.

Considering that the presence of European Caucasians, Africans, Amerindians, and other ethnic groups, such as the Sephardic Jews and Turks, have influenced the composition of the Colombian population in several regions of the country[21], we would expect that future C/T_{-13910} and G/A_{-22018} SNPs analysis of other samples of the Colombian population will reveal different haplotype compositions and allele segregations. We suggest that genotypic frequencies and association analysis be carried out to assess the behavior of SNPs in Colombia and their potential to be used as a diagnostic tool for lactase persistence.

ACKNOWLEGMENTS

The authors thank Marena Rodríguez for genetic tests assessment and Dr. Jorge Luis Bilbao for his valuable observations to this manuscript.

REFERENCES

1. Sahi T. Hypolactasia and lactase persistence. Historical review and terminology. *Scand J Gastroenterol.* 1994; *29 Suppl 202:* 1-6.
2. Holden C, Mace R. Phylogenetic analysis of the evolution of lactose digestion in adults. *Hum Biol.* 1997; *69:* 605-28.
3. Hollox EJ, Poulter M, Zvarik M, Ferak V, Krause A, Jenkins T, *et al*. Lactase haplotype diversity in the old world. *Am J Hum Genet.* 2001; *68:* 160-72.
4. Lee M, Krasinski S. Human adult-onset lactase decline: An update. *Nutr Rev.* 1998; *156:* 1-8.
5. Dahlqvist A. Assay of intestinal disaccharidases. *Scand J Clin Lab Invest.* 1984; *44:* 169-72.
6. 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.
7. Bulhoes AC, Goldani HAS, Oliveira FS, Matte US, Mazzuca RB. Correlation between lactose absorption and the C/T $_{-13910}$ and G/A $_{-22018}$ mutations of the lactase-phlorizin hydrolase (LCT) gene in adult-type hypolactasia. *Braz J Med Biol Res.* 2007; *40:* 1441-6.
8. Krawczyk M, Wolska M, Schwartz S, Gruenhage F, Terjung B, Portincasa P, *et al*. Concordance of genetic and breath tests for lactose intolerance in a tertiary referral centre. *J Gastrointestin Liver Dis.* 2008; *17:* 135-9.
9. Ingram CJE, Mulcare CA, Itan Y, Thomas MG, Swallow DM. Lactose digestion and the evolutionary genetics of lactase persistence. *Hum Genet.* 2009; *124:* 579-91.
10. Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG. The origins of lactase persistent in Europe. *PLoS Comput Biol.* 2009; *5:* e1000491.
11. Mulcare CA, Weale ME, Jones AL, Connell B, Zeitlyn D, Tarekegn Á, *et al* . 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.
12. Ingram CJE, Elamin MF, Mulcare CA, Weale ME, Tarekegn A, Raga TO, *et al*. A novel polymorphism associated with lactose tolerance in Africa: multiple causes for lactase persistence? *Hum Genet.* 2007; *120:* 779-88.
13. Imtiaz F, Savilahti E, Sarnesto A, Trabzuni D, Al-Kahtani K, Kagevi I, *et al*. T/G $_{-13915}$ variant upstream of the lactase gene (LCT) is the founder allele of adult-type hypolactasia in an Arab population(s). *J Med Genet.* 2007; *44:* e89.
14. Bedoya G, Montoya P, García J, Soto I, Bourgeois S, Carvajal L, *et al*. Admixture dynamics in Hispanics: A shift in the nuclear genetic ancestry of a South American population isolate. *PNAS.* 2006; *103:* 7234-9.
15. Coelho M, Luiselli D, Bertorelle G, Lopes AI, Seixas S, Destro-Bisol G, *et al*. Microsatellite variation and evolution of human lactase persistence. *Hum Genet.* 2005; *117:* 329-39.
16. Excoffier L, Laval G, Schneider S. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform. Online* 2005; *1:* 47-50.
17. American Society for Clinical Nutrition. The acceptability of milk and milks products in populations with a high prevalence of lactose intolerance. *Am J Clin Nutr.* 1988; *48:* 1086-98.
18. Ángel LA, Calvo E, Muñoz Y. Prevalencia de hipolactasia tipo adulto e intolerancia a la lactosa en adultos jóvenes. *Rev Col Gastroenterol.* 2005; *20:* 35-47.
19. Enattah NS, Jensen TG, Nielsen M, Lewinski R, Kuokkanen M, Rasinpera H, *et al*. 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.
20. Hai-ming S, Yuan-dong Q, Feng C, Li-dan X, Jing B, Song-bin F. The lactase gene $_{-13910}$T allele can not predict the lactase-persistence phenotype in north China. *Asia Pac J Clin Nutr.* 2007; *16:* 598-601.
21. Carvajal-Carmona LG, Soto ID, Pineda N, Ortiz-Barrientos D, Duque C, Ospina-Duque JH, *et al*. Strong Amerind/Caucasoid gender bioa and evidence of a sepharadic contribution among the founders of a population in North West Colombia. *Am J Hum Genet.* 2000; *67:* 1287-95.
22. Bersaglieri T, Sabeti PC, Patterson N, Vanderploeg T, Schaffner SF, Drake JA, *et al*. Genetic signatures of strong recent positive selection at the lactase gene. *Am J Hum Genet.* 2004; *74:* 1111-20
23. Fay JC, Wu C. Hitch-hiking under positive Darwinian selection. *Genetics.* 2000; *155:* 1405-13.
24. Rasinpera H, Savilahti E, Enattah NS, Kuokanen M, Totterman N, Lindahl H, *et al*. A genetic test which can be used to diagnose adult-type hypolactasia in children. *Gut.* 2004; *53:* 1571-6.