Lactase persistence DNA variant enhances lactase promoter activity in vitro: functional role as a cis regulatory element

Lynne C. Olds and Eric Sibley*

Department of Pediatrics, Stanford University Medical Center, Stanford, CA 94305, USA

Received May 2, 2003; Revised and Accepted July 7, 2003

Lactase persistence is a heritable, autosomal dominant, condition that results in a sustained ability to digest the milk sugar lactose throughout adulthood. The majority of the world’s human population experiences a decline in production of the digestive enzyme lactase-phlorizin hydrolase during maturation. However, individuals with lactase persistence continue to express high levels of the lactase gene into adulthood. Lactase persistence has been strongly correlated with single nucleotide genetic variants, C/T\textsubscript{13910} and G/A\textsubscript{22018}, located 13.9 and 22 kb upstream from the lactase structural gene. We aimed to characterize a functional role for the polymorphisms in regulating lactase gene transcription. DNA in the region of the C/T\textsubscript{13910} or G/A\textsubscript{22018} human lactase variants was cloned upstream of the 3.0 kb rat lactase gene promoter in a luciferase reporter construct. Human intestinal Caco-2 cells were transfected with the lactase variant/promoter–reporter constructs and assayed for promoter activity. A 200 bp region surrounding the C\textsubscript{13910} variant, associated with lactase non-persistence, results in a 2.2-fold increase in lactase promoter activity. The T\textsubscript{13910} variant, associated with lactase persistence, results in an even greater 2.8-fold increase. The DNA sequence of the C/T\textsubscript{13910} variants differentially interacts with intestinal cell nuclear proteins on EMSAs. AP2 co-transfection results in a similar repression of the C/T\textsubscript{13910} variant/promoter–reporter constructs. The DNA region of the C/T\textsubscript{13910} lactase persistence/non-persistence variant functions in vitro as a cis element capable of enhancing differential transcriptional activation of the lactase promoter. Such differential regulation by the C and T variants is consistent with a causative role in the mechanism specifying the lactase persistence/non-persistence phenotypes in humans.

INTRODUCTION

Intestinal lactase-phlorizin hydrolase (LPH, lactase) is the absorptive enterocyte membrane glycoprotein essential for digestive hydrolysis of lactose in milk. Lactase is present predominantly along the brush border membrane of differentiated enterocytes lining the villi of the small intestine. Expression of the lactase gene is temporally regulated during gut development and maturation (1,2). Lactase enzyme activity is maximal in the small intestine of pre-weaned mammals and declines markedly during maturation. In laboratory rodents, the maturational decline in lactase activity has been shown to occur at the time of weaning and contrasts with a maturational increase in enzymatic activity of other intestinal hydrolases essential for digestion of solid foods (comprehensively reviewed in 3). In humans, the activities of lactase and most of the other digestive hydrolases are maximal at birth. The age of onset of the post-weaning maturational decline of lactase activity in humans is variable, ranging from the toddler years to young adulthood. Lactose is the predominant carbohydrate in breast milk. The lactose disaccharide consists of a glucose and a galactose molecule linked by a β(1–4) glycosidic bond. The digestive enzyme lactase catalyzes the hydrolysis of lactose to yield glucose and galactose that can then be absorbed across the intestinal mucosa membrane. The maturational decline in lactase activity renders most of the world’s adult human population intolerant to milk and other dairy products (4,5). Lactose present in dairy products cannot be digested in the small intestine, due to the reduced lactase level, and instead is fermented by bacteria in the distal ileum and colon. The fermentative products result in symptoms of diarrhea, gas bloat, flatulence and abdominal pain.

In a minority of adults, high levels of lactase activity persist in adulthood. This hereditary persistence of lactase is common

*To whom correspondence should be addressed at: Department of Pediatrics, Stanford University School of Medicine, 750 Welch Road, Suite 116, Palo Alto, CA 94304, USA. Tel: +1 6507235070; Fax: +1 6504985608; Email: erc@stanford.edu

Human Molecular Genetics, Vol. 12, No. 18 © Oxford University Press 2003; all rights reserved
primarily in people of northern European descent and is attributed to inheritance of an unidentified autosomal dominant mutation that prevents the normal maturational decline in lactase expression (6). Enattah et al. (7), using linkage disequilibrium and haplotype analysis in humans, have recently reported the identification of genetic variants located −13 910 and −22 018 bp upstream of the human lactase gene that are associated with hereditary lactase persistence/non-persistence in Finnish families (7). The −13 910 bp C/T −13 910 variant is located 5′ to the lactase gene within intron 13 of the adjacent MCM6 gene (8) on chromosome 2. The intron 13 variant is a single nucleotide polymorphism, C to T. The authors report complete correlation between the lactase non-persistence phenotype and homozygosity for the C variant. Similarly, complete correlation is reported between the lactase persistence phenotype and the presence of the T variant allele. In the same study, a second variant (G/A −22 018) was associated with the lactase persistence trait. Although not completely associated, this second variant may be involved in regulating the lactase activity. The C/T −13 910 variant, however, is a strong candidate for the regulatory element involved in mediating lactase persistence. In a recent report, the lactase persistence allelic genotype T−13 910/A−22 018 has been shown to be associated with increased levels of lactase mRNA transcripts compared with the non-persistence allelic genotype C−13 910/G−22 018 (9). While strongly correlated with the lactase persistence/non-persistence phenotypes, the C/T −13 910 variant has not yet been shown to be causative of the traits and a functional mechanism has not been defined (10). Enattah et al. (7) have postulated that the C/T −13 910 variant affects an AP2 site, noting that the C allele associated with lactase non-persistence is in the consensus binding motif whereas the T variant disrupts this motif. In the present paper, we report that the C/T −13 910 lactase persistence/non-persistence variant functions as a cis element capable of directing differential transcriptional activation of the lactase promoter.

RESULTS

The lactase persistence C/T −13 910 variant region enhances lactase promoter activity in intestinal cell culture

We have previously demonstrated that a 3.0 kb rat lactase promoter fragment (11) is capable of driving transcription in Caco-2 cells, a human adenocarcinoma-derived cell-line that mimics a small intestinal enterocyte phenotype with respect to expression of several digestive hydrolases including lactase (12). In order to determine whether the DNA sequence surrounding the C/T −13 910 human lactase persistence variant is capable of regulating lactase promoter activity, Caco-2 cells were transfected with lactase promoter–reporter constructs harboring the C or T variant nucleotides. The constructs were generated by cloning the C/T −13 910 region of the human lactase promoter upstream of the 3.0 kb rat lactase promoter in the luciferase reporter plasmid pGL3Basic. The C/T −13 910 region (−14 017 to −13 800) was PCR-amplified from Caco-2 cell genomic DNA in order to generate the T variant reporter constructs p3kLac-T (forward orientation) and p3kLac-T′ (reverse orientation). Similarly the C/T −13 910 region was amplified from human RP11-34L23 BAC clone DNA to yield the C variant constructs p3kLac-C and p3kLac-C′. The C/T −13 910 region in each construct was sequenced to confirm differences only at the −13 910 single nucleotide variant. Transfected Caco-2 cell extracts were assayed for relative luciferase activity 48 h after transfection as shown in Figure 1.

Transfection with the lactase non-persistence associated C variant construct, p3kLac-C, results in a 2.2-fold induction of promoter activity driven by the 3 kb lactase promoter. Transfection with the lactase persistence-associated T variant construct, p3kLac-T, results in a 2.8-fold induction of promoter activity driven by the 3.0 kb lactase promoter. While both C/T −13 910 regions are capable of enhancing promoter activity, the lactase persistence-associated T variant possesses a significantly greater activity compared to the non-persistence-associated C variant (P < 0.00001, p3kLac-T versus p3kLac-C). Placement of the C/T −13 910 variant region in both forward and reverse orientations results in comparable activation of the lactase promoter. Differential enhancement of lactase promoter activity mediated by the C/T −13 910 variant region suggests a functional role for the DNA variant in specifying the lactase persistence/non-persistence phenotypes.

The G/A −22 018 variant region directs minimal enhancement of lactase promoter activity in intestinal cell culture

A second genomic variant, G/A −22 018, located upstream of the C/T −13 910 variant region has also been associated with the lactase persistence/non-persistence traits. Although not completely associated, this second variant may also have a role in regulating lactase promoter activity. To determine whether the G/A −22 018 variant region is involved in regulating the human lactase promoter, the G/A −22 018 region (22 133–21 910) was cloned in both orientations upstream of the 3.0 kb rat lactase promoter to generate p3kLac-G, p3kLac-G′, p3kLac-A and p3kLac-A′. Transfection with the lactase non-persistence associated G variant constructs (p3kLac-G and p3kLac-G′) or with the lactase persistence-associated A variant construct (p3kLac-A) results in minimal, ~1.2- to 1.4-fold, induction of promoter activity (Fig. 2). The A variant region cloned in the reverse orientation (p3kLac-A′) results in an ~0.8-fold reduction in promoter activity.

To determine whether the G/A −22 018 variant region is capable of regulating the human lactase promoter in the context of the C/T −13 910 variant, the G/A −22 018 region was cloned upstream of the C/T −13 910 region to generate p3kLac-GC and p3kLac-AT. The G −22 018 variant and the C −13 910 variant are both associated with lactase non-persistence. The addition of the G −22 018 region upstream of the C −13 910 region does not result in a significant change in lactase promoter activity (Fig. 3). The A −22 018 variant and the T −13 910 variant are both associated with lactase persistence. Addition of the A −22 018 region upstream of the T −13 910 Region results in a minimal, ~1.2-fold, increase in lactase promoter activity. Thus, the G/A −22 018 variant region appears to direct minimal enhancement of the lactase promoter in context of the C/T −13 910 variant region in the transfected reporter constructs.
AP2 overexpression represses lactase promoter activity in intestinal cell culture

It has been postulated that the C/T\textsubscript{13910} variant may affect an AP2 transcription factor binding site. The C allele, associated with lactase non-persistence, is in the consensus binding motif whereas the T variant disrupts this motif (7). In order to determine whether the transcription factor AP2 is involved in mediating the C/T\textsubscript{13910} region transcriptional activation of the lactase promoter, Caco-2 cells were co-transfected with the C/T\textsubscript{13910} variant/promoter constructs and AP2\textalpha and AP2γ expression constructs (Fig. 4). Co-transfection with the AP2\textalpha or AP2γ construct results in a 1.7-fold or greater repression in promoter activity driven by both C/T\textsubscript{13910} variant/promoter–reporter constructs as well as the 3.0 kb rat lactase promoter. This result suggests that AP2 regulatory cis element sequences may reside in the proximal rat lactase promoter but does not support a role for AP2 in mediating the lactase C/T\textsubscript{13910} variant effect.

Differential interaction between the C/T\textsubscript{13910} variant cis elements and intestinal cell nuclear proteins

In order to identify interaction between the C/T\textsubscript{13910} Variant cis element and intestinal cell nuclear proteins, we employed the electrophoretic mobility shift assay. C and T variant binding region oligonucleotide probes were incubated in the presence of Caco-2 cell (~90% confluent) nuclear extract and then migrated through a non-denaturing acrylamide gel. The EMSA in Figure 5 reveals the position of the rapidly migrating unbound probe at the base of the gel and DNA–protein complexes of slower mobility formed after incubation with Caco-2 nuclear extract in the absence of competition (lanes 4 and 5). A slow migrating DNA–protein complex, indicated by arrow, is more abundant in incubation with the T\textsubscript{13910} cis element probe (lane 4) compared with the C\textsubscript{13910} probe (lane 5). The T\textsubscript{13910} element DNA–protein complex is competed by 200-fold excess unlabeled T\textsubscript{13910} oligonucleotide (lane 3), but is not competed away by 200-fold excess of an unrelated Sp1 oligonucleotide (lane 2). The nuclear protein bound to the T\textsubscript{13910} cis element probe in Caco-2 cells therefore represents a specific DNA–protein interaction. In contrast, the C\textsubscript{13910} element DNA–protein complex is much less abundant and is not competed away by 200-fold excess C\textsubscript{13910} (lane 6) or Sp1 (lane 7) oligonucleotides. To confirm specificity of binding to the T\textsubscript{13910} cis element, EMSAs reactions were also performed with the T\textsubscript{13910} cis element probe in the presence of C\textsubscript{13910} cis element competitor DNA. The T\textsubscript{13910} cis element DNA–protein complex was not competed by 200-fold excess C\textsubscript{13910} cis element oligonucleotide (Fig. 6, lane 1). The nuclear protein interacting with the T\textsubscript{13910} cis element probe is thus a candidate factor that may be involved in differential transcriptional activation mediated by the C/T\textsubscript{13910} variant.
variant region. EMSA reactions performed with nuclear extract isolated from 7 day post-confluent Caco-2 cells result in a comparable T/C0 \(_{13910}\) cis element specific DNA–protein complex and AP2 antisera does not detect proteins interacting with the C/T \(_{13910}\) elements on gel supershifts (data not shown).

**DISCUSSION**

Hereditary persistence of lactase is common primarily in people of northern European descent and has been attributed to inheritance of a hitherto unidentified autosomal dominant mutation that prevents the normal maturational decline in lactase expression (6). The recent identification by Enattah et al. (7) of genetic variants associated with lactase persistence has prompted interest in determining whether the single nucleotide variants are causative. The intron 13 variant is a single nucleotide polymorphism (C to T) located at nucleotide \(_{13910}\) relative to the lactase start site and was completely correlated with lactase persistence/non-persistence. With respect to regulation of the lactase persistence/non-persistence phenotype, previous studies have suggested both transcriptional and post-transcriptional control mechanisms (6,13–18).

In the present study, we have focused on characterization of the role of the C/T \(_{13910}\) and G/A \(_{22018}\) regions in regulating the lactase gene promoter in intestinal cell culture. Specifically, the DNA surrounding the C/T \(_{13910}\) and G/A \(_{22018}\) polymorphisms was assayed for the ability to function as a transcriptional activator or repressor in transfected Caco-2 cells. With respect to the C/T \(_{13910}\) region, DNA fragments harboring both the C and T variants are capable of enhancing lactase promoter (Fig. 1). Of note, however, the T variant results in greater promoter activation as compared to the C variant. Such an enhanced activity mediated by the T variant would be consistent with a functional role in the mechanism of lactase persistence. It is possible that a more dramatic effect was not seen due to the nature of transient transfection studies. Transfected Caco-2 cells contain multiple non-replicating episomal copies of the reporter constructs. Transiently transfected constructs are not always subject to regulatory mechanisms that involve complex chromatin structure, in contrast to the diploid chromosomal complement of the endogenous gene (19). In addition, the transfection constructs assayed likely lack other essential co-regulatory sequences located outside the 200 bp C/T \(_{13910}\) region. Caco-2 cells, while a model of intestinal cell differentiation, are not identical to adult enterocytes *in vivo* and as such may also lack essential co-regulatory proteins needed to specify maximal phenotypic expression patterns. Thus, the C/T \(_{13910}\) variants may exert a significantly greater differential regulatory effect *in vivo* in the context of the endogenous lactase gene. Compared with the C/T \(_{13910}\) region, the G/A \(_{22018}\) variant region results in minimal enhancement of the lactase promoter (Fig. 2). Unlike the C/T \(_{13910}\) variant, the G/A \(_{22018}\) variant was not completely correlated with the lactase

![Figure 2](https://academic.oup.com/hmg/article-abstract/12/18/2333/2355640/2336)
persistence/non-persistence phenotype in the original report by Enattah et al. (7). Therefore, the G/A/C022018 variant may be associated with but not causative of the lactase phenotypes.

Enattah et al. (7) have postulated that the C/T/C013910 variant affects an AP2 binding site, noting that the C allele, associated with lactase non-persistence, is in the consensus binding motif, whereas the T variant disrupts this motif. To determine whether AP2 overexpression differentially regulates the C/T/C013910 variant element, we co-transfected Caco-2 cells with the variant–promoter constructs and with AP2α and AP2γ. Interestingly, AP2 over-expression results in a repression of the lactase promoter reporter constructs (Fig. 4). However, the fold repression is comparable for both the C and T allelic constructs. In addition, AP2 antisera do not detect proteins interacting with the C/T/C013910 element on gel supershifts (data not shown). These results do not support a direct role for AP2 binding.

The DNA sequence surrounding the C/T/C013910 variants, however, does interact differentially with intestinal cell nuclear proteins on gel shift assays. A slow migrating DNA–protein complex is more abundant in gel shift reactions performed with the T/C013910 cis element probe than in reactions with the C/C013910 probe (Fig. 5, compare lanes 4 and 5). Such a differential gel shift binding pattern combined with the differential transactivation results suggests a possible mechanism to account for the lactase persistence/non-persistence phenotypes. Namely, the DNA sequence surrounding the T/C013910 variant may create a binding site for a transactivating protein that is capable of enhancing lactase transcription in adults with lactase persistence, thereby preventing the maturational decline. Such a mechanism would be consistent with a dominant mode of inheritance for lactase persistence associated with the T/C013910 allelic polymorphism. In individuals with lactase non-persistence, i.e. homozygous for the C/C013910 variant, the C nucleotide may disrupt the transactivating protein binding site and thus not allow for sustained lactase transcription into adulthood. There are likely to be additional regulatory control regions that specify the maturational decline in lactase in these individuals. Studies looking at the effect of the variants in stably transfected cell lines or in living animals will be necessary to confirm a role for the C/T/C013910 variant region in specifying lactase persistence or non-persistence. In addition, to characterize a potential regulatory mechanism directed by the C/T/C013910 region, further studies must be directed at identification of nuclear proteins interacting with this DNA region.

MATERIALS AND METHODS

Materials and reagents

Restriction endonucleases were purchased from Life Technologies (Rockville, MD, USA). Reagents for PCR were obtained from Qiagen (Valencia, CA, USA). Oligonucleotides
were synthesized by the protein and nucleic acid (PAN) facility of the Stanford University Medical Center.

PCR amplification of the C/T\textsubscript{-13910} and G/A\textsubscript{-22018} regions of the human lactase gene locus

A 218 bp fragment of the human lactase LCT gene C/T\textsubscript{-13910} variant region was PCR-amplified using a forward oligonucleotide corresponding to nt -14017 to -13994, 5'-AGACGTAAGTTACCATTTAATC-3', and a reverse oligonucleotide corresponding to nt -13800 to -13821, 5'-CGTAAATACCCACTGACCTATC-3'. Both primers were synthesized with a 5' terminal Mlu restriction site for subsequent cloning. The C/T\textsubscript{-13910} region (-14017 to -13800) was PCR-amplified from Caco-2 cell genomic DNA to generate the T variant region DNA fragment. Similarly the C/T\textsubscript{-13910} region was amplified from RP11-34L23, a human genomic DNA BAC clone (BACPAC Resources), to yield the C variant region DNA fragment. A 224 bp fragment of the lactase gene G/A\textsubscript{-22018} variant region was PCR-amplified using a forward oligonucleotide corresponding to nt -22133 to -22113, 5'-TAAGAACATTCTACTCTTT-3', and a reverse oligonucleotide corresponding to nt -21910 to -21930, 5'-AGAAAAATGGGTTTTCGCAATG-3'. Both primers were synthesized with a 5' terminal KpnI restriction site. The G/A\textsubscript{-22018} region (-22133 to -21910) was PCR-amplified from Caco-2 cell genomic DNA to generate the A variant region fragment. The G/A\textsubscript{-22018} region was amplified from RP11-34L23 BAC clone DNA to yield the G variant region fragment.

Subcloning of the lactase variant–promoter–reporter constructs

C/T\textsubscript{-13910} and G/A\textsubscript{-22018} variant region promoter–reporter constructs were generated by cloning the PCR-amplified C/T\textsubscript{-13910} or G/A\textsubscript{-22018} variant region PCR products upstream of the lactase promoter in the previously described p3kLac construct (11). The p3kLac construct contains a 3.0 kb fragment of the rat lactase promoter cloned upstream of the luciferase cDNA in the reporter plasmid pGL3Basic (Promega). Specifically, the internal 218 bp MluI fragment of the C/T\textsubscript{-13910} region PCR products was cloned into p3kLac to generate p3kLac-C and p3kLac-T (forward orientation) and p3kLac-C' and p3kLac-T' (reverse orientation). Similarly, the 224 bp KpnI fragment of the G/A\textsubscript{-22018} region PCR products was cloned into p3kLac to generate p3kLac-G, p3kLac-G', p3kLac-A and p3kLac-A'. The combination G/A\textsubscript{-22018} C/T\textsubscript{-13910} variant region reporter constructs were generated by cloning the C/T\textsubscript{-13910} region MluI fragments into p3kLac-G and p3kLac-A downstream of the G/A\textsubscript{-22018} region to yield p3kLac-G-C and p3kLac-A-T. Incorporation of the C/T\textsubscript{-13910} and G/A\textsubscript{-22018} variant regions was confirmed by sequencing for all constructs generated.

Transient transfection assays

Caco-2 cells were cultured in Dulbecco’s modified Eagle’s medium (DMEM) with 10% fetal bovine serum. Forty-eight hours prior to transfection the cells were split and 35 mm dishes were seeded with 2 x 10\textsuperscript{5} cells. For each reporter construct, a DNA transfection mixture was prepared consisting of 0.4 pmol of the reporter construct, 0.1 µg of pRL-CMV (Promega) as an internal control, and pBluescript KS+ II to adjust to a final of 2.1 µg total DNA. The individual DNA mixtures were transfected in triplicate wells into cells (50–80% confluent) with Lipofectamine Reagent (BRL) according to the protocol of the manufacturer. For AP2 co-transfection experiments,
coligonucleotide probe alone (lanes 3 and 4) or in the presence of 100/2C2-labeled specific T

TM

protein (8.0

luciferase activity (means

promoter–reporter activities were normalized to the activity of

measurement of both reporter enzymes. Experimental lactase

pRL-CMV control) allowed for simultaneous expression and

with the dual reporters (firefly luciferase for the lactase

manufacturer, in a Monolight 3010 luminometer. Transfection

Reporter Assay System (Promega) as described by the

Extract Kit from Active Motif (Carlsbad, CA, USA). Caco-2 cell

and 7 days post confluency were extracted using the Nuclear

Electrophoretic mobility shift assays

harvested 48 h following transfection (70–90% confluent) and

with 0.4 pmol of the luciferase reporter constructs. Cells were

were detected using the LightShift Chemiluminescent EMSA

Kit (Pierce Biotechnology, Rockford, IL, USA) and 3' end-
biotinylated probes. Specifically, nuclear extracts (2 µg) were

incubated with 20 fmol double-stranded, biotinylated DNA probes

and 500 ng of poly(deoxyinosinate-deoxycytidylate) in a buffer

containing 10 mM Tris, 50 mM KCl, 1 mM DTT, 0.05% NP-40 and

5 mM MgCl2 in a total volume of 20 µl for 25 min at room

temperature. Complexes were separated in a 5% native poly-

acrylamide gel in 0.5 x Tris-borate-EDTA buffer, transferred to a

nylon membrane and cross-linked at 120 mJ/cm2. Detection of the

complexes was per the manufacturer’s instructions. Competition

EMSAs used 200-fold molar excess of unlabeled oligonucleotide

or a non-specific competitor oligonucleotide, Sp1, 5'-GATC

GGGGCGGGGCGGGGCGATC-3'. EMSAs were also performed with corresponding 32P labeled fill-in oligonucleotides

generated as previously described (11) and visualized by

autoradiography.

ACKNOWLEDGEMENTS

This work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases grants DK60074 and DK60715.

REFERENCES


