Published in final edited form as:

Hum Genet. 2012 July ; 131(7): 1153–1159. doi:10.1007/s00439-012-1140-z.

The human lactase persistence-associated SNP –13910*T enables *in vivo* functional persistence of lactase promoter-reporter transgene expression

Lin Fang, Jong Kun Ahn, Dariusz Wodziak, and Eric Sibley

Department of Pediatrics (Gastroenterology), Stanford University School of Medicine, Stanford, CA 94305

Abstract

Lactase is the intestinal enzyme responsible for digestion of the milk sugar lactose. Lactase gene expression declines dramatically upon weaning in mammals and during early childhood in humans (lactase nonpersistence). In various ethnic groups, however, lactase persists in high levels throughout adulthood (lactase persistence). Genetic association studies have identified that lactase persistence in Northern Europeans is strongly associated with a single nucleotide polymorphism (SNP) located 14 kb upstream of the lactase gene: –13910*C/T. To determine whether the –13910*T SNP can function *in vivo* to mediate lactase persistence, we generated transgenic mice harboring human DNA fragments with the –13910*T SNP or the ancestral –13910*C SNP cloned upstream of a 2 kb rat lactase gene promoter in a luciferase reporter construct. We previously reported that the 2 kb rat lactase promoter directs a post-weaning decline of luciferase transgene expression similar to that of the endogenous lactase gene. In the present study, the post-weaning decline directed by the rat lactase promoter is impeded by addition of the –13910*T SNP human DNA fragment, but not by addition of the –13910*C ancestral SNP fragment. Persistence of transgene expression associated with the –13910*T SNP represents the first *in vivo* data in support of a functional role for the –13910*T SNP in mediating the human lactase persistence phenotype.

Keywords

lactase; gene regulation; SNP; transgenic mice

INTRODUCTION

Lactase is a digestive enzyme present on the apical membrane of intestinal epithelial cells of nearly all mammals and is responsible for hydrolyzing the lactose disaccharide in milk to yield glucose and galactose monosaccharides for mucosal absorption. Lactase gene expression is spatially restricted to the proximal and middle portions of the small intestine. Temporally lactase levels peak in newborn mammals and decline dramatically post weaning during maturation (Buller et al. 1990; Freund et al. 1990; Krasinski et al. 1994; Krasinski et al. 1997; Lee et al. 2002). In a majority of humans, the lactase maturational decline occurs during childhood (lactase nonpersistence) (Buller and Grand 1990; Fajardo et al. 1994; Sebastio et al. 1989). In various human populations, however, lactase persists in high levels through adulthood (lactase persistence) (Enattah et al. 2008; Enattah et al. 2002; Ingram et

Correspondence: Eric Sibley, M.D. Ph.D., Department of Pediatrics (Gastroenterology), Stanford University School of Medicine, 300 Pasteur Drive, Grant Building G-310, Stanford, CA 94305-5208, Telephone: (650) 723-5070, Fax: (650) 724-3106, eric.sibley@stanford.edu.

al. 2007; Swallow 2003; Tishkoff et al. 2007). Lactase persistence is a heritable autosomal dominant condition and has been strongly correlated with several single nucleotide polymorphisms (SNPs) located 14 kb upstream of the lactase gene in different ethnic populations: –13910*T in Europeans and –13907*G, –13915*G, and –14010*C in several African and Arabian ethnic pastoral populations (Enattah et al. 2008; Enattah et al. 2002; Ingram et al. 2007; Swallow 2003; Tishkoff et al. 2007).

In vitro studies have shown that the –14 kb region DNA sequence corresponding to the lactase persistence SNPs can enhance transcriptional activation of the lactase promoter compared to the ancestral sequence in cell culture (Jensen et al.; Lewinsky et al. 2005; Olds et al.; Olds and Sibley 2003; Troelsen et al. 2003). The transcription factor Oct-1 has been shown to bind the –13907*G, –13910*T and –13915*G SNP region sequences with greater affinity than the ancestral sequence (Ingram et al. 2007; Jensen et al.; Lewinsky et al. 2005; Olds et al.). Due to the inability to study maturational changes in cell culture and *in vitro*, however, it remains to be determined whether and how the lactase persistence-associated SNPs affect lactase gene expression during development and aging in living animals.

To address these questions, we sought to generate transgenic mice carrying a luciferase reporter gene driven by a rat 2 kb lactase gene promoter fused with fragments of human DNA sequence corresponding to the lactase persistence-associated –13910*T SNP or the ancestral nonpersistence-associated –13910*C SNP. In a previous report, we established a transgenic mouse line carrying a luciferase reporter gene driven by the 2 kb rat lactase promoter (Lee et al. 2002). The 2 kb promoter is sufficient to drive *in vivo* luciferase transgene expression exclusively in intestinal epithelial cells with maximal expression in the proximal and middle sections of the small intestine. Temporally, the 2 kb promoter drives high-level transgene expression in newborn pups followed by a sharp post-weaning decline in adults similar to the endogenous lactase gene. In the present study, we therefore examined how addition of human lactase gene fragments carrying the lactase persistence –13910*T SNP or the ancestral –13910*C SNP affects the post-weaning decline mediated by the rat lactase promoter.

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.

Subcloning of the lactase SNP region-promoter-reporter transgene constructs

DNA fragments corresponding to the nucleotide sequence surrounding the –13910*C/T SNP region of the human lactase gene were generated by PCR amplification as previously described (Olds and Sibley 2003). Specifically, 218 bp fragments of the human lactase gene –13910*C/T SNP region were PCR-amplified using a forward oligonucleotide corresponding to nt –14017 to –13994, 5'-AGACGTAAGTTACCATTTAATAC-3', and a reverse oligonucleotide corresponding to nt –13800 to –13821, 5'-CGTTAATACCCACTGACCTATC-3'. Both primers were synthesized with a 5' terminal MluI restriction site for subsequent cloning. The –13910*C/T SNP region (–14017 to –13800) was PCR-amplified from Caco-2 cell genomic DNA to generate the –13910*T SNP region fragment. Similarly, the –13910*C/T SNP region was amplified from RP11-34L23, a human genomic DNA BAC clone (BACPAC Resources), to yield the –13910*C SNP region fragment. The –13910*C/T SNP region lactase promoter-reporter

constructs were generated by cloning the PCR-amplified -13910*C/T SNP region PCR products upstream of the lactase promoter in the previously described gLac2.0k construct (Lee et al. 2002). The gLac2.0k construct contains a 2.0 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 -13910*C and -13910*T SNP region PCR products was cloned into gLac2.0k to generate pLacLuc-2kC and p2kLacLuc-2kT respectively. Incorporation of the -13910*C/T SNP regions was confirmed by sequencing the 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 35mm dishes were seeded with 2×10^5 cells. For each reporter construct, a DNA transfection mixture was prepared consisting of 0.4pmol of the reporter construct, 0.1µg 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 Caco-2 cells (70–80% confluent) with Lipofectamine Reagent (BRL) according to the protocol of the manufacturer. Cells were harvested 48h following transfection (90% confluent) and luciferase activity was measured by the Dual-Luciferase™ Reporter Assay System (Promega) as described by the manufacturer, in a Monolight 3010 luminometer (BD Pharmingen). Transfection with the dual reporters (firefly luciferase for the lactase promoter-reporter plasmids and renilla luciferase for the pRL-CMV control) allowed for simultaneous expression and measurement of both reporter enzymes. Experimental lactase promoter-reporter activities were normalized to the activity of the pRL-CMV internal control and expressed as relative luciferase activity (means±SD). Statistical significance (P-value) was determined using Student's unpaired t-test.

Production and identification of transgenic mice

The plasmids pLacLuc-2kC and pLacLuc-2kT were linearized by digestion with XmnI and purified transgene DNA was injected into fertilized mouse eggs, that were reimplanted into pseudopregnant mice, and allowed to develop to term. Transgenic animals were generated through the Transgenic Core facility of the Stanford University School of Medicine. Mice carrying transgenes were identified by PCR amplification of genomic DNA using oligonucleotide primers specific for the rat lactase promoter (Verhave et al. 1995) (GenBank™ accession number S77839) and the firefly luciferase reporter gene (GenBank™ accession number U47295): glac (sense), 5'-TATCCTAGATAACCCAGTTAAA-3'; luc (antisense), 5'-CTTTATGTTTTTGGCGTCTTCC-3'. Transgenic lines were established by outbreeding founder mice.

Transgene copy number determination by quantitative PCR

Quantitative PCR was applied to determine the transgene copy numbers of the transgenic mouse lines. Primers for the *LacLuc* transgene were qLuc-F: 5'-GCCTGAAGTCTCTGATTAAGT-3' and qLuc-R: 5'-CGGGAAGACCTGCGACACCTG-3'. Primers for the internal control reference gene GAPDH were qGAPDH-F: 5'-GTCTCCTGCGACTTCAAC-3' and qGAPDH-R: 5'-TCATTGTCATACCAGGAAATGAGC-3'. Fermentas SYBR Green 2× Master Mix (K0221) was used and PCR amplifications were run on an Applied Biosystem Step One Plus Realtime PCR machine. To determine the copy number, the relative quantification by the comparative C_T method ($2^{-\Delta\Delta C_T}$) (Bubner and Baldwin 2004; Bubner et al. 2004) was adopted and modified to take into consideration different amplification efficiencies for the *LacLuc* transgenes and the internal control gene GAPDH. A previously characterized

transgenic line, *LacLuc-1.3kD*, with a known copy number ($N_{Calibrator}=11$) was used as the calibrator (Lee et al. 2002). For each transgenic line, four repeat PCR reactions were set up with 50 ng of genomic DNA as template. For the calibrator, five two-fold serial dilutions of genomic DNA were used to generate a standard curve and to determine amplification efficiency for the *LacLuc* transgenes and the internal control GAPDH gene. The C_T value, cycle number at which the amplification curve met the threshold line, was calculated by the ABI Step One Plus program. For the *LacLuc* and GAPDH genes, amplification efficiency E was calculated using the standard curve generated from the calibrator line. Linear regression analysis was performed with $\log_{10}(\text{fold dilution})$ as the X value and C_T as the Y value. The slope was retrieved and the efficiency of the reaction (E) was calculated as $E = 10^{(-1/\text{slope})} - 1$ (Bubner et al. 2004). The copy number of each *LacLuc* transgenic line N_{LacLuc} was then calculated using the formula: $N_{LacLuc} = N_{Calibrator} \times (1 + E_{LacLuc})^{(C_T \text{ Calibrator} - C_T \text{ LacLuc})} \times (1 + E_{GAPDH})^{(C_T \text{ GAPDH} - C_T \text{ GAPDH.Calibrator})}$ where N_{LacLuc} and $N_{Calibrator}$ are the copy numbers for the *LacLuc* lines (C, T1, T2) and calibrator line respectively; E_{LacLuc} and E_{GAPDH} are the amplification efficiency of *LacLuc* and GAPDH respectively; $C_T \text{ LacLuc}$, $C_T \text{ GAPDH}$, $C_T \text{ Calibrator}$ and $C_T \text{ GAPDH.Calibrator}$ are the C_T values of *LacLuc* and GAPDH in the *LacLuc* and calibrator lines respectively.

Luciferase enzymatic assay

The small intestine was harvested from transgenic mice and assayed for luciferase activity using the Luciferase™ Reporter Assay System (Promega, Madison, WI) as previously described (Lee et al. 2002). Harvested small intestinal tissue (~50 mg) was homogenized in cell lysis buffer (500 µl), and lysate protein concentration was determined. Lysate aliquots were incubated in the presence of luciferin substrate as described by the reporter assay system, and luciferase activity was quantified in a Monolight 3010 luminometer (BD PharMingen). The results are reported as relative light units per µg of total protein.

In vivo bioluminescence imaging

Transgenic mice were anesthetized using isoflurane and an aqueous solution of the substrate D-luciferin (150 mg/kg) was injected into the peritoneal cavity 5 min before imaging. The mice were then placed in a light-tight chamber and imaged with an intensified charge-coupled device (ICCD) camera (IVIS; Xenogen). Photons emitted from luciferase activity within the mice and then transmitted through the tissue were collected for a period of 5 min. A pseudocolor image representing light intensity was generated and superimposed over a gray scale whole body reference image similar to the method described previously (Lee et al. 2002). LivingImage software (Xenogen) was used to collect and analyze the bioluminescent images generated.

RESULTS

Generation of lactase promoter-luciferase transgenic mice harboring DNA fragments corresponding to the -13910*T or -13910*C SNPs

To investigate the role of the -13910*C/T SNPs associated with lactase nonpersistence/persistence in mediating temporal gene regulation in transgenic mice, we initially generated DNA constructs comprised of the 2 kb rat lactase promoter fused with a 218 bp fragment of the human lactase gene corresponding to the -13910*C (ancestral) or -13910*T (persistence) SNPs cloned upstream of a luciferase reporter gene (Fig. 1a). Prior to use for transgenesis, we performed preliminary transfection studies in cell culture with the constructs. Intestinal Caco-2 cells were transfected with the lactase promoter-luciferase constructs and assayed 48 hours after transfection. Transfection with the lactase nonpersistence-associated ancestral -13910*C SNP construct, pLacLuc-2kC, results in a 1.67-fold induction of promoter activity driven by the 2 kb lactase promoter (Fig. 2).

Transfection with the lactase persistence-associated –13910*T SNP construct, pLacLuc-2kT, results in a 2.66-fold induction of promoter activity driven by the 2 kb lactase promoter. Consistent with reports with different constructs (Olds and Sibley 2003), both –13910*C/T region fragments are capable of enhancing promoter activity while the lactase persistence-associated –13910*T construct possesses a significantly greater activity compared to the nonpersistence-associated ancestral –13910*C construct ($P<0.0005$, pLacLuc-2kC versus pLacLuc-2kT).

Transgenic mice were generated harboring the *LacLuc-2kC* or *LacLuc-2kT* transgenes and were screened for transgene expression by *in vivo* bioluminescent detection (Fig. 1b) and by assaying for luciferase activity in intestine homogenates. Among the transgenic mice generated, one *LacLuc-2kC* line (C line) and two *LacLuc-2kT* lines (T1 and T2 lines) express the luciferase reporter gene. The number of transgenes integrated into the DNA of each founder mouse varies from 4 to 11 copies/cells (Table 1).

Spatial restriction of *LacLuc-2kC/T* transgene expression along the length of the intestine

To define the regional expression of the *LacLuc-2kC* and *LacLuc-2kT* transgenes in the small intestine, luciferase activity was analyzed in tissue segments harvested along the length of the intestine in 2 week old preweaned pups and 8 week old adult mice. Specifically, the small intestine from mice of the C, T1 and T2 lines was divided into equal one-eighth segments (P1–8) beginning with the proximal duodenum and extending to the distal ileum. In preweaned mice, for all three lines, luciferase activity was detected throughout the length of the intestine as shown in Fig. 3a–c, pups. The broad spatial pattern of transgene expression is similar to that of the endogenous lactase gene in neonates (Lee et al. 2002, Krasinski et al. 1994). In adult mice, for the *LacLuc-2kC* line C, maximal luciferase activity was detected in the distal duodenum and jejunum (P2 to P5 segments) with peak activity in the P2 and P3 segments as shown in Fig. 3a, adults. For the *LacLuc-2kT* line T1, maximal luciferase activity was detected in the P2 and P3 segments (Fig. 3b, adults). For the *LacLuc-2kT* line T2, maximal luciferase activity was detected in the P2 to P4 segments with peak activity in the P3 segment (Fig. 3c, adults). Of note, the level of luciferase expression for the *LacLuc-2kC* line is much lower than the level of expression for the *LacLuc-2kT* lines T1 and T2 due possibly to integration site or copy number effects.

LacLuc-2kC mediates post weaning decline of transgene expression while *LacLuc-2kT* mediates persistent expression

Expression of the endogenous lactase gene declines dramatically post weaning (~3 weeks) in mice. We have previously shown that the 2 kb rat lactase promoter region is capable of directing a similar post-weaning decline in mice harboring the *LacLuc-2k* transgene (Lee et al. 2002). To determine how the human –14 kb region containing the –13910*C/T SNPs affects the temporal decline, the proximal-jejunum (P3) of 2 week old pups and 2 month old adult transgenic mice was harvested and transgene expression was quantified by measuring luciferase activity (Fig. 4).

In the *LacLuc-2kC* line C, luciferase activity declines ~16 fold in adult mice ($p=0.0005$). The post-weaning decline of transgene expression mediated by the *LacLuc-2kC* promoter is similar to the temporal pattern of expression mediated by the endogenous lactase gene promoter and by the 2 kb rat lactase promoter in the previously described *LacLuc-2k* line (5). By contrast, in the *LacLuc-2kT* line T1, luciferase activity is similar in pups and adults. In the T2 line, luciferase activity in adults is increased ~1.5 fold compared to pups ($p=0.025$). The persistent luciferase expression mediated by the *LacLuc-2kT* transgene in the two independent lines is consistent with the single nucleotide switch at –13910 from C

to T functioning to abolish the dramatic post-weaning decline and allow for persistent expression.

DISCUSSION

Although a strong association has been established between Northern European lactase persistence/nonpersistence and the $-13910^*C/T$ SNPs, there have been no *in vivo* studies to address the role of the SNPs in mediating temporal regulation of lactase gene expression. In our transgenic mouse study, the -14 kb region containing the lactase nonpersistence-associated -13910^*C SNP mediates a post-weaning decline of reporter gene expression in the *LacLuc-2kC* line. By contrast, the same region containing the lactase persistence-associated -13910^*T SNP allows for persistence of transgene expression in the two independent *LacLuc-2kT* lines. Our *in vivo* results support a causal role for the $-13910^*C/T$ SNPs identified from genetic association studies in determining lactase gene temporal regulation in humans and thereby specify the lactase persistence/nonpersistence phenotype.

The precise mechanism by which the -13910^*T SNP functions to mediate persistent expression of the human lactase gene remains to be determined. Potential mechanisms might involve a transcriptional repressor that binds to the -14 kb DNA region ancestral sequence or epigenetic modification of the region upon maturation that functions to mediate a decline in lactase gene expression. The -13910^*T SNP might function to prevent binding of the repressor to the region or prevent epigenetic modification thereby allowing for persistent gene expression. *In vitro* studies have shown that the -14 kb region DNA sequence corresponding to the lactase persistence SNPs can enhance transcriptional activation of the lactase promoter compared to the ancestral sequence in cell culture (Jensen et al. 2011; Lewinsky et al. 2005; Olds et al. 2011; Olds and Sibley 2003; Troelsen et al. 2003). In addition, previous reports have indicated that the SNP sequences can interact differentially with intestinal cell nuclear proteins and with the transcription factor Oct-1 (Ingram et al. 2007; Jensen et al. 2011; Lewinsky et al. 2005; Olds et al. 2011; Olds and Sibley 2003; Troelsen et al. 2003) that is capable of stimulating lactase promoter activity in cell culture (Ingram et al. 2007; Lewinsky et al. 2005; Olds et al. 2011). Binding sites for GATA, HNF4alpha, Cdx-2 and Fox transcription factors have also been identified in the -13910 SNP region (Lewinsky et al. 2005).

In addition to mediating distinctly different temporal expression patterns, mouse lines generated from *LacLuc-2kC* and *LacLuc-2kT* lines have distinct spatial transgene expression patterns. Consistent with the broad intestinal spatial expression pattern of the endogenous lactase gene in newborn mice that narrows to the proximal and middle sections of the small intestine upon maturation, the *LacLuc-2kC* and *LacLuc-2kT* lines similarly display a broad expression pattern in pups that narrows upon maturation in adults. Transgene expression for the C, T1 and T2 lines is restricted, however, to a more anterior spatial pattern with localization most proximal for the T1 and T2 lines. We conclude that addition of the human -14 kb region containing either -13910^*C or -13910^*T does not affect gross targeting of transgene expression to the small intestine but does influence fine patterning of spatial restriction along the anterior-posterior gut axis.

We have previously demonstrated that the *LacLuc-2k* lactase transgene is accurately expressed in enterocytes of the small intestine (Lee et al., 2002). While similar enterocyte expression is anticipated for the *LacLuc-2kC* and *LacLuc-2kT* transgenes, it is possible that addition of the human $-13910^*C/T$ fragments may result in ectopic transgene expression. Transgene expression may also be affected by chromosomal integration or copy number effects. Lack of a significant difference in copy number between the C, T1 and T2 lines, however, argues against a copy number effect to account for their different developmental

expression patterns. Specifically, the copy numbers of the C and T2 lines are similar (4.4 ± 1.7 and 5.3 ± 2.2 respectively) while their temporal expression patterns upon maturation are dramatically different.

In addition to providing *in vivo* support for the role of the -13910^*T and -13910^*C SNPs in functioning to mediate lactase persistence and nonpersistence respectively, the *LacLuc-2kT* and *LacLuc-2kC* transgenic lines provide model systems to study mechanisms involved in spatiotemporal gene expression during development and aging that is not possible in mammalian cell culture and *in vitro*.

Acknowledgments

The authors would like to thank Lynne Olds, Lydia Tsai, Christine Oo, Chris Brian Jackson, Alpna Kak and Charalambos Maravelias for excellent technical support and animal husbandry. This work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases Grants DK60715, DK72416 and DK56339.

REFERENCES

- Bubner B, Baldwin IT. Use of real-time PCR for determining copy number and zygosity in transgenic plants. *Plant Cell Rep.* 2004; 23:263–271. [PubMed: 15368076]
- Bubner B, Gase K, Baldwin IT. Two-fold differences are the detection limit for determining transgene copy numbers in plants by real-time PCR. *BMC Biotechnol.* 2004; 4:14. [PubMed: 15251044]
- Buller HA, Grand RJ. Lactose intolerance. *Annu Rev Med.* 1990; 41:141–148. [PubMed: 2109963]
- Buller HA, Kothe MJ, Goldman DA, Grubman SA, Sasak WV, Matsudaira PT, Montgomery RK, Grand RJ. Coordinate expression of lactase-phlorizin hydrolase mRNA and enzyme levels in rat intestine during development. *J Biol Chem.* 1990; 265:6978–6983. [PubMed: 1691182]
- 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. [PubMed: 18179885]
- Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L, Jarvela I. Identification of a variant associated with adult-type hypolactasia. *Nat Genet.* 2002; 30:233–237. [PubMed: 11788828]
- Fajardo O, Naim HY, Lacey SW. The polymorphic expression of lactase in adults is regulated at the messenger RNA level. *Gastroenterology.* 1994; 106:1233–1241. [PubMed: 8174885]
- Freund JN, Torp N, Duluc I, Foltzer-Jourdainne C, Danielsen M, Raul F. Comparative expression of the mRNA for three intestinal hydrolases during postnatal development in the rat. *Cell Mol Biol.* 1990; 36:729–736. [PubMed: 2128044]
- 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; 120:779–788. [PubMed: 17120047]
- Jensen TG, Liebert A, Lewinsky R, Swallow DM, Olsen J, Troelsen JT. The -14010^*C variant associated with lactase persistence is located between an Oct-1 and HNF1alpha binding site and increases lactase promoter activity. *Hum Genet.* 2011
- Krasinski SD, Estrada G, Yeh KY, Yeh M, Traber PG, Rings EH, Buller HA, Verhave M, Montgomery RK, Grand RJ. Transcriptional regulation of intestinal hydrolase biosynthesis during postnatal development in rats. *Am J Physiol.* 1994; 267:G584–G594. [PubMed: 7943323]
- Krasinski SD, Upchurch BH, Irons SJ, June RM, Mishra K, Grand RJ, Verhave M. Rat lactase-phlorizin hydrolase/human growth hormone transgene is expressed on small intestinal villi in transgenic mice. *Gastroenterology.* 1997; 113:844–855. [PubMed: 9287976]
- Lee SY, Wang Z, Lin CK, Contag CH, Olds LC, Cooper AD, Sibley E. Regulation of intestine-specific spatiotemporal expression by the rat lactase promoter. *J Biol Chem.* 2002; 277:13099–13105. [PubMed: 11812796]

- Lewinsky RH, Jensen TG, Moller J, Stensballe A, Olsen J, Troelsen JT. T-13910 DNA variant associated with lactase persistence interacts with Oct-1 and stimulates lactase promoter activity in vitro. *Hum Mol Genet.* 2005; 14:3945–3953. [PubMed: 16301215]
- Olds LC, Ahn JK, Sibley E. 13915*G. DNA polymorphism associated with lactase persistence in Africa interacts with Oct-1. *Hum Genet.* 2011; 129:111–113. [PubMed: 20960210]
- Olds LC, Sibley E. Lactase persistence DNA variant enhances lactase promoter activity in vitro: functional role as a cis regulatory element. *Hum Mol Genet.* 2003; 12:2333–2340. [PubMed: 12915462]
- Sebastio G, Villa M, Sartorio R, Guzzetta V, Poggi V, Auricchio S, Boll W, Mantei N, Semenza G. Control of lactase in human adult-type hypolactasia and in weaning rabbits and rats. *Am J Hum Genet.* 1989; 45:489–497. [PubMed: 2485006]
- Swallow DM. Genetics of lactase persistence and lactose intolerance. *Annu Rev Genet.* 2003; 37:197–219. [PubMed: 14616060]
- 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. [PubMed: 17159977]
- Troelsen JT, Olsen J, Moller J, Sjoström H. An upstream polymorphism associated with lactase persistence has increased enhancer activity. *Gastroenterology.* 2003; 125:1686–1694. [PubMed: 14724821]
- Verhave M, Krasinski SD, Maas SM, Smiers FJ, Mishra K, Breeuwsma NG, Naim HY, Grand RJ. Further characterization of the 5'-flanking region of the rat lactase-phlorizin hydrolase gene. *Biochem Biophys Res Commun.* 1995; 209:989–995. [PubMed: 7537502]

Figure 1. Transgenic constructs and founder mice

Transgenic mice were established with the luciferase reporter constructs shown in the schematic (a). The 2 kb rat lactase promoter was fused with a 218 bp human gene sequence –14 kb upstream of the lactase gene harboring the –13910*C or –13910*T SNPs and then cloned upstream of the firefly luciferase reporter gene in pGL3-Basic to generate LacLuc-2kC and LacLuc-2kT respectively. (b) To screen for pups carrying the transgenes, luciferin was injected intraperitoneally into 7 day old pups, and light emission was detected with an intensified charge-coupled device camera. *LacLuc-2kC* line C transgenic (PCR+) mice that emit light localized to the abdomen are shown in *blue and red*. The non-transgenic

(PCR-) F₁ littermates do not express luciferase and do not emit light. *Blue* illustrates the least intense and *red* the most intense light emission signals.

Figure 2. Luciferase activity of *LacLuc-2kC* and *LacLuc-2kT* transgenes in cell culture

Caco-2 cells were transfected with pGL3, promoterless luciferase reporter construct, or with lactase promoter-luciferase reporter constructs and assayed for luciferase expression. pLacLuc-2k, 2.0 kb of the rat lactase promoter cloned in pGL3. pLacLuc-2kC and pLacLuc-2kT, human lactase gene -13910*C/T SNP region harboring a C or T corresponding to nucleotide -13910 cloned in pLacLuc-2k. Transfection efficiencies were normalized to renilla luciferase expression from a co-transfected pRL-CMV vector and expressed as relative luciferase activity (means $\pm$ SD, $n=3$). * $P<0.05$; ** $P<0.05$; *** $P<0.05$.

Figure 3. Spatial restriction of *LacLuc-2kC* and *LacLuc-2kT* transgene expression in the small intestine of 2 week old pups and adults

Tissue along the length of the gastrointestinal tract was harvested from *LacLuc-2kC* line C1 (a) and *LacLuc-2kT* line T1 (b) and T2 (c) mice at 2 weeks of age (pups) and 2 months of age (adults). The small intestine was divided into equal one-eighth segments (P1–8) beginning with the proximal duodenum and extending to the distal ileum. Luciferase activity is expressed as arbitrary units per ug of total protein (means $\pm$ S.D., $n \geq 3$).

Figure 4. Temporal expression the *LacLuc-2kC* and *LacLuc-2kT* transgenes during gut maturation

Proximal-jejunum (segment P3) was harvested from *LacLuc-2kC* line C and *LacLuc-2kT* line T1 and T2 transgenic littermates sacrificed at 2 weeks (pups) and 2 months (adults) of age. Luciferase activity per ug of protein is normalized to average activity at 2 weeks for each line. For line C, n=6 pups and 6 adults. For line T1, n=6 pups and 4 adults. For line T2, n=13 pups and 13 adults. * p=0.0005

Table 1 Copy number *LacLuc-2kC* and *LacLuc-2kT* transgenic lines

	Line	Copy Number
<i>LacLuc-2kC</i>	C	4.4+/-1.7
<i>LacLuc-2kT</i>	T1	10.1+/-3.5
	T2	5.3+/-2.2