Mutations in the Translated Region of the Lactase Gene (LCT) Underlie Congenital Lactase Deficiency

Mikko Kuokkanen,1,2 Jorma Kokkonen,5 Nabil Sabri Enattah,1,2 Tero Ylisaukko-oja,1,2 Hanna Komu,2 Teppo Varilo,1,2 Leena Peltonen,1,2,6 Erkki Savilahti,3 and Irma Jarvela2,4

1Department of Molecular Medicine, National Public Health Institute, 2Department of Medical Genetics and 3Hospital for Children and Adolescents, University of Helsinki, and 4Helsinki University Central Hospital, Laboratory Services HUSLAB, Laboratory of Molecular Genetics, Helsinki; 5Department of Paediatrics, Oulu University Hospital, Oulu, Finland; and 6Broad Institute of MIT and Harvard, Boston

Congenital lactase deficiency (CLD) is a severe gastrointestinal disorder characterized by watery diarrhea in infants fed with breast milk or other lactose-containing formulas. We initially assigned the CLD locus by linkage and linkage disequilibrium on 2q21 in 19 Finnish families. Here we report the molecular background of CLD via characterization of five distinct mutations in the coding region of the lactase (LCT) gene. Twenty-seven patients out of 32 (84%) were homozygous for a nonsense mutation, c.4170T→A (Y1390X), designated “Fin major.” Four rare mutations—two that result in a predicted frameshift and early truncation at S1666fsX1722 and S218fsX224 and two point mutations that result in substitutions Q268H and G1363S of the 1,927-aa polypeptide—confirmed the lactase mutations as causative for CLD. These findings facilitate genetic testing in clinical practice and enable genetic counseling for this severe disease. Further, our data demonstrate that, in contrast to common adult-type hypolactasia (lactose intolerance) caused by a variant of the regulatory element, the severe infancy form represents the outcome of mutations affecting the structure of the protein inactivating the enzyme.
late to the major haplotype. On the basis of information about shared haplotypes, we concluded that a critical region would contain the LCT gene, and we proceeded to perform sequence analysis of this gene. The transcript map of the critical DNA region for CLD is shown in figure 2.

We identified five distinct mutations in the LCT gene in 32 patients from 24 families. Three mutations were predicted to lead to a premature truncation of lactase, and two were missense mutations that resulted in amino acid substitutions (table 1 and figs. 3 and 4). All the chromosomes with the major disease haplotype carried a nonsense mutation, c.4170T>A (Finmajor), resulting in Y1390X, a premature stop codon in exon 9 predicting the truncation of 537 amino acids. Finmajor was detected in 27 of 32 patients who were homozygous for the conserved founder haplotype or its modification. A total of 38 parents (5 did not participate) and 14 healthy siblings were found to be carriers of the Finmajor mutation. Two patients had a deletion of four nucleotides, c.4998_5001delTGAG, in their paternal disease chromosome in exon 14, leading to a frameshift and a premature stop codon after 55 altered amino acids (S1666fsX1722). These patients were compound heterozygous for Finmajor and S1666fsX1722. On the basis of genealogical studies, both of these mutations are known to originate from central and northern Finland. The third mutation, c.653_654delICT, is a deletion of two nucleotides in exon 2, predicting a frameshift change at codon 218 and protein truncation at codon 224, S218fsX224. One patient carried this mutation in her maternal disease chromosome. The fourth mutation is a c.804G>C transversion at codon 268, leading to an amino acid substitution of histidine for glutamine, Q268H, in the last nucleotide of exon 3. One patient carried this mutation in his maternal disease chromosome. The fifth mutation is a c.4087G>A transition resulting in a missense substitution of serine for an uncharged glycine, G1363S, at codon 1363 in exon 9. One patient carried this mutation in his maternal disease chromosome. A healthy sibling of the family also carried this rare mutation. These patients were compound heterozygous for the Finmajor and their respective rare mutations. On the basis of the birthplaces of the great-grandparents of the patients, mutations three, four, and five would be assumed to originate from eastern Finland.

The CLD mutations and the corresponding lactase activities measured from duodenal biopsy specimens are shown in table 1.

Ninety percent (43/48) of the disease chromosomes...
carried the Finmajor mutation. The carrier frequency of Finmajor was determined among 556 anonymous blood donors representing both the early and the later settlement regions of Finland. The highest carrier frequency of 1:35 (4/140) was seen in a little town, Nilsiä, in central Finland (fig. 5). This particular geographical region shows an enrichment of ancestors on genealogical studies (Järvelä et al. 1998). No carriers for other mutations (S218fsX224, S1666fsX1722, Q268H, and G1363S) were observed in any regional subpopulation screened.

The nascent lactase polypeptide contains four (I-IV) conserved structural and functional regions, but after posttranslational processing only two (III-IV) remain in the mature lactase enzyme (Mantei et al. 1988). The polypeptide has two catalytically active sites: phlorizin hydrolase activity is located in region III and lactase activity in region IV (Zecca et al. 1998; Arribas et al. 2000). Finmajor, Y1390X, truncates lactase polypeptide in the beginning of region IV. We employed an allele-specific minisequencing method to characterize the potential impact of the mutation at the steady-state transcript level of the LCT allele carrying the nucleotide change resulting in Y1390X (Rasinperä et al. 2005). We used a duodenal biopsy sample of a patient who was heterozygous for the nonsense mutation resulting in Y1390X and heterozygous for C/T [H1100213910], the SNP associated with adult-type hypolactasia (MIM 223100) (Enattah et al. 2002; Rasinperä et al. 2005). We were able to determine that the transcript level of the allele carrying the premature stop codon was at the same level as the transcript of the C/H11002 [H1100213910] allele, which typically represents ca 8% of the expressed LCT mRNA (Kuokkanen et al. 2003). We expect this to be caused by nonsense-mediated mRNA decay, a pathway which degrades mRNA induced by premature termination codons, eliminating the production of harmful truncated proteins.

Table 1

<table>
<thead>
<tr>
<th>DNA Mutation</th>
<th>Protein Mutation</th>
<th>Exon</th>
<th>Patient</th>
<th>Mean Lactase Protein Activity (U/g of Protein)</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.4170T→A</td>
<td>Y1390X</td>
<td>9</td>
<td>1–13, 15, 18–26, 28–31</td>
<td>2.1 (range 0–7)(^a)</td>
<td>Homozygote</td>
</tr>
<tr>
<td>c.4998_5001delTGAG</td>
<td>S1666fsX1722</td>
<td>14</td>
<td>17, 32</td>
<td>5 (range 3–7)(^b)</td>
<td>Compound heterozygote(^a)</td>
</tr>
<tr>
<td>c.653_654delCT</td>
<td>S218fsX224</td>
<td>2</td>
<td>16</td>
<td>2</td>
<td>Compound heterozygote(^a)</td>
</tr>
<tr>
<td>c.804G→C</td>
<td>Q268H</td>
<td>3</td>
<td>27</td>
<td>5</td>
<td>Compound heterozygote(^a)</td>
</tr>
<tr>
<td>c.4087G→A</td>
<td>G1363S</td>
<td>9</td>
<td>14</td>
<td>3</td>
<td>Compound heterozygote(^a)</td>
</tr>
</tbody>
</table>

\(^a\) Lactase activities of patients 11, 18, 20, 26, and 29 were not available.
\(^b\) Compound heterozygote for Y1390X.
Figure 3  CLD mutations in the LCT gene. A–E, DNA sequence chromatograms of five identified CLD mutations. One mutation, Y1390X, (A) is a homozygote, whereas the rest of the mutations are heterozygotes. The first row shows affected sequences and the second row wild-type sequences. In addition, c.655G→A, a SNP (rs3754689) leading to V219I, is shown in panel E (Boll et al. 1991). PCR was performed using the genomic DNA-amplifying promoter region, exons, flanking intron sequences, and 3′-UTRs of the LCT gene (primers are available on request). Sequencing was performed in both directions, and the sequenced products were electrophoresed on an ABI 3730 DNA analyzer (Applied Biosystems) in accordance with the manufacturer’s instructions and were analyzed using ABI Sequencing Analysis 3.3 (Applied Biosystems) and Sequencher 4.1 (Gene Codes). (Maquat 2005). This mechanism has been seen to modulate human disease phenotypes caused by disease-associated premature nonsense or frameshift mutations (Baserga and Benz 1988; Frischmeyer and Dietz 1999; Inoue et al. 2004). Furthermore, three of our CLD patients with the Y1390X mutation were included in an earlier study in which protein patterns of brush-border fragments were preliminarily examined (Freiburghaus et al. 1976). CLD patients had a complete or nearly complete absence of brush-border lactase in studied duodenal biopsies (Freiburghaus et al. 1976), which is in good agreement with our result that the mutant Y1390X LCT allele is down-regulated by nonsense-mediated decay. The second mutation, S1666fsX1722, truncates lactase in the middle of region IV, and the third mutation, S218fsX224, truncates polypeptide early, at region I. In theory, these mutant alleles may also be subjected to nonsense-mediated decay. The fourth mutation, Q268H, hits region I, substituting the basic histidine for an uncharged glutamine. Histidine has an imidazole ring that can be uncharged or positively charged, depending on its microenvironment. Although Q268H does not affect the mature enzyme, it may affect the structure of the lactase α-profragment, which has been demonstrated to act during lactase-folding processes as an intramolecular chaperone towards the mature enzyme (Naim et al. 1994; Jacob et al. 2002). The importance of the Q268 residue is also implied by cross-species conservation from human through chimpanzee, rabbit, cow, rat, and mouse. The fifth mutation, G1363S, replaces the smallest amino acid, glycine, with serine, which has a hydroxyl group that makes it polar and reactive. The replacement of the small glycine by the larger serine may alter the three-dimensional structure of the polypeptide. Since the G1363S mutation is located in the mature lactase at the end of region III, near the catalytically active sites, it may have serious functional consequences. Also, G1363 is conserved across species. The schematic structure of lactase polypeptide and the location of the five CLD mutations are illustrated in fig. 4. Opposite to the rare congenital lactase deficiency, adult-type hypolactasia (MIM 223100) is the most com-
Figure 4  Structure of lactase and the location of the identified CLD mutations. The genomic size of the \textit{LCT} gene is \(\sim 55\) kb and is composed of 17 exons (Boll et al. 1991). The size of messenger RNA (mRNA) is 6,274 bases, and the primary translation product (prolactase) is 1,927 amino acids. The prolactase contains a cleavable signal sequence from Met\(_1\) to Gly\(_{19}\), that guides the polypeptide to the endoplasmic reticulum (von Heijne 1986; Mantei et al. 1988). The region from Ser\(_{20}\) to Thr\(_{1882}\) consists of four homologous domains (I–IV). The prolactase is processed by two proteolytic cleavages. The first, an intracellular cleavage, occurs between Arg\(_{734}\) and Leu\(_{735}\). The second cleavage in the intestinal lumen between Arg\(_{868}\) and Ala\(_{869}\) generates lactase\(_{\text{b}}\), the mature enzyme (Jacob et al. 1996; Wüthrich et al. 1996). The figure is modified from Naim (2001).

Figure 5  Carrier frequency of Y1390X in four subpopulations in the early and later settlement regions of Finland. The carrier frequencies of the CLD mutations were determined from 556 anonymous blood donors obtained from the Finnish Red Cross Blood Transfusion Service.
Web Resources

URLs for data presented herein are as follows:

- Baylor College of Medicine (BCM) Search Launcher, http://searchlauncher bcm.tmc.edu/
- Marshfield Clinic Research Foundation Genetic Map, http://research.marshfieldclinic.org/
- UCSC Genome Browser, http://genome.ucsc.edu/

References