Assignment of the Locus for Congenital Lactase Deficiency to 2q21, in the Vicinity of but Separate from the Lactase-Phlorizin Hydrolase Gene

Irma Järvelä,[1,*] Nabil Sabri Enattah,[1,*] Jorma Kokkonen,[3] Teppo Varilo,[1] Erkki Savilahti,[2] and Leena Peltonen[1]

[1]Department of Human Molecular Genetics, National Public Health Institute, and Department of Medical Genetics, University of Helsinki, and [2]Hospital for Children and Adolescents, University of Helsinki, Helsinki; and [3]Department of Pediatrics, Oulu University Hospital, Oulu, Finland

Summary

Congenital lactase deficiency (CLD) is an autosomal recessive, gastrointestinal disorder characterized by watery diarrhea starting during the first 1–10 d of life, in infants fed lactose-containing milks. Since 1966, 42 patients have been diagnosed in Finland. CLD is the most severe form of lactase deficiency, with an almost total lack of lactase-phlorizin hydrolase (LPH) activity on jejunal biopsy. In adult-type hypolactasia, the most common genetic enzyme deficiency in humans, this enzyme activity is reduced to 5%–10%. Although the activity of intestinal LPH has been found to be greatly reduced in both forms, the molecular pathogenesis of lactase deficiencies is unknown. On the basis of the initial candidate-gene approach, we assigned the CLD locus to an 8-cM interval on chromosome 2q21 in 19 Finnish families. At the closest marker locus, a specific allele 2 was present in 92% of disease alleles. On the basis of a genealogical study, the CLD mutation was found to be enriched in sparsely populated eastern and northern Finland, because of a founder effect. The results of both the genealogical study and the haplotype analysis indicate that one major mutation in a novel gene causes CLD in the Finnish population. Consequently, the critical region could be restricted further, to an ~350-kb interval, by ancient-haplotype and linkage-disequilibrium analyses. Surprisingly, the LPH gene was shown to lie outside the critical CLD region, excluding it as a causative gene for CLD. The LPH locus was found to reside >2 Mb from the critical CLD region.

Received May 13, 1998; accepted for publication July 27, 1998; electronically published September 18, 1998.

Address for correspondence and reprints: Dr. Irma Järvelä, Department of Human Molecular Genetics, National Public Health Institute, Mannerheimintie 166, 00300 Helsinki, Finland. E-mail: Irma.Jarvela@ktl.fi

*These two authors contributed equally to the work.

0002-9297/98/6304-0021$02.00

Introduction

Congenital lactase deficiency (CLD; MIM 223000) is a severe gastrointestinal disorder in newborns that is inherited as an autosomal recessive trait. The first patients were described by Holzel et al. (1959). Thereafter, 16 patients have been described in Finland (Savilahti et al. 1983). To date, a total of 35 CLD families, with 42 patients, have been diagnosed in Finland. The estimated incidence of CLD is 1:60,000 newborns in the Finnish population. CLD is considered to belong to the so-called Finnish disease heritage, together with ~30 rare, monogenic disorders enriched in this population because of a founder effect and genetic drift (Nevanlinna 1972; Norio et al. 1973).

CLD is characterized by watery diarrhea starting rapidly, within a day of the introduction of lactose-containing infant feed, breast milk, or regular, cow's milk–based formulas. The severe diarrhea is caused by unabsorbed lactose in a high concentration (20–90 g/liter) and is followed by dehydration, acidosis, and weight loss, diagnosed at a mean age of 36 d. The lactase activity in a jejunal biopsy specimen from affected patients is reduced to 0–10 U/g protein. Jejunal morphology is normal. On a lactose-free diet the children are free of symptoms, and development is normal (Savilahti et al. 1983). The molecular basis of the nonfunctioning or absent lactase enzyme has remained unknown.

Owing to the significant decrease of lactase activity in CLD patients, we focused our initial linkage studies on chromosome 2q, where the lactase-phlorizin hydrolase (LPH) gene has been localized (Kruse et al. 1988; Harvey et al. 1993). In this article, we report assignment of the CLD gene to the long arm of chromosome 2 in 19 Finnish families. The localization of the CLD gene was refined further by extended haplotype analysis using seven polymorphic markers on chromosome 2q21, which facilitated the utilization of ancestral recombinations in the positioning of the disease locus. An obvious candidate, the LPH gene was excluded as a causative gene for CLD, on the basis of radiation-hybrid (RH) and physical mapping over the critical CLD region.

Subjects, Material, and Methods

CLD Families and DNA Samples

A total of 27 CLD patients, from 19 families, participated in the linkage study (fig. 1). The diagnosis of all patients was based on clinical symptoms. Estimation of lactase activity in the jejunal biopsy specimen was performed for all but one patient (Savilahti et al. 1983). DNA was extracted from frozen peripheral blood, in accordance with standard protocols (Vandenplas et al. 1984). Thirty-one families participated in the genealogical study. Three generations of ancestors were traced, on the basis of local church registries (Varilo et al. 1996*b*). Samples were taken in accordance with the Helsinki Declaration.

Analysis of Microsatellite Markers

We analyzed 10 highly polymorphic microsatellite markers, from the Généthon Resource Center, on 2q that are near the lactase gene (Dib et al. 1996); the genetic distances are as follows: cen-D2S114-1 cM-D2S1334-0 cM-D2S2196-0 cM-D2S442-2 cM-D2S314-2 cM-D2S2385-1 cM-D2S2288-1 cM-D2S397-1 cM-D2S150-1 cM-D2S132. The order of the markers was obtained mostly from the physical YAC contig map of chromosome 2 (Chumakov et al. 1995), supplemented with data from the Généthon map. PCR was performed in a total volume of 15 μl containing 12 ng template DNA, 5 pmol of the primers, 0.2 mM of each nucleotide, 20 mM TrisHCl (pH 8.8), 15 mM $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, 0.1% Tween 20, 0.01% gelatin, and 0.25 U *Taq* polymerase (Dynazyme, Finnzymes). One of the primers was radiolabeled, at the 5′ end, with $\gamma[^{32}P]$-ATP. The reactions were performed in a multiwell microtiter plate, for 35 cycles of denaturation at 94°C for 30 s, annealing at various temperatures, depending on the primers, for 30 s, and extension at 72°C for 30 s; denaturation was set at 3 min, and final extension was set at 5 min. The amplified fragments were separated on 6% polyacrylamide gels, and autoradiography was performed.

Linkage and Linkage-Disequilibrium Analyses

Pairwise LOD scores were calculated by use of the MLINK option of the LINKAGE program package (Lathrop et al. 1984). An autosomal recessive mode of inheritance (Nevanlinna 1972; Savilahti et al. 1983) with 100% penetrance and a disease-allele frequency of .001 was assumed. We considered locus heterogeneity to be highly unlikely, since CLD is one of the rare disorders enriched in the Finnish population.

Linkage-disequilibrium analyses were performed by use of the program HRRLAMB (Terwilliger 1995, 1996). This program applies a likelihood-ratio test for linkage disequilibrium, calculated as parameter λ, that is independent of the number of alleles and, by extension, of the number of marker loci over a small chro-

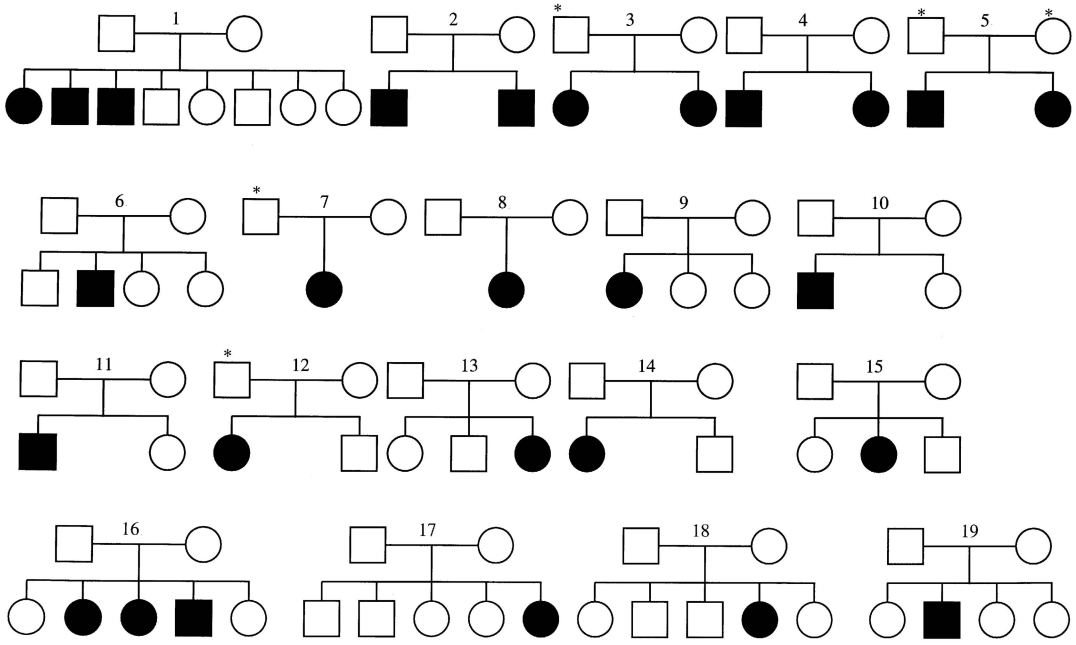

Figure 1 Finnish CLD families studied. Blackened symbols indicate affected individuals. An asterisk (*) indicates that no sample was available.

mosomal region. In addition, all genotyping data were analyzed by use of the HRRMULT program, which is designed for joint association analysis of multiple loci. In this method, the recombination fraction between any given map position and each of the marker loci is fixed, and the likelihood is maximized, at that map position, over α (proportion of disease alleles originally associated with a certain allele) and n (number of generations since introduction of the founder disease allele into the population).

RH Mapping

A high-resolution whole genome (TNG) RH panel (Research Genetics) comprising 90 human-hamster hybrid lines was analyzed together with a human lymphoblastoid cell line (RM; positive control) and nonirradiated hamster recipient cells (A3; negative control). Microsatellites D2S114, AFM338YE5, D2S442, D2S314, and D2S2385 (Dib et al. 1996) and two primer sets (from exons 1 and 17; Stanford Human Genome Center [SHGC] 10723) designed from cDNA of the LPH gene (Mantei et al. 1988) were amplified, by PCR, from each RH. PCRs were performed with 15 ng template, 60 pmol of the primers, 200 μM each dNTP, and 0.5 U *Taq* polymerase (Dynazyme, Finnzymes) in 50 μl buffer, under the conditions described above. The PCR products were visualized on 1.5% agarose gels. The screening results for the TNG panel were analyzed by use of the RH server at SHGC. Because of the high-resolution power of the panel, two markers must be closer together (~350 kb from each other) to show linkage; thus, the mapping results are based on two-point maximum-likelihood linkage for the marker screened with the SHGC G3 map of chromosome 2, resulting in a LOD score $\geqslant 6$. Because the mapping services did not provide assignments for markers D2S114 and D2S442 or the distance between them, the RHMAP statistical package for RH mapping (Cox et al. 1990; Boehnke et al. 1991) also was used.

Physical Mapping

The initial YAC contigs were assembled by CEPH/Généthon. YAC clones previously assigned to chromosome 2q were ordered from the Sanger Centre (Medical Research Council, United Kingdom). The presence of the known markers of the CLD locus and three primer pairs designed from exons 1, 8, and 17 (SHGC-10723) of the LPH cDNA (Mantei et al. 1988) were tested by PCR amplification. Similarly, a P1-derived artificial chromosome (PAC) library on 321 multiwell microtitration plates (kindly provided by Prof. Peter de Jong, Roswell Park Cancer Institute; Ioannou et al. 1994) was screened for the presence of the LPH gene and the five closest markers, D2S114, D2S442, AFMA338YE5, D2S314, and D2S2385. The PCR-positive clones were picked up and cultured in Luria broth supplemented with 25 μg kanamycin/ml, and DNA was extracted from these cultures, in accordance with the standard alkaline-lysis method (Sambrook et al. 1989). The PCR conditions were similar to those used in family studies and are described above.

Results

Distribution of CLD in the Finnish Population

The distribution of birthplaces of the great-grandparents of 31 Finnish CLD families shows that the CLD mutation is enriched in sparsely populated eastern and northern Finland (fig. 2), which was inhabited during the late settlement after the 16th century (Nevanlinna 1972; Norio et al. 1973).

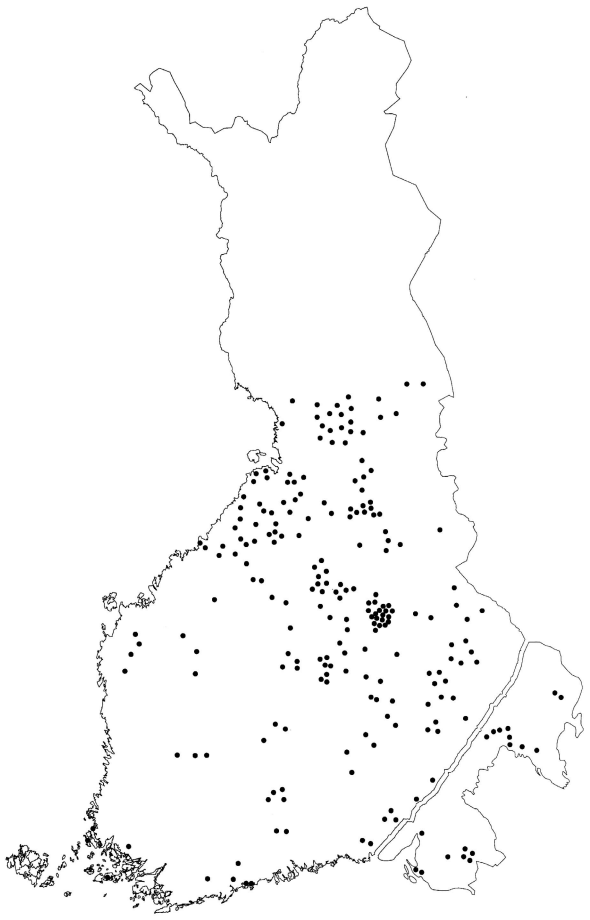

Figure 2 Birthplaces of great-grandparents of 31 CLD families in Finland. Families with more than one affected child have been marked only once.

Linkage Analyses

The segregation of a total of 10 highly informative microsatellite markers covering 9 cM flanking the LPH gene was analyzed in 19 CLD families. The LOD scores obtained from pairwise linkage analyses between CLD and markers on 2q21 are given in table 1. The highest LOD score, 7.93, was obtained with marker D2S2385. The LOD scores significant for linkage were also observed with seven markers, D2S1334, D2S2196, D2S442, D2S314, D2S397, D2S2288, and D2S150. Obligatory recombination events (both occurring in family 5) were detected with marker D2S114, which defines the centromeric boundary of the CLD locus, and D2S150, which defines the corresponding telomeric boundary of the CLD locus (table 1). Marker AFMA338YE5 was not polymorphic in our family material and was not included in the linkage analyses.

Ancient-Haplotype and Linkage-Disequilibrium Analysis

Figure 3 shows all 19 extended haplotypes formed by the seven linked markers spanning ~6 cM. The order of the markers was confirmed by RH mapping using the TNG RH panel (fig. 4) and by physical mapping using YACs assigned to this chromosomal region (fig. 5). The mutual order of D2S1334 and D2S2196 remained unknown. From 38 recent CLD chromosomes, 19 different haplotypes could be formed with seven microsatellite markers linked to the CLD gene. The suggested founder haplotype, cen-6-4-4-2-2-3-5-tel, was present in 13 (34%) of the disease chromosomes. The other haplotypes were separated from this putative founder haplotype by ancient recombination events. These ancient recombinations restricted the CLD locus to between markers D2S314 and D2S2385, within a 2-cM region. A single core haplotype, 2-2, formed by these markers

Marker	Haplotype																			λ
D2S1334	6	6	7	5	6	4	5	5	6	6	6	5	5	4	7	6	4	7	5	0.46
D2S2196	4	4	4	4	3	4	4	4	4	4	4	3	3	4	2	4	3	3	2	0.64
D2S442	4	4	4	4	4	4	4	4	4	4	3	4	4	3	5	4	3	3	3	0.78
D2S314	2	2	2	2	2	2	2	2	2	2	2	2	5	5	5	5	5	5	4	0.80
D2S2385	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	7	7	7	0.85
D2S2288	3	4	3	3	3	3	3	5	4	3	4	3	5	3	5	4	3	3		0.49
D2S397	5	5	5	5	5	3	4	5	2	5	5	2	5	1	5	5	4	2		0.75
Total	13	4	3	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	= 38

Figure 3 All CLD haplotypes and numbers of observations of the 38 recent disease chromosomes. The ancestral founder haplotype is shaded, and its identifiable tracers are boxed. On the basis of ancestral recombinations, the CLD locus could be restricted to between markers D2S314 and D2S2385. The proportion of excess of a certain allele, in the disease chromosomes, is expressed by the λ value, which measures the linkage disequilibrium between each marker and the disease gene.

was present on 82% (31/38) of the affected chromosomes and on none of the unaffected chromosomes. Allele 2 of D2S314 was present on 82% (31/38) of the affected chromosomes and on 6% (2/38) of the unaffected chromosomes. Allele 2 of D2S2385 was present on 92% (35/38) of the affected chromosomes and on 6% (2/38) of the unaffected chromosomes. Three parents (8% of all parents), from families 14, 19, and 21, carried allele 7 at D2S2385, and one parent from family 5 carried allele 1 at this locus. A total of seven markers, spanning an area >8 cM, were in linkage disequilibrium ($P \leqslant .001$) with the disease locus (table 1). The P and λ values resulting from the HRRLAMB program are shown in table 1 and figure 3, respectively. Data from all markers were combined by multipoint association analysis using the HRRMULT program (Terwilliger 1995, 1996). The likelihood curve peaked between markers D2S314 and D2S2385, with a maximum multipoint LOD score of 39.9, thus confirming the position of the CLD gene (fig. 6). The estimated α value was 1.00, indicating locus homogeneity.

RH Mapping

The SHGC mapping service was able to localize 5 of the 7 markers screened with the TNG panel of the SHGC G3 map of chromosome 2: LPH exons 1 and 17 (SHGC-10723) and marker AFMA338YE5 are linked to marker SHGC-30722; D2S314 is linked to marker SHGC-1625; and D2S2385 is linked to marker SHGC-20039. Thus, the marker order is as follows: cen-LPH-AFM338YE5-D2S314-D2S2385-tel. Analysis by means of the RHMAP statistical package resulted in the same marker order. On the RH map, the LPH gene was located between markers D2S114 and D2S442 (fig. 4). The distance between markers D2S314 and D2S2385, which

Table 1

Two-Point LOD Scores and Allelic Association Analyses for the Finnish CLD Families

Marker	LOD Score at Recombination Fraction of					
	.0	.1	.2	.3	.4	*P* Value
D2S114	$-\infty$	2.695	2.144	1.2115	.3633	7×10^{-7}
D2S1334	7.4111	5.2754	3.3366	1.6836	.4723	.006
D2S2196	4.9193	3.6076	2.3636	1.2583	.382	.001
D2S442	4.5603	3.2086	1.9951	.9856	.2706	.00001
D2S314	7.5608	5.5431	3.6256	1.9089	.5666	9×10^{-11}
D2S2385	7.9287	5.86	3.8469	2.0251	.5999	3×10^{-12}
D2S2288	4.629	3.3637	2.1511	1.0894	.3056	.028
D2S397	6.3058	4.6755	3.075	1.6085	.4669	.000001
D2S150	$-\infty$	4.9699	3.391	1.8271	.5486	7×10^{-7}
D2S132	$-\infty$	1.878	1.4583	.7909	.2282	.5

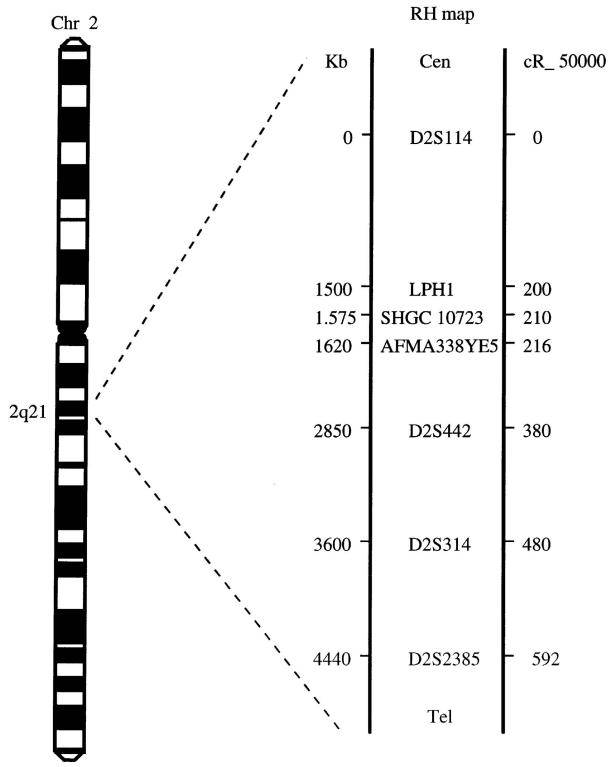

Figure 4 Graphic representation of RH-mapping results from the TNG panel. The LPH gene is shown to be located outside the critical CLD region on 2q21 (when 1 cR ≈ 7.5 kb is assumed; "50,000" is the x-ray dose used for the construction of the panel).

restricted the CLD locus, was calculated to be 112 cR, corresponding to ~840 kb (when an average relation of 1 cR ≈ 7.5 kb is assumed). On the basis of the RH-mapping data, the LPH gene was localized to 265 cR (~2 Mb) centromeric to D2S314 (fig. 4). The marker closest to the LPH gene, D2S442, was localized to 100 cR (~750 kb) centromeric to D2S314.

Physical Map of the CLD Region and Location of the LPH Gene

Figure 5 illustrates the genomic clones over the CLD critical region and their orientation to each other. A YAC contig was constructed with eight closely linked markers (D2S150, D2S397, D2S2288, D2S2385, D2S314, D2S442, D2S2196, and D2S1334), one uninformative CA repeat marker, AFMA338YE5, and the LPH gene and was extended ~4.5 Mb. The presence of LPH was analyzed by PCR amplification using three different primer pairs, covering exons 1, 8, and 17 (SHGC-10723) of the LPH cDNA (Mantei et al. 1988). Primers detecting exons 1 and 8 of the LPH cDNA were positive with YAC 944H9, which also contained markers D2S442 and D2S314. The primers covering exons 1, 8, and 17 (SHGC-10723) of the LPH cDNA were positive with YAC 743C2, which also contained two centromeric microsatellite markers, D2S2196 and D2S1334. An uninformative CA-repeat marker, AFMA338YE5, was positive with PACs 69K1, 71E19, and 93I13 and with YACs 743C2 and 944H9. Consequently, the LPH gene was assigned unequivocally to the centromeric end of YAC 944H9, between markers D2S442 and D2S2196/D2S1334, and definitively outside the core haplotype. Three YACs (850D5, 948D3, and 775E6) contained both of the closest markers, D2S314 and D2S2385. Two of these YACs (850D5 and 948D3) also were positive for markers D2S442 and D2S2288. On the basis of the published size of the YACs, the LPH gene should reside >2 Mb from the closest marker, D2S2385 (fig. 5).

Discussion

We describe the refined assignment of the CLD locus to human chromosome 2q by linkage, linkage-disequilibrium, and ancient-haplotype analyses. We also provide evidence that a novel gene in this region causes CLD, since we excluded the LPH gene as a causative gene for CLD, on the basis of RH-mapping results and the physical location of the LPH gene in the critical chromosomal region.

Adult-type hypolactasia (MIM 223100) is the most frequent genetic enzyme deficiency in humans. The frequency varies between different populations, from 2.6%–6.6% in Danes to 100% in healthy Thais (Sahi et al. 1972). CLD, however, is very rare. It is enriched in the Finnish population, with a total of 32 CLD families diagnosed so far, whereas only 18 cases have been reported outside Finland (Savilahti et al. 1983). The activity of lactase and its accompanying enzyme, phlorizin hydrolase, is reduced in adult-type hypolactasia to 5%–10% of that at birth. In 70% of CLD patients, LPH activity is reduced more, to 0%–2% of that at birth. In the rest of the CLD patients, LPH activity is 3%–10% of that at birth, which overlaps with the LPH-activity values for adult-type hypolactasia. In these borderline cases, it might be difficult to distinguish CLD from the more common adult-type hypolactasia. The age at onset of adult-type hypolactasia varies between different populations, starting at 2–4 years of age among Thais and at school age among Caucasians. In the Finnish population, the age at onset of adult-type hypolactasia varies within the range 5–20 years (Sahi et al. 1972). The gastrointestinal symptoms of adult-type hypolactasia usually are mild or not present, whereas, without proper treatment, CLD often would be lethal: the diarrhea is so severe that 90% of infants remain below their birth weight as long as they eat lactose-containing milks. All but one affected child in this study was hospitalized as a newborn, and the diagnosis of CLD was based on

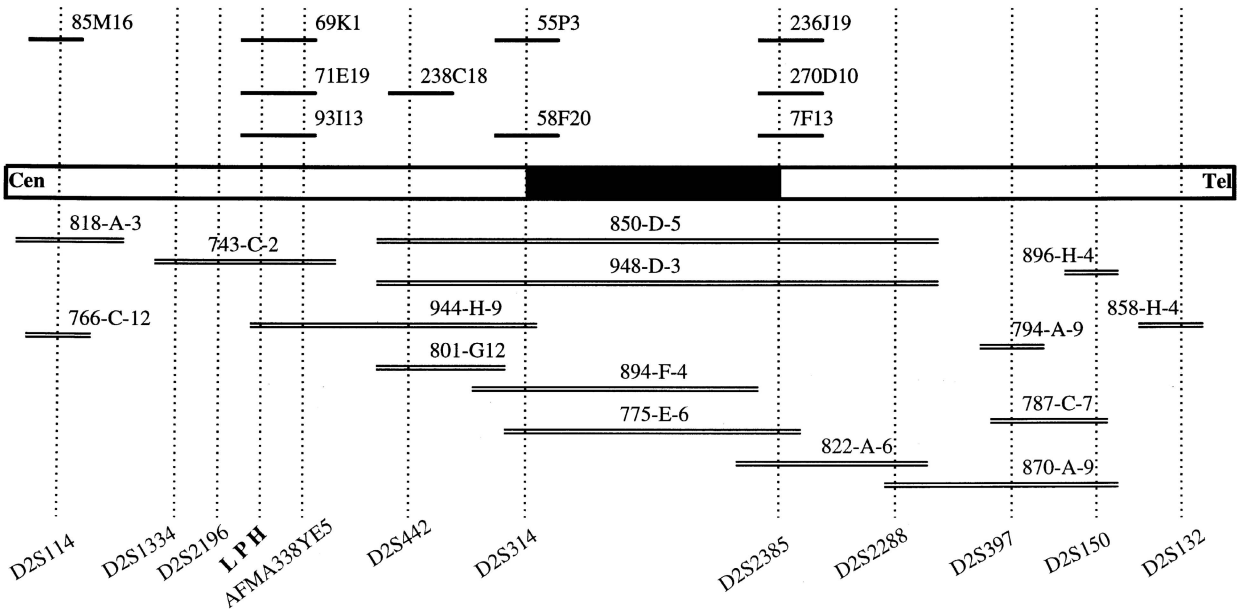

Figure 5 Physical map of the CLD locus on 2q21. The PAC clones are shown above the horizontal axis, and the YAC clones are shown below it. Vertical lines indicate the microsatellites present in the clones. Old recombinations detectable in extended disease haplotypes restricted the CLD region to between D2S314 and D2S2385 (*blackened rectangle*). The primers used to detect the LPH-positive PAC and YAC clones were nt 1640–1660 and nt 1770–1751 for exon 1, nt 2540–2560 and nt 2660–2641 for exon 8, and human STS SHGC-10723 for exon 17 of the LPH cDNA (Mantei et al. 1988).

determination of lactase activity in their jejunal biopsy specimens (Savilahti et al. 1983). One patient suffered from continuous diarrhea and was studied at the hospital at 4 mo of age, because of poor weight gain. Careful analysis of clinical phenotype should minimize the genetic heterogeneity in our family material.

The LPH gene had been assigned to 2q21 (Kruse et al. 1988; Harvey et al. 1993), which guided us to analyze this region as a candidate region for CLD. After analyzing 10 highly polymorphic microsatellite markers flanking the LPH gene, we obtained significant LOD scores with 8 of them. Seven markers, spanning an area of >8 cM, were in strong linkage disequilibrium ($P <$ .001) with the disease locus, thus confirming the close linkage. We further utilized the characteristics of the isolated Finnish population by analyzing the ancient haplotypes surrounding the CLD locus. Previous molecular genetic studies of the Finnish diseases have shown that, characteristically, one major haplotype or allele has been enriched in disease chromosomes, reflecting one founder mutation (Peltonen and Uusitalo 1997). In CLD, core haplotype 2-2, formed by the closest markers to the CLD locus, D2S314 and D2S2385, was present in 82% of the affected chromosomes and in none of the unaffected chromosomes. Furthermore, 92% of the affected chromosomes carried a single allele 2 of the closest marker, D2S2385, supporting the hypothesis of one major mutation at the CLD locus among Finns. The genealogical data shows that the ancestors of the three parents (8%) carrying the more rare disease allele 7 of D2S2385 were born in three neighboring villages close to the border between Russia and Finland. This allele might represent another, minor mutation underlying CLD. One parent carried allele 1 at this locus. This could be explained by a microsatellite mutation event or by a potential third mutation underlying CLD (fig. 3).

The order of the analyzed markers was defined by the TNG RH-mapping panel of the SHGC G3 map of chromosome 2 and by the RHMAP statistical package (Cox et al. 1990; Boehnke et al. 1991). The distance between critical markers D2S314 and D2S2385 was estimated to be ~840 kb. The LPH gene was found to be positioned between D2S114 and D2S442, by use of two different markers (primers of exons 1 and 17 [SHGC-10723]) of the LPH gene and marker AFMA338YE5, which was positive with LPH-containing PACs; also, the LPH gene is located >2 Mb centromeric to D2S314, outside the critical region.

The physical location of the LPH gene also was determined by sequence-tagged–site (STS) mapping of chromosomes 2q21–specific YACs. The LPH gene was positioned centromeric to marker D2S442, which shows a genetic distance of ~2 cM from marker D2S314. This distance was supported by the RH-mapping data. In

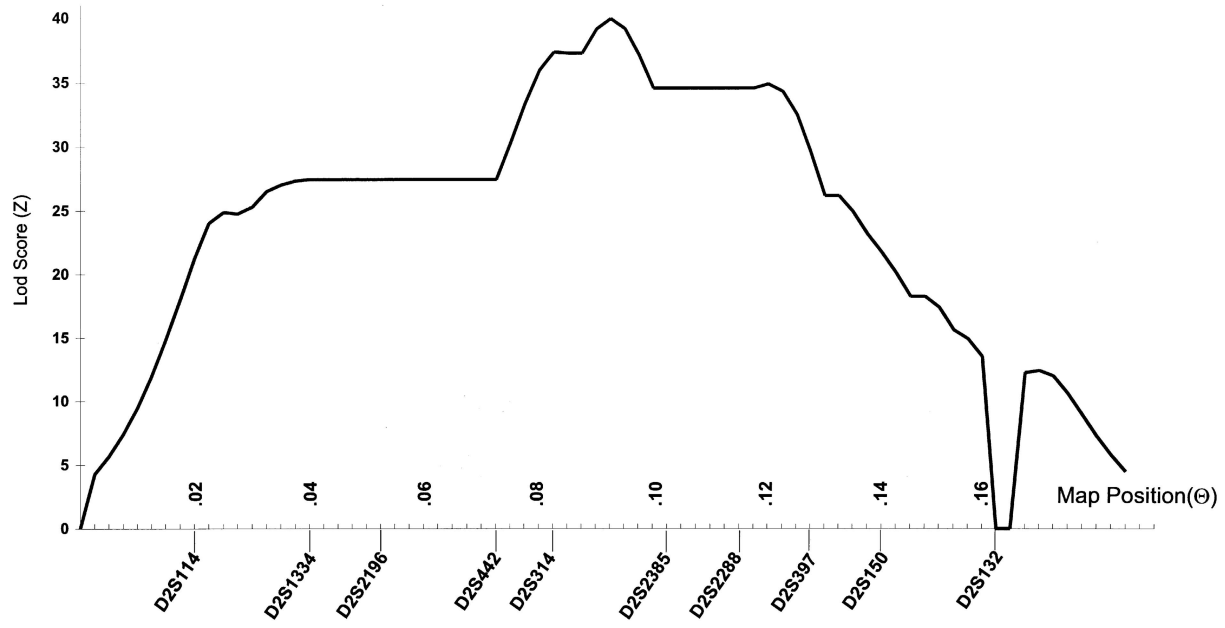

Figure 6 Multipoint likelihood-ratio statistic for the CLD gene. The graph shows the most likely location of the CLD gene, relative to the known fixed positions of the marker loci.

addition, on the basis of the size of the YACs, the corresponding physical distance was determined to be >2 Mb from marker D2S2385, closest to the hypothetical CLD locus. The close proximity of LPH to markers D2S2385 and D2S314 was excluded further by STS mapping of PACs: all three LPH-positive PACs (69K1, 71E19, and 93I13) were negative with these markers. Thus, both genetic and physical data prove unequivocally that the LPH gene lies outside the ancient core haplotype 2-2 and, consequently, can be excluded as a causative gene. The cDNA code and the ~1-kb regulatory region of the LPH gene have been cloned and sequenced (Mantei et al. 1988). Mutation analysis of the cDNA coding for the LPH gene and its promoter region has not revealed any sequence differences in the cDNA either of one of our Finnish CLD patients (Poggi and Sebastio 1991) or of patients with adult-type hypolactasia (Boll et al. 1991), which further supports our conclusion.

The results of the genealogical study showed that the majority of the ancestors of CLD patients was concentrated in the sparsely populated areas of eastern and northern Finland, which were inhabited during and after the internal migration movement in the 16th century (Norio et al. 1973). This is compatible with the observed, relatively wide linkage-disequilibrium interval of >8 cM. The geographic distribution and the extent of linkage disequilibrium of CLD indicate a young mutation, the history of which closely resembles that of infantile cerebellar ataxia, which has been estimated to have been introduced into the Finnish population 30–40 generations ago (infantile onset spinal cerebellar ataxia; Nikali et al. 1994; Varilo et al. 1996*a*). If the CLD mutation was enriched during the same period ($n = 30$), the CLD gene should lie ~350 kb (total range 0–1,580 kb) from the closest marker, D2S2385, according to the linkage-disequilibrium analysis (Terwilliger 1995; Varilo et al. 1996*b*).

The assignment of the CLD locus to chromosome 2q21 will facilitate the identification of the novel gene underlying the congenital form of lactase deficiency. Moreover, it will provide new tools that can be used to analyze potential similarities to the pathogenesis/molecular genetics of adult-type hypolactasia.

Acknowledgments

We are grateful to the CLD families for their participation in this study. We thank Anu Wartiovaara for critically reading the manuscript. This work was supported by the Academy of Finland, the Paulo Foundation, the Finnish Cultural Foundation, and the Foundation of Paediatric Research (Ulla Hjelt Fond), Helsinki.

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

CEPH/Généthon, http://www.genethon.fr (for YAC contigs)
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim (for CLD [MIM 223000] and adult-type hypolactasia [MIM 223100])
Stanford Human Genome Center, http://shgc-www.stanford.edu/Mapping/rh/search.html (for markers SHGC-10723, -30722, -1625, and -20039 and for the G3 map of chromosome 2)

References

Boehnke M, Lange K, Cox DR (1991) Statistical methods for multipoint radiation hybrid mapping. Am J Hum Genet 49: 1174–1188

Boll W, Wagner P, Mantei N (1991) Structure of the chromosomal gene and cDNAs coding for lactase-phlorizin hydrolase in humans with adult-type hypolactasia or persistence of lactose. Am J Hum Genet 48:889–902

Chumakov IM, Rigault P, Le Gall, Bellanne-Chantelot C, Billaut A, Guillou S, Soularue P, et al (1995) A YAC contig map of the human genome. Nature Suppl 377:175–183

Cox DR, Burmeister M, Price ER, Kim S, Myers RM (1990) Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution map of mammalian chromosomes. Science 250:245–250

Dib C, Fauré S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, et al (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380:152–154

Harvey CB, Fox MF, Jeggo PA, Mantei N, Povey S, Swallow DM (1993) Regional localization of the lactase-phlorizin hydrolase gene, LCT, to chromosome 2q21. Ann Hum Genet 57:179–185

Holzel A, Schwarz V, Sutcliffe KW (1959) Defective lactose absorption causing malnutrition in infancy. Lancet i: 1126–1128

Ioannou PA, Amemiya CT, Garnes J, Kroisel PM, Shizuya H, Chen C, Batzer MA, et al (1994) A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. Nat Genet 6:84–89

Kruse TA, Bolund L, Grzeschik K-H, Ropers HH, Sjöström H, Noren O, Mantei N, et al (1988) The human lactase-phlorizin hydrolase gene is located on chromosome 2. FEBS Lett 240:123–126

Lathrop GM, Lalouel J-M, Julier C, Ott J (1984) Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci USA 81:3443–3446

Mantei N, Villa M, Enzler T, Wacker H, Boll W, James P, Hunziker W, et al (1988) Complete primary structure of human and rabbit lactase-phlorizin hydrolase: implications for biosynthesis, membrane anchoring and evolution of the enzyme. EMBO J 7:2705–2713

Nevanlinna H (1972) The Finnish population structure. Hereditas 71:195–236

Nikali K, Koskinen T, Suomalainen A, Pihko H, Peltonen L (1994) Infantile onset spinocerebellar ataxia represents an allelic disease distinct from other hereditary ataxias. Pediatr Res 36:607–612

Norio R, Nevanlinna HR, Perheentupa J (1973) Hereditary diseases in Finland. Ann Clin Res 5:109–141

Peltonen L, Uusitalo A (1997) Rare disease genes: lessons and challenges. Genome Res 7:765–767

Poggi V, Sebastio G (1991) Molecular analysis of the lactase gene in the congenital lactase deficiency. Am J Hum Genet Suppl 49:105

Sahi T, Isokoski M, Jussila J, Launiala K (1972) Lactose malabsorption in Finnish children of school age. Acta Paediatr Scand 61:11–16

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2d ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Savilahti E, Launiala K, Kuitunen P (1983) Congenital lactase deficiency. Arch Dis Child 58:246–252

Terwilliger JD (1995) A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. Am J Hum Genet 56: 777–787

——— (1996) Reply to Sham et al. Am J Hum Genet 58: 1095–1096

Vandenplas S, Wiid I, Grobler-Rabie A, Brebner K, Ricketts G, Wallis A, Bester C, et al (1984) Blot hybridization analysis of genomic DNA. J Med Genet 21:164–172

Varilo T, Nikali K, Suomalainen A, Lönnqvist T, Peltonen L (1996*a*) Tracing an ancestral mutation: genealogical and haplotype analysis of the infantile onset spinocerebellar ataxia locus. Genome Res 6:870–875

Varilo T, Savukoski M, Norio R, Santavuori P, Peltonen L, Järvelä I (1996*b*) The age of human mutation: genealogical and linkage disequilibrium analysis of the CLN5 mutation in the Finnish population. Am J Hum Genet 58:506–512