Am. J. Hum. Genet. 63:1396–1403, 1998

Assignment of the Locus for a New Lethal Neonatal Metabolic Syndrome to 2q33-37

Ilona Visapää,[1,2] Vineta Fellman,[3] Teppo Varilo,[1,2] Aarno Palotie,[4] Kari O. Raivio,[3] and Leena Peltonen[1,2]

[1]Department of Human Molecular Genetics, National Public Health Institute, and [2]Department of Medical Genetics, University of Helsinki; and [3]Hospital for Children and Adolescents and [4]Department of Clinical Chemistry, Helsinki University Central Hospital, Helsinki

Summary

A new neonatal syndrome characterized by intrauterine growth retardation, lactic acidosis, aminoaciduria, liver hemosiderosis, and early death was recently described. The pathogenesis of this disease is unknown. The mode of inheritance is autosomal recessive, and so far only 17 cases have been reported in 12 Finnish families. Here we report the assignment of the locus for this new disease to a restricted region on chromosome 2q33-37. We mapped the disease locus in a family material insufficient for traditional linkage analysis by using linkage disequilibrium, a possibility available in genetic isolates such as Finland. The primary screening of the genome was performed with samples from nine affected individuals in five families. In the next step, conventional linkage analysis was performed in eight families, with a total of 12 affected infants, and finally the locus assignment was proved by demonstrating linkage disequilibrium to the regional markers in 20 disease chromosomes. Linkage analysis restricted the disease locus to a 3-cM region between markers D2S164 and D2S2359, and linkage disequilibrium with the ancestral haplotype restricted the disease locus further to the immediate vicinity of marker D2S2250.

Introduction

A new neonatal metabolic disease has recently been described (Fellman et al. 1998). This disease is characterized by severe intrauterine growth retardation, fulminant lactic acidosis during the first day of life, Fanconi-type aminoaciduria, and abnormalities in iron metabolism, including liver hemosiderosis. Affected infants fail to thrive, and they die neonatally or in early infancy. The disease is distinct from other lactic acidoses, neonatal hemochromatosis, and hepatitis (Fellman et al. 1998). The pathogenesis of this disorder is unknown.

Seventeen patients in 12 families have been diagnosed in Finland since 1965. The parents of the patients are healthy, and at least eight families have one or two healthy children. The male/female ratio of the patients has been 5/12. The occurrence of the disease in sibships and in both sexes in families with healthy parents is consistent with autosomal recessive inheritance. Since this disease has not been described elsewhere in the world, it represents a new member of the Finnish disease heritage, a group of 30 rare monogenic diseases enriched or encountered only in Finland (Norio et al. 1973; de la Chapelle 1993). The genealogies of the affected families have been traced back to the mid-19th century—and the oldest genealogy to the late 17th century—by using church records. Most of the ancestors were born in rural communities in eastern and central Finland, remote from the majority of the population (fig. 1), and some ancient connections between the families have been established (Fellman et al. 1998).

The incidence in Finland can be roughly estimated, from the 14 cases diagnosed 1985–1997, to be 1/56,000 newborns, with a disease allele frequency of .004 and a carrier frequency of 1/120. However, the disease is probably underdiagnosed and the true allele frequency may be higher.

Here we report the assignment of the locus of this new metabolic syndrome to a restricted region on chromosome 2q33-37. The initial genome screen was carried out by genotyping just nine affected individuals. This modification of the haplotype-sharing method (Houwen et al. 1994) has been successfully used in the identification of other disease loci of the Finnish disease heritage (Mäkelä-Bengs et al. 1998; Pekkarinen et al. 1998). Further linkage analyses with markers on the promising chromosomal region were performed in eight families. This family material was too restricted to establish a

Received August 4, 1998; accepted for publication September 14, 1998; electronically published October 23, 1998.

Address for correspondence and reprints: Dr. Leena Peltonen, Department of Human Molecular Genetics, National Public Health Institute, Mannerheimintie 166 FIN-00300, Helsinki, Finland. E-mail: leena.peltonen@ktl.fi

0002-9297/98/6305-0016$02.00

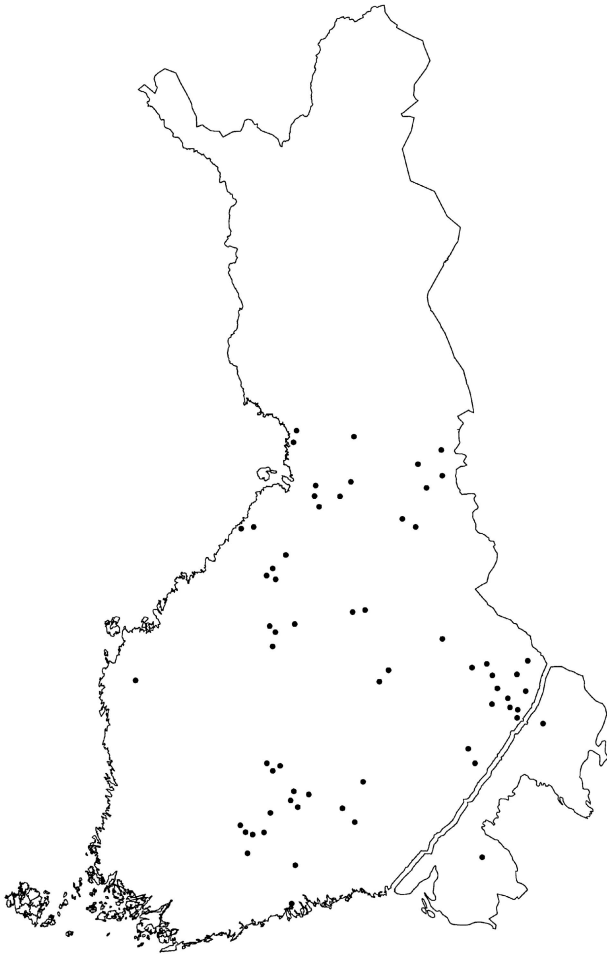

Figure 1 Birthplaces of the patients' great-grandparents on a map of Finland.

confirmed linkage. However, since this syndrome belongs to the Finnish disease heritage, and since the patients most probably have one common ancestor, we were able to confirm the locus assignment by demonstrating linkage disequilibrium to the regional markers.

Subjects and Methods

Patients, Families, and DNA Samples

All of the 14 affected infants diagnosed 1985–1997 were included in this study (Fellman et al. 1998). The infants were born near term but were severely growth-retarded, with an average birth weight of 1,560 g, which corresponds to a −4 SD score for the gestational age. At birth they had a normal umbilical pH, but during the first day of life they developed severe lactic acidosis; by the age of 24 h, the mean arterial blood pH was 7.0 (normal, 7.35–7.45), base excess −23 mmol/l (normal, ±2.5 mmol/l), and lactate 12 mmol/l (normal, 0.7–1.8 mmol/l). All the infants had Fanconi-type aminoaciduria and hemosiderosis of the liver and spleen. In four infants studied, the serum ferritin concentration was increased tenfold, transferrin was low (0.54–0.76 g/l; normal, 1.75–3.13 g/l), and transferrin saturation increased (61%–100%; normal, <52%). No deficiency in respiratory-chain oxygen consumption or enzyme activity was found. No neurological abnormality was detected in these infants. Despite intensive care and alkali therapy, seven infants died at the age of 1–3 days, and seven at the age of 0.5–4 months (Fellman et al. 1998).

Cultured fibroblast specimens were available from 12 patients, and paraffin-embedded tissue samples could be used from the 2 remaining patients. Blood samples from parents and healthy siblings were collected after written informed parental consent was obtained. The pedigrees of the individuals included in this study are shown in figure 2. In one family (family 1) the parents were second cousins, and two families (mother in family 4 and father in family 5) had ancestors with the same surname living in neighboring houses at the end of the 18th century, but the consanguinity could not be confirmed, since the presumable common ancestor had lived before the establishment of local church records (Fellman et al. 1998). Families 1–8, with 12 affected and 26 healthy family members, were included in the linkage studies. For the linkage disequilibrium studies, the family material was completed with the affected infants of families 9 and 10.

As controls for the linkage disequilibrium analyses, 46 parental DNA samples collected for the mapping of infantile-onset spinocerebellar ataxia (IOSCA) and hy-

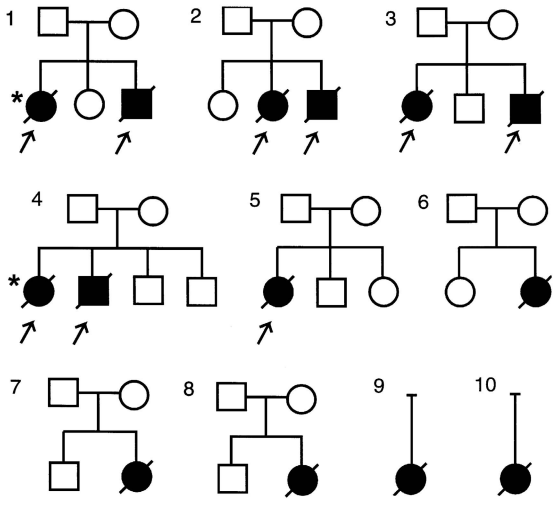

Figure 2 Pedigrees used in the linkage and linkage-disequilibrium studies. Arrows indicate individuals whose DNA samples were analyzed in the initial genome screen. Asterisks denote individuals whose DNA was isolated from paraffin-embedded tissue samples.

drolethalus syndrome were used. These diseases also belong to the Finnish disease heritage, and the birthplace patterns of ancestors are highly similar to that observed in this new disease (Salonen 1986; Varilo et al. 1996). Consequently, these control samples ought to represent the same subpopulation of Finland in which this disease is enriched.

Total DNA was extracted from cultured fibroblasts and peripheral leukocytes according to standard procedures. The paraffin-embedded tissue samples were deparaffinized and the DNA released from cells by the rapid lysis method (Isola et al. 1994).

Marker Genotyping

A slightly modified Weber screening set (Sheffield et al. 1995) of 380 polymorphic microsatellite markers was used in the genome screen. The primers were purchased from the Nordic Human Genome Initiative. The markers were amplified by PCR, with a pipetting robot used for the PCR reactions (Biomek 2000, Beckman). One of the primers was fluorescence 5′–end labeled (cyp5), and the PCR products were separated on an automated laser-fluorescence DNA sequencer (ALF Express, Pharmacia). Genotype data were processed with the ALLELELINKS computer program (Pharmacia).

In the denser mapping of the critical chromosomal region, marker maps from Généthon (Dib et al. 1996), the Genetic Location Database of the University of Southampton, and the Whitehead Institute were used. The primer sequences of markers D2S104, D2S433, and D2S1354 originated from the Genome Database. The primer sequences of the other markers originated from the Généthon collection (Dib et al. 1996). The primers were 5′–end labeled with ^{32}P, and the PCR reactions and the size analysis of the PCR products were performed as described elsewhere (Aaltonen et al. 1993).

Radiation Hybrid Panels

In the radiation hybrid mapping (Goss and Harris 1975; Cox et al. 1990), Stanford medium- and high-resolution RH panels were used (G3 and TNG, respectively). The raw hybrid data of the G3 panel, with markers D2S434, D2S173, D2S433, and D2S163, and of the TNG panel, with marker D2S173, available on the Stanford Human Genome Center web pages, were used. The other ordered markers were typed in both panels twice, or three times if the two typings were ambiguous. The markers were amplified by PCR and the products were separated by electrophoresis on 1.5% agarose gels stained with ethidium bromide.

Statistical Analyses

Two-point linkage analyses were performed with the MLINK option of the LINKAGE package computer programs, version FASTLINK 2.2 (Lathrop and Lalouel 1984), assuming autosomal recessive inheritance, complete penetrance, no phenocopies, and a disease allele frequency of .01. The consanguinity of the parents of family 1 was taken into account in the linkage analysis.

Two-point linkage disequilibrium analyses were performed with the DISLAMB computer program, and multipoint linkage disequilibrium analysis was performed with the DISMULT computer program (Terwilliger 1995). A disease allele frequency of .01 was assumed in the analyses. In the DISMULT program, the distances between the analyzed markers are fixed, and the likelihood is maximized over α (the proportion of disease alleles originally associated with a certain allele) and n (the number of generations since the introduction of the founder disease allele into the population).

The RH mapping data were analyzed with the FORTRAN program RHMAXLIK of the RHMAP package version 3.0 (Boehnke et al. 1991; Lange et al. 1995; Lunetta et al. 1996). This program orders the markers and calculates marker distances in centirays (cR). It allows the use of data from several RH panels simultaneously, which was very important in this study because the ordered markers were too close to each other to be ordered with the Stanford G3 RH panel only, and too distant from each other to be ordered with the Stanford TNG RH panel alone. We used the data of both these panels and the branch and bound ordering option, which ensures that the best locus order is identified. By comparing the raw RH data (G3 panel) of the markers, ordered in relation to each other, to the raw RH data of the flanking markers (available on the Stanford Human Genome Center web pages), we concluded which markers must be located at the ends of the ordered region. Then we chose from the probable locus orders given by computer analysis the most probable order matching this conclusion. The marker distances were converted from centirays to centimorgans by dividing the known total length (in centimorgans) of the linked region proportional to the marker distances in centirays.

Results

Genome Screen and Linkage Analysis

DNA samples from just nine individuals were used in the initial genome scan: four affected sibpairs and the affected infant of family 5 (indicated by arrows in fig. 2). The affected infant of family 5 was included because the mother of family 4 and the father of family 5 were believed to be distant relatives, and consequently they could share a long chromosomal segment around the disease locus. The parents of family 1 were second cousins, and we assumed that the affected sibs in this family

were homozygotes with markers flanking the potential disease locus.

The analysis of the genotypes in the initial genome scan was performed in a very simple and rapid way by simply comparing the genotypes of sibs in each family and looking for markers on which each sibpair carried identical genotypes. After analyzing just 25 markers of the genome scan panel, we found two adjacent markers on chromosome 2q—D1S1649 (which maps to chromosome 2 despite its name) and D2S434, 5-cM apart–revealing identical genotypes in each of four sibpairs. In addition, marker D2S1363, 14 cM away from marker D2S434, revealed identical genotypes in three sibpairs. In family 1, whose parents were second cousins, the sibs were homozygous for all these three markers. These markers were then genotyped in the complete family material. No obligatory recombinations could be identified between marker D2S434 and the disease (maximum LOD score 3.0 at recombination fraction [θ] 0), whereas markers D1S1649 and D2S1363 showed obligatory recombinations between the markers and the disease.

The next step was to analyze several markers around marker D2S434 on chromosome 2q33-37 in the complete family material. Multiple markers between D2S164 and D2S2359 provided positive LOD scores (maximum LOD scores 0.91–3.88 at $\theta = 0$), suggesting linkage on a 3-cM region. In the constructed haplotypes in the families, no obligatory recombination events could be detected between the disease and a total of 11 markers on this 3-cM region. The results of two-point linkage analysis with 15 markers on 2q33-37 are summarized in figure 3.

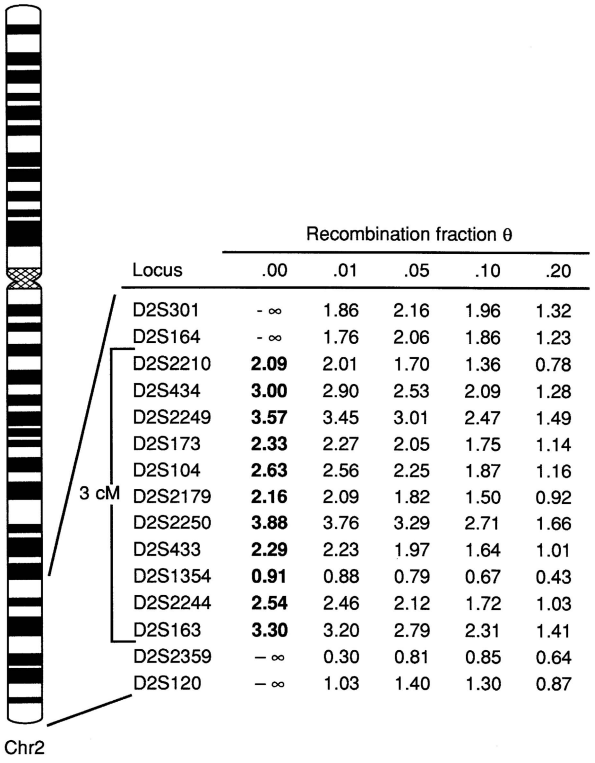

Figure 3 Pairwise LOD scores between the disease and 15 marker loci on 2q33-37 in eight families.

Linkage Disequilibrium and Radiation Hybrid Mapping of the Critical Region

To confirm the linkage and to restrict the critical chromosomal region, we performed linkage disequilibrium analysis with all the markers providing maximum two-point LOD scores at $\theta = 0$. Since we had only 20 affected and 16 nonaffected chromosomes in our family material, we obtained additional nonaffected chromosomes from regionally selected control samples. In total, we used 20 disease and 108 nondisease chromosomes in the linkage disequilibrium studies.

The results of the linkage disequilibrium analysis for each of the 11 markers on the linkage region on 2q33-37 are presented in table 1. Marker D2S2250 showed the most significant linkage disequilibrium; all the disease chromosomes carried the same allele 4, the proportion of which in the nondisease chromosomes was 21%, with a P value of 10^{-12}. This confirmed the assignment of the disease locus to the vicinity of this marker. Markers D2S433 and D2S2244 also provided significant P values in the linkage disequilibrium analysis, and several flanking markers showed suggestive, but not significant, evidence of linkage disequilibrium.

To be able to carry out multipoint linkage disequilibrium analysis and to correctly build the ancestral haplotypes, we performed radiation hybrid mapping of the critical markers to confirm their order. The 11 linked markers were too close to each other to be ordered with the medium-resolution RH panel, but the gap between markers D2S2179 and D2S2250 turned out to be so long that all the markers were not linked in the high-resolution RH panel. We typed the markers in both panels and performed the RH data analysis with a computer program that allows the use of data from both panels simultaneously. Based on this analysis, the most probable marker order and the approximate marker distances (converted to centimorgans) over the linked region are as follows: cen–D2S2210–0.4 cM–D2S434–0.3 cM–D2S2249–0.2 cM–D2S173–0.1 cM–D2S104–0.1 cM–D2S2179–0.4 cM–D2S2250–0.1 cM–D2S433–0.3 cM–D2S1354–0.4 cM–D2S2244–0.4 cM–D2S163–tel. The total map length from D2S2210 to D2S163 was 60 $cR_{10,000}$.

The multipoint linkage disequilibrium analysis carried out with these markers resulted in a maximum LOD

score of 30.9 in the immediate vicinity of marker D2S2250 (fig. 4).

Haplotype Analysis

The haplotypes of disease chromosomes are shown in table 2.The haplotypes of families 1–8 are presented as paternal and maternal disease chromosomes. In families 9 and 10 the parental data were missing, and consequently only the genotypes of the affected infants are shown. For these cases, we constructed the disease haplotypes based on the infants' expected descent from the same ancestors as the other affected individuals.

The 1-cM core haplotype cen-1-2-4-2-3-tel (D2S104-D2S2179-D2S2250-D2S433-D2S1354) was observed in 18/20 of the disease chromosomes. One of the haplotypes that differed from the core haplotype had different alleles with markers D2S104 and D2S2179, and the other had a different allele with marker D2S433. The disease locus is probably very close to marker D2S2250, since all the disease chromosomes carried the same allele 4, which was present in 21% of the nondisease chromosomes. In addition, with marker D2S1354 all the disease chromosomes carried the same allele 3, but this allele was also very common (71%) in the nondisease chromosomes. On the centromeric side of marker D2S2250, the critical DNA region was clearly restricted by marker D2S2179, since the disease haplotype of the mother of family 4 was entirely different from the core haplotype with marker D2S2179 and the markers centromeric from it. On the telomeric side of marker D2S2250, the critical chromosomal region remained less definitive. With marker D2S433, the infant of family 10 carried allele 3 instead of the core haplotype allele 2. However, allele 3 differs from allele 2 by one tetranucleotide repeat only, and so a marker mutation in this disease chromosome cannot be excluded, especially since with marker D2S2244 this infant again carried allele 1, which shows strong linkage disequilibrium with the disease. Therefore, the most critical DNA region is not definitely restricted by marker D2S433, but lies between markers D2S2179 and D2S2244. Most probably, the disease locus is in the immediate vicinity of marker D2S2250.

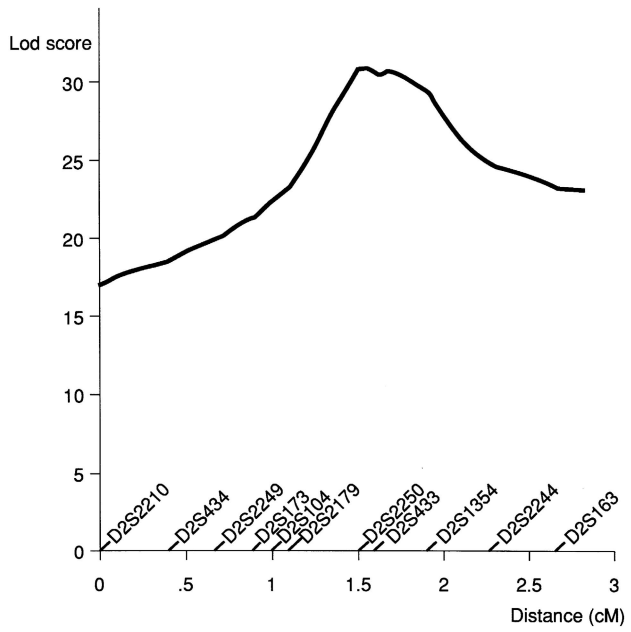

Figure 4 Multipoint linkage disequilibrium analysis of 11 markers in the 3-cM linkage region with 20 disease and 108 nondisease chromosomes.

Table 1

Allele Frequencies of Disease and Nondisease Chromosomes and Two-Point Linkage Disequilibrium Analyses of 11 Marker Loci on 2q33-37

	Percentage of Disease Chromosomes (n = 20) / Percentage of Control Chromosomes (n = 108)									χ^2	P	
Locus	1	2	3	4	5	6	7	8	9	LRT[a]	Value	λ[b]
D2S2210	15/23	35/18	50/57	0/1	0/1	…	…	…	…	.1	.4	.13
D2S434	0/20	40/17	10/33	45/20	5/6	0/4	…	…	…	2.7	.05	.27
D2S2249	0/9	55/16	0/29	35/20	10/17	0/3	0/1	0/5	…	8.7	.002	.45
D2S173	5/29	30/7	60/53	0/7	5/2	0/2	…	…	…	2.4	.06	.23
D2S104	95/66	5/8	0/12	0/8	0/1	0/3	0/1	0/1	…	8.0	.002	.85
D2S2179	0/7	95/66	5/14	0/1	0/4	0/5	0/1	0/2	…	8.1	.002	.85
D2S2250	0/5	0/21	0/29	100/21	0/15	0/3	0/2	0/1	0/3	49.1	<.0001	1.00
D2S433	0/3	95/25	5/56	0/10	0/3	0/3	…	…	…	34.5	<.0001	.93
D2S1354	0/9	0/4	100/71	0/14	0/1	0/1	…	…	…	11.6	.0003	1.00
D2S2244	65/9	35/75	0/10	0/2	0/3	0/1	…	…	…	24.0	<.0001	.61
D2S163	5/8	5/13	10/34	15/11	10/9	55/20	0/3	0/2	…	6.6	.005	.43

[a] χ^2 statistic of the likelihood-ratio test (Terwilliger 1995).
[b] Proportion of excess of a certain allele in the disease-carrying chromosomes.

Table 2

Haplotypes of the Disease Chromosomes in the 3-cM Linkage Region

Family	D2S2210	D2S434	D2S2249	D2S173	D2S104	D2S2179	D2S2250	D2S433	D2S1354	D2S2244	D2S163
1 P	2	4	4	2	1	2	4	2	3	1	6
1 M	2	4	4	2	1	2	4	2	3	1	6
2 P	3	2	2	3	1	2	4	2	3	2	6
2 M	1	3	2	5	1	2	4	2	3	1	4
3 P	3	2	2	3	1	2	4	2	3	1	6
3 M	3	4	2	3	1	2	4	2	3	2	6
4 P	3	2	2	3	1	2	4	2	3	1	6
4 M	1	4	5	1	2	3	4	2	3	2	5
5 P	3	2	2	3	1	2	4	2	3	1	1
5 M	3	4	4	3	1	2	4	2	3	1	4
6 P	2	4	4	2	1	2	4	2	3	1	3
6 M	2	2	2	3	1	2	4	2	3	2	6
7 P	2	4	4	2	1	2	4	2	3	1	5
7 M	3	2	2	3	1	2	4	2	3	1	6
8 P	3	2	2	3	1	2	4	2	3	2	6
8 M	3	3	2	3	1	2	4	2	3	2	6
9 O	1	5	2	3	1	2	4	2	3	2	2
9 O	2	4	4	2	1	2	4	2	3	1	3
10 O	3	2	5	3	1	2	4	2	3	1	6
10 O	2	4	4	2	1	2	4	3	3	1	4

NOTE.—P denotes paternal disease chromosome; M, maternal disease chromosome; and O, haplotype formed from genotypes of the affected offspring (parental data missing). Underlining indicates alleles, which are part of the core haplotype shared with all disease chromosomes.

Discussion

In this study the disease locus was found by using samples from just nine affected individuals in the initial search. This efficient method of monitoring for shared DNA regions has been used successfully in the identification of other monogenic disease loci of the Finnish disease heritage (Mäkelä-Bengs et al. 1998; Pekkarinen et al. 1998). As in these earlier studies, we did not find a conserved haplotype shared by all affected individuals with the marker density used in the first stage of the genome scan, but it proved successful to monitor for shared haplotypes between siblings only and not in all the affected individuals, although a common ancestor was suspected in the distant past.

For the linkage analysis, we had DNA samples from eight families, only four of which had two affected children, and completely reliable disease locus assignment could not be determined by conventional linkage analysis. However, the strong linkage disequilibrium, identified with the regional markers and providing a LOD score of 30.9 in the multipoint analysis, confirmed the locus assignment to the suggestively linked chromosomal region. Linkage disequilibrium analysis has been used in several mapping projects in the Finnish disease heritage in the process of restricting the critical chromosomal region before initiating the physical mapping or candidate gene analyses (Hästbacka et al. 1992; Hellsten et al. 1993; Sulisalo et al. 1994; Höglund et al. 1995). Here too, the linkage disequilibrium analysis restricted the critical chromosomal region from the initial 3 cM to 1 cM, but we used it not only for the fine mapping of the disease locus but also to prove the presumptive linkage in this very small family material. To initiate a gene-mapping attempt with such family material, one would have to assume (*a*) that the families have one common ancestor and consequently there is strong linkage disequilibrium between the markers close to the disease locus, and (*b*) that the mutation is young enough for linkage disequilibrium to be detected with a reasonable marker density. In many diseases in the Finnish disease heritage, >90% of the disease chromosomes carry the same ancestral mutation and linkage disequilibrium can be detected on chromosomal regions 2–13 cM, and we rightly expected this to be the case with this disease also (Peltonen et al. 1995).

Since 18/20 of the disease chromosomes shared the same 1-cM core haplotype with five markers, and the 2 other disease chromosomes also partly shared this ancestral haplotype, all the families probably have one common ancestor and carry the same disease mutation. Significant linkage disequilibrium spans an interval of 1 cM in two-point analyses, a distance comparable to the significant two-point linkage disequilibrium observed in disease chromosomes in another Finnish disease, IOSCA (Nikali et al. 1995). The birthplace patterns of ancestors also resemble each other in these two diseases. The mutations were most probably introduced into the Finnish population around the same time period, some 30–40 generations ago (Varilo et al. 1996). In the 16th century,

the great east-to-west migration of the population and the inhabitation of wilderness spread rare genes to rural subpopulations, which remained surprisingly stable until World War II and present-day industrialization (Norio 1973; Solantie 1988). The clustering of the birthplaces of the ancestors of this disease mirrors this internal migration in Finland.

As expected, the disease haplotypes of the parents of family 1 are identical over a long chromosomal region, because the parents are second cousins. On the other hand, and contrary to our assumption, the disease haplotypes of the mother of family 4 and the father of family 5 did not show similarity on a region longer than that of the other disease chromosomes, even though we assumed that they had a common ancestor ~10 generations ago. Our assumption of distant consanguinity may have been incorrect, the disease mutation may have descended to these families from other ancestors, or abnormally numerous recombinations may have taken place in this chromosomal region over a relatively short time.

We are in the process of constructing a physical map of the critical chromosomal region so we can restrict the gene locus more accurately and then localize and analyze the possible candidate genes and expressed sequence tags (ESTs) in this region. The total RH map length of the 3-cM linkage region was 60 $cR_{10,000}$. Since, on average, 1 $cR_{10,000}$ is 25 kb (Lunetta et al. 1996), the physical length of the most critical 1-cM region may be ~500 kb, which is feasible for transcript mapping. In the Human Gene Map of the National Center for Biotechnology Information, there are four cloned genes and 20 ESTs mapped between markers D2S164 and D2S163, an interval corresponding to our linkage region. The four genes are KIAA0173 and KIAA0236 genes, the function of which has not yet been well defined (Nagase et al. 1996*a*, 1996*b*); the villin gene (VIL1), which is expressed mainly in cells that develop a brush border (Pringault et al. 1991); and the gene coding for the interleukin 8 receptor alpha (IL8RA). Considering the clinical characteristics of the disease, neither villin nor an interleukine receptor represents a tempting candidate gene. Confirmation of the causative character of the disease mutation may not be easy, since all the disease chromosomes most probably carry the same mutation and the same polymorphisms. The discovery of individuals with this disease elsewhere in the world would make the task easier.

Cloning of the disease-causing gene would be essential to reveal the pathogenesis of this disease, which still is a mystery. Assignment of the disease locus to a well-defined chromosomal region already provides the possibility of prenatal diagnostics for families with at least one affected child as well as carrier diagnostics for unaffected sibs and other relatives.

Acknowledgments

We are grateful to the Finnish families for participating in the study and providing blood samples and genealogic data. Dr. Juhani Rapola and the University Hospital of Kuopio kindly provided the patient samples and Drs. Kaisu Nikali, Tuula Lönnqvist, and Riitta Salonen the control samples. We would like to thank Dr. Joseph Terwilliger for his valuable advice in the statistical treatment of the data and Dr. Reijo Norio for his genealogical collaboration. Anne Jokiaho, Mira Kyttälä, Mervi Eeva, Elli Kempas, Pekka Ellonen, Anne Vikman, Arja Terola, Jaana Hartiala, and Mari Sipilä are thanked for their skillful technical assistance. This study was supported by the Academy of Finland and the Foundation for Pediatric Research, Ulla Hjelt Fund.

Electronic-Database Information

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

Généthon, http://www.genethon.fr (for marker maps)
Genetic Location Database of University of Southampton, http://cedar.genetics.soton.ac.uk/public_html/index.html (for marker maps)
Genome Database, http://www.gdb.org/ (for for primer sequences, accession numbers GDB:185492, GDB:686616, and GDB:365051)
National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/ (for the Gene Map of the Human Genome)
Stanford Human Genome Center, http://shgc-www.Stanford.edu/ (for raw RH data)
Whitehead Institute/MIT Center for Genome Research, http://www-genome.wi.mit.edu/ (for marker maps)

References

Aaltonen J, Komulainen J, Vikman A, Palotie A, Wadelius C, Perheentupa J, Peltonen L (1993) Autoimmune polyglandular disease type I: exclusion map using amplifiable multiallelic markers in a microtiter well format. Eur J Hum Genet 1:164–171

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

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

de la Chapelle A (1993) Disease gene mapping in isolated human populations: the example of Finland. J Med Genet 30:857–865

Dib C, Faure 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

Fellman V, Rapola J, Pihko H, Varilo T, Raivio KO (1998) Iron-overload disease in infants involving fetal growth re-

tardation, lactic acidosis, liver haemosiderosis, and aminoaciduria. Lancet 351:490–493

Goss SJ, Harris H (1975) New method for mapping genes in human chromosomes. Nature 255:680–684

Hästbacka J, de la Chapelle A, Kaitila I, Sistonen P, Weaver A, Lander E (1992) Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. Nat Genet 2:204–211

Hellsten E, Vesa J, Speer MC, Mäkelä TP, Järvelä I, Alitalo K, Ott J, et al (1993) Refined Assignment of the infantile neuronal ceroid lipofuscinosis (INCL, CLN1) locus at 1p32: incorporation of linkage disequilibrium in multipoint analysis. Genomics 16:720–725

Höglund P, Sistonen P, Norio R, Holmberg C, Dimberg A, Gustavson K, de la Chapelle A, et al (1995) Fine Mapping of the Congenital Chloride Diarrhea Gene by Linkage Disequilibrium. Am J Hum Genet 57:95–102

Houwen RHJ, Baharloo S, Blankenship K, Raeymaekers P, Juyn J, Sandkuijl LA, Freimer NB (1994) Genome screening by searching for shared segments: mapping a gene for bening recurrent intrahepatic cholestasis. Nat Genet 8:380–386

Isola J, De Vries SD, Chu L, Ghazvini S, Waldman F (1994) Analysis of changes in DNA sequence copy number by comparative genomic hybridization in archival paraffin-embedded tumor samples. Am J Path 145:1301–1308

Lange K, Boehnke M, Cox DR, Lunetta KL (1995) Statistical methods for polypoid radiation hybrid mapping. Genome Res 5:136–150

Lathrop GM, Lalouel JM (1984) Easy calculations of lod scores and genetic risks on small computers. Am J Hum Genet 36:460–465

Lunetta K, Boehnke M, Lange K, Cox DR (1996) Selected locus and multiple panel models for radiation hybrid mapping. Am J Hum Genet 59:717–725

Mäkelä-Bengs P, Järvinen N, Vuopala K, Suomalainen A, Ignatius J, Sipilä M, Herva R, et al (1998) Assignment of the disease locus for the lethal congenital contracture syndrome to a restricted region of chromosome 9q34, by genome scan using five affected individuals. Am J Hum Genet 63:506–516

Nagase T, Seki N, Ishikawa K, Ohira M, Kawarabayasi Y, Ohara O, Tanaka A, et al (1996*a*) Prediction of the coding sequences of unidentified human genes. VI. The coding sequences of 80 new genes (KIAA0201–KIAA0280) deduced by analysis of cDNA clones from cell line KG-1. DNA Res 3(5):321–329, 341–354

Nagase T, Seki N, Ishikawa K, Tanaka A, Nomura N (1996*b*) Prediction of the coding sequences of unidentified human genes. V. The coding sequences of 40 new genes (KIAA0161–KIAA0200) deduced by analysis of cDNA clones from human cell line KG-1. DNA Res 3(1):17–24

Nikali K, Suomalainen A, Terwilliger J, Koskinen T, Weissenbach J, Peltonen L (1995) Random search for shared chromosomal regions in four affected individuals: the assignment of a new hereditary ataxia locus. Am J Hum Genet 56: 1088–1095

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

Pekkarinen P, Hovatta I, Hakola P, Järvi O, Kestilä M, Lenkkeri U, Adolfsson R, et al (1998) Assignment of the locus for PLO-SL, a frontal-lobe dementia with bone cysts, to 19q13. Am J Hum Genet 62:362–372

Peltonen L, Pekkarinen P, Aaltonen J (1995) Messages from an isolate: lessons from the Finnish gene pool. Biol Chem Hoppe-Seyler 376:697–704

Pringault E, Robine S, Louvard D (1991) Structure of the human villin gene. Proc Natl Acad Sci 88:10811–10815

Salonen R (1986) The Meckel and hydrolethalus syndromes diagnostic, genetic, and epidemiological studies. PhD thesis, University of Helsinki, Helsinki

Sheffield VC, Weber JL, Buetow KH, Murray JC, Even DA, Wiles K, Gastier J, et al (1995) A collection of tri- and tetranucleotide repeat markers used to generate high quality, high resolution human genome-wide linkage maps. Hum Mol Genet 4:1837–1844

Solantie R (1988) Climatic conditions for the cultivation of rye with reference to the history of settlement in Finland. Fennoscandia Archaeologica 5:3–20

Sulisalo T, Klockars J, Mäkitie O, Francomano CA, de la Chapelle A, Kaitila I, Sistonen P (1994) High-resolution linkage-disequilibrium mapping of the cartilage-hair hypoplasia gene. Am J Hum Genet 55:937–945

Terwilliger J (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

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