A Genomewide Screen in a Four-Generation Dutch Family with Celiac Disease: Evidence for Linkage to Chromosomes 6 and 9

Martine J. van Belzen, Ph.D., Martine M. Vrolijk, M.Sc., Jos W.R. Meijer, M.D., J. Bart A. Crusius, Ph.D., Peter L. Pearson, Ph.D., Lodewijk A. Sandkuijl, M.D.†, Roderick H. J. Houwen, M.D., Ph.D., and Cisca Wijmenga, Ph.D.

Departments of Biomedical Genetics and Pediatric Gastroenterology, University Medical Center Utrecht, Utrecht; Department of Pathology, Rijnstate Hospital, Arnhem; Laboratory for Immunogenetics, VU Medical Center, Amsterdam, The Netherlands

OBJECTIVES: Celiac disease is caused by the interaction of multiple genes and environmental factors. Inheritance of the disease shows a complex pattern with a 10% sibling recurrence risk. The HLA-region is a major genetic risk locus in celiac disease, but genes outside this region are expected to contribute to the disease risk as well. The aim of this study was to identify the loci causing celiac disease in one large Dutch family with apparent dominant transmission of the disease.

METHODS: The family comprised 17 patients in four generations, with possible transmission of the disease by both grandparents. Microsatellite markers evenly spread over all chromosomes were genotyped and linkage analysis was performed using both dominant and recessive disease models and a model-free analysis.

RESULTS: Disease susceptibility in the family was linked to the HLA-region (lod score of 2.33) and all patients were HLA-DQ2. A dominantly inherited non-HLA locus with a maximum lod score of 2.61 was detected at 9p21-13, which was shared by 16 patients. Model-free analysis identified another possible non-HLA locus, at 6q25.3, which was shared by 14 patients ($p = 0.01$). Neither of these regions was detected in a genomewide screen in Dutch affected sibpairs, but the 9p21 locus has been implicated in Scandinavian families.

CONCLUSIONS: Two potential non-HLA loci for celiac disease were identified in this large Dutch family. Our results provide replication of the Scandinavian 9p21 locus, and suggest that this locus plays a role in celiac disease patients from different Caucasian populations.

INTRODUCTION

Celiac disease is an autoimmune disorder with a multifactorial etiology. Ingestion of gluten, present in wheat, barley, and rye, leads to a range of aberrations in the small intestinal mucosa of these patients. The range of abnormalities can be classified according the modified Marsh classification. The Marsh type III lesion is the most severe lesion, with the presence of intraepithelial lymphocytosis with damage of the mucosal epithelial cells, crypt hyperplasia and partial (IIa), subtotal (IIb) or total (IIc) villous atrophy (1, 2). A wide spectrum of clinical symptoms can be associated with this lesion, including diarrhea, abdominal pain, and bloating. Symptoms due to malabsorption of nutrients, such as fatigue, weight loss, anemia, and osteopenia, are also frequently present. Most celiac disease patients that adhere strictly to the gluten-free diet (GFD) show improvement or complete disappearance of the clinical symptoms and recovery of the small intestinal mucosa (3). Recently, it has been recognized that less severe lesions, like lymphocytosis without or with crypt hyperplasia (Marsh types I and II, respectively), can also be associated with the same clinical spectrum (1, 3, 4).

Celiac disease is a strongly inheritable disorder, with a relative risk of approximately 10% for siblings of a patient. One important genetic factor is the HLA-region, with the majority of patients expressing HLA-DQ2, and almost all of the remaining patients expressing HLA-DQ8 (5). However, the genetic contribution of the HLA-region to celiac disease has been estimated at only $\sim$40% (6, 7). Therefore, non-HLA genes must also contribute to the disease.

Identification of susceptibility genes for multifactorial disorders is hampered by multigenic etiology and genetic heterogeneity (8). It is expected that a causative variant in a susceptibility gene for these diseases will be common in the population, and most carriers of this variant will not develop the disease since they do not carry all the necessary disease
susceptibility genes to pass the disease threshold. Furthermore, different genes can cause identical phenotypes in different families. To overcome these difficulties, large numbers of cases and controls are needed in association studies, and large numbers of families with multiple patients are needed for linkage analysis. Within a single family the disease is expected to be genetically homogeneous, as the same genes are likely to cause the disease in all patients from that family. Therefore, a large family with many patients in different generations may provide a unique opportunity to identify genes that cause these multifactorial diseases.

We present here a family with 17 celiac disease patients in four generations. Ten out of thirteen (76%) siblings from the second generation have celiac disease, which is much more than would be expected from the average sibling relative risk. We therefore hypothesized that a single non-HLA gene, with high penetrance, causes celiac disease in this family and a genomewide screen was performed to localize this gene locus.

SUBJECTS AND METHODS

The Study Family

This study was approved by the Medical Ethics Committee of the University Medical Center Utrecht and written informed consent was obtained from all the participants. The family is of Dutch origin and lives in the north of the Netherlands. At the time of ascertaining, only five patients were diagnosed with celiac disease: individuals (IDs) 05, 09, 14, 19, and 20 (Fig. 1). Serological screening was offered to 36 additional family members and they were screened for antigliadin IgA (AGA), antigliadin IgG (AGG), and antientomysium (Ema) antibodies. Antitissue transglutaminase (tTG) antibodies could only recently be determined, and this test was performed in only three individuals. Twelve additional celiac disease patients were diagnosed by a duodenal biopsy. All biopsy specimens, including those from the initial patients, were retrieved for reevaluation and Marsh classification by one experienced pathologist (JWRM). The histological lesions were classified according to Marsh (2). The Marsh III category was subdivided into three groups; i.e., partial (Marsh IIIa), subtotal (Marsh IIIb), and total (Marsh IIIc) villous atrophy (9). All patients were genotyped at the HLA-DQA1 and DQB1 loci as described before (10). An overview of symptoms, serology, and Marsh classification before and on a GFD, together with the HLA-DQ type, is given in Table 1. Serology after GFD was not available.

Linkage analysis was performed using an “affecteds-only” strategy, so only 23 individuals were included in this study: the 17 celiac disease patients, five spouses, and the
## Table 1. Patient Characteristics

<table>
<thead>
<tr>
<th>ID*</th>
<th>Year of Birth</th>
<th>Symptoms Before GFD</th>
<th>Serology Before GFD</th>
<th>DQ Type</th>
<th>Symptoms After GFD</th>
<th>Marsh After GFD</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>1920</td>
<td>Fatigue, DH, cannot eat bread</td>
<td>AGA, AGG</td>
<td>ND</td>
<td>No GFD</td>
<td></td>
</tr>
<tr>
<td>02</td>
<td>1923</td>
<td>Fatigue, anemia</td>
<td>AGA, Ema</td>
<td>II</td>
<td>No GFD</td>
<td></td>
</tr>
<tr>
<td>05</td>
<td>1947</td>
<td>Fatigue, anemia, weight loss, abdominal pain/bloating</td>
<td>ND</td>
<td>IIIb</td>
<td>Disappeared</td>
<td>IIa</td>
</tr>
<tr>
<td>07</td>
<td>1949</td>
<td>Fatigue, diarrhea, irritable</td>
<td>AGA, Ema</td>
<td>IIIc</td>
<td>Disappeared</td>
<td>ND</td>
</tr>
<tr>
<td>08</td>
<td>1950</td>
<td>Fatigue, abdominal pain</td>
<td>AGA, Ema</td>
<td>IIIa</td>
<td>Disappeared</td>
<td>ND</td>
</tr>
<tr>
<td>09</td>
<td>1952</td>
<td>Diarrhea, abdominal pain</td>
<td>ND</td>
<td>IIIb</td>
<td>Disappeared</td>
<td>ND</td>
</tr>
<tr>
<td>10</td>
<td>1953</td>
<td>Fatigue</td>
<td>Ema</td>
<td>IIIa</td>
<td>Disappeared</td>
<td>Improved</td>
</tr>
<tr>
<td>11</td>
<td>1954</td>
<td>Fatigue, diarrhea</td>
<td>AGA, AGG, Ema</td>
<td>IIIb</td>
<td>Improved</td>
<td>ND</td>
</tr>
<tr>
<td>12</td>
<td>1955</td>
<td>Fatigue, abdominal pain/bloating</td>
<td>AGA, Ema</td>
<td>IIIa</td>
<td>Disappeared</td>
<td>II</td>
</tr>
<tr>
<td>13</td>
<td>1957</td>
<td>Fatigue, abdominal bloating, irritable</td>
<td>AGA</td>
<td>IIIa</td>
<td>Disappeared</td>
<td>ND</td>
</tr>
<tr>
<td>14</td>
<td>1960</td>
<td>Fatigue, diarrhea, anemia, weight loss, DH</td>
<td>ND</td>
<td>NA</td>
<td>Disappeared</td>
<td>ND</td>
</tr>
<tr>
<td>16</td>
<td>1962</td>
<td>Fatigue, diarrhea</td>
<td>Ema</td>
<td>IIIa</td>
<td>Disappeared</td>
<td>Improved</td>
</tr>
<tr>
<td>19</td>
<td>1971</td>
<td>Fatigue, bleeding gums</td>
<td>ND</td>
<td>II</td>
<td>Disappeared</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>1976</td>
<td>Fatigue, diarrhea, weight loss, irritable, tooth enamel defects</td>
<td>ND</td>
<td>NA</td>
<td>Disappeared</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>1978</td>
<td>Fatigue</td>
<td>Ema</td>
<td>II</td>
<td>Disappeared</td>
<td>ND</td>
</tr>
<tr>
<td>32</td>
<td>1989</td>
<td>Fatigue</td>
<td>Ema</td>
<td>IIIb</td>
<td>Disappeared</td>
<td>ND</td>
</tr>
<tr>
<td>41</td>
<td>1998</td>
<td>Fatigue, irritable</td>
<td>AGA, Ema, tTG</td>
<td>IIIb</td>
<td>Disappeared</td>
<td>ND</td>
</tr>
<tr>
<td>55</td>
<td>1978</td>
<td>None</td>
<td>Ema</td>
<td>IIIb</td>
<td>No GFD</td>
<td></td>
</tr>
</tbody>
</table>

*ID refers to the identification numbers used in Figure 1.

GFD = gluten-free diet; DH = dermatitis herpetiformis; AGA = antigliadin IgA; AGG = antigliadin IgG; Ema = antiendoymysium IgA; ND = not done; NA = not available; and tTG = antitissue transglutaminase IgA. Note: tTG testing was only performed in ID 01 and 41.

grandfather (Fig. 1 and Table 1). Twelve of the patients had a Marsh III lesion and all of them responded well to the GFD. Three individuals, IDs 02, 19, and 24, showed a Marsh II lesion. The grandmother (ID 02) does not adhere to a GFD, but the presence of the celiac disease-specific lesion.

Linkage Analysis

A parametric linkage analysis was performed with the MLINK program of the LINKAGE package (12) using an “affecteds-only” approach. This approach is preferred in complex diseases, as it is possible that healthy individuals also carry some of the disease susceptibility loci. All 23 individuals who were genotyped were included in the linkage analysis and the disease status of the 17 patients, either with a Marsh II lesion, a Marsh III lesion, or with unavailable biopsy specimens, was affected. The five spouses were not tested for

### Genotyping Microsatellite Markers

A genomewide screen was performed in all 23 individuals (Fig. 1). A total of 321 microsatellite markers were genotyped, evenly spread over the genome with an average distance of 15 cM. A marker spacing <5 cM was obtained for regions that had been implicated in our genomewide screen in Dutch affected sibpairs (6q21-22 and 19p13.1) (11) and regions showing linkage in other studies (CELIAC1 in the HLA-region, CELIAC2 at 5q31-33, and CELIAC3 in the CTLA4/CD28 region at 2q33-34). Regions with lod scores >1.0 were selected for finemapping, and 50 additional markers were genotyped in these regions. Our marker set consisted mainly of screening set 6 from the Marshfield Center for Medical Genetics, complemented with markers selected from the Marshfield and Ensembl genetic maps. DNA from the family members was amplified by PCR, together with three CEPH reference samples and a negative control. The reaction volume of 10 μl contained 25 ng of DNA, 0.2 mM dNTPs, 2.5 mM MgCl2, 50 ng fluorescence-labeled primer and 0.4 U AmpITaq Gold (PE Applied Biosystems, Foster City, CA). The PCR products were pooled and separated on a 3700 DNA sequencer (PE Applied Biosystems) and analyzed by Genescan 3.5 and Genotyper 2.0 software (PE Applied Biosystems). All genotypes were checked independently by two researchers. The identity of the markers was verified by comparing genotypes of the CEPH reference samples to the CEPH genotype database. A Mendelian inheritance check was performed and markers with Mendelian errors were excluded from the linkage analysis.
celiac disease and were labeled unknown. The disease
status of the grandfather (ID 01) was also unknown, as celiac
disease was not histologically proven in this individual.

A two-point linkage analysis was performed for all mark-
ers using a dominant inheritance model as well as a re-
escissive model (parametric linkage analysis). A multipoint
analysis was performed in regions with a lod score > 1.0 us-
ing the FASTLINK program. For these analyses, the model
that produced linkage in the two-point analysis was used. A
maximum of three successive markers was included in the
parametric multipoint analysis because of calculation time
considerations. Power calculations were performed with the
SLINK and MSIM programs. For all analyses, the pene-
trance for the normal genotype was equal to the population
frequency of celiac disease (0.005) and the penetrance of
the disease genotype was 0.8. Disease allele frequency was
0.001, based on the hypothesis that a rare variant of a single
susceptibility gene is causing the exceptionally large number
of patients in this family. Allele frequencies were set equal
for all alleles, since all the parental genotypes were available
and therefore the allele frequencies in the data
file were not
used in the analysis. The order and location of the markers
on the chromosomes was based on the Marshfield genetic
map.

Although the disease transmission in this family suggests
autosomal dominant inheritance, other models cannot be ex-
cluded. To check for the possibility that linkage was missed
because of applying a wrong model, a model-free (nonpara-
metric) analysis was also performed using the Genehunter
program (13). When calculating the nonparametric lod score
(NPL) statistic, a disease model is not used but the sharing
of marker alleles by all possible pairs of affected individuals
is determined and compared to the expected values based on
the familial relationship. The study family is too large for
the Genehunter program, and was therefore divided into two
smaller families (family A: IDs 01, 02, 05, 07, 08, 09, 18,
19, 20, 21, 24, 40, and 41; family B: IDs 01, 02, 10, 11, 12,
13, 14, 16, 31, 32, 54, and 55). For this purpose, the affection
status of both grandparents (ID 01 and 02) was set as un-
known. NPL values obtained for both family branches were
totaled.

**RESULTS**

To determine the power of the study family, simulation stud-
ies were performed using the same parameters as selected for
the linkage analysis. The maximum attainable lod score was
4.17 assuming a dominant model of transmission and 3.65 as-
suming a recessive model. These simulation lod scores were
obtained with 100% informativity of the marker and no re-
combination between the marker and the disease locus. The
probability of obtaining lod scores of at least 1.0, 2.0, or 3.0
were 53%, 25%, and 8% respectively. Because of the low
probability of obtaining high lod scores, all regions with a
lod score > 1.0 were selected for finemapping.

Two-point lod scores were calculated for all 321 mi-
crosatellite markers and seven regions with a lod score > 1
were identified on chromosomes 3, 6, 9, 10, 15, 16, and 19.
Fifty additional markers were typed in these regions, and
the highest lod scores obtained in each region are shown in
Table 2, together with the p-value corresponding to the NPL
statistic from the two-point model-free analysis. The HLA
region on chromosome 6 was linked to celiac disease in

<table>
<thead>
<tr>
<th>Location</th>
<th>Position *cM</th>
<th>Mb</th>
<th>Marker</th>
<th>Lod Score</th>
<th>NPL p-Value</th>
<th>Information Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>3q12.3</td>
<td>119.1</td>
<td>101.6</td>
<td>D3S2459</td>
<td>1.13</td>
<td>0.02</td>
<td>0.91</td>
</tr>
<tr>
<td>6p21.3</td>
<td>–</td>
<td>31.3</td>
<td>MIB</td>
<td>2.33</td>
<td>0.01</td>
<td>0.95</td>
</tr>
<tr>
<td>9p21-13</td>
<td>55.3</td>
<td>31.1</td>
<td>D9S43</td>
<td>1.55</td>
<td>0.03</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>58.3</td>
<td>32.1</td>
<td>D9S1118</td>
<td>1.64</td>
<td>0.02</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>59.3</td>
<td>34.0</td>
<td>D9S1817</td>
<td>2.61</td>
<td>0.0005</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>59.9</td>
<td>35.3</td>
<td>D9S163</td>
<td>2.61</td>
<td>0.0003</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>60.6</td>
<td>36.9</td>
<td>D9S50</td>
<td>2.61</td>
<td>0.0003</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>61.4</td>
<td>37.4</td>
<td>D9S1804*</td>
<td>0.78</td>
<td>0.19</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>38.2</td>
<td>D9S200</td>
<td>1.43</td>
<td>0.02</td>
<td>0.95</td>
</tr>
<tr>
<td>10q26.2</td>
<td>156.2</td>
<td>–</td>
<td>D10S1223</td>
<td>1.33</td>
<td>0.23</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>160.0</td>
<td>129.1</td>
<td>D10S1676</td>
<td>1.33</td>
<td>0.08</td>
<td>0.82</td>
</tr>
<tr>
<td>15q22.3</td>
<td>62.4</td>
<td>59.7</td>
<td>D15S153*</td>
<td>1.36</td>
<td>0.12</td>
<td>0.61</td>
</tr>
<tr>
<td>16q23.2</td>
<td>108.3</td>
<td>82.4</td>
<td>D16S3098</td>
<td>1.17</td>
<td>0.004</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>111.1</td>
<td>83.9</td>
<td>D16S422</td>
<td>1.17</td>
<td>0.23</td>
<td>0.66</td>
</tr>
<tr>
<td>19p13.1</td>
<td>45.5</td>
<td>17.6</td>
<td>D19S899</td>
<td>1.17</td>
<td>0.42</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>47.3</td>
<td>18.3</td>
<td>D19S915</td>
<td>1.17</td>
<td>0.42</td>
<td>1.0</td>
</tr>
</tbody>
</table>

*The genetic locations are based on the Marshfield genetic map and the physical locations are based on the Ensembl map (March 2003 release). Markers without a position were not present in that map.

†Markers marked by an asterisk were not informative in ID 02.

‡Lod scores are based on the dominant transmission model, except for the markers at 10q26.2, which were obtained with the recessive model.

§p-Value corresponding to the NPL statistic of the nonparametric linkage analysis.
this family, with a lod score of 2.33 at marker MIB. Multipoint linkage analysis resulted in the same lod score. The most promising non-HLA region is located on chromosome 9, which reached a maximum lod score of 2.61 at markers D9S50, D9S1817, and D9S163. D9S1817 was completely informative in all meioses, so the maximum multipoint lod score in this region was also 2.61. Model-free analysis also demonstrated increased allele sharing among the affected individuals, with an NPL score of 7.81 ($p = 0.0005$) at D9S1817. Haplotypes of eight markers from this region are depicted in Figure 1. Sixteen out of seventeen patients share a five-marker haplotype ranging from D9S1817 to D9S1874. Only one patient, ID 11, inherited the other haplotype from the grandmother. The maximum size of this candidate region is 6.1 Mb and is defined by markers D9S1118 and D9S200. Multipoint linkage analysis of the 3q12.3, 10q26.2, 16q23.2, and 19p13.1 regions resulted in lod scores identical to the single point lod scores. A lower multipoint lod score was obtained for 15q22.3 (0.67).

Nonparametric linkage analysis identified three additional regions with nominal $p$-values $<0.05$ on chromosomes 6, 10, and 11. The NPL value at 6q25.3 was 4.58 ($p = 0.01$) and peaked at marker D6S969. This locus was inherited from the grandfather (ID 01) and is shared by 14 of 16 affected offspring. Only IDs 13 and 16 inherited the other haplotype. Marker D10S1227 at 10q21.1 produced an NPL value of 4.22 ($p = 0.02$), this locus was inherited from the grandmother (ID 02) and was transmitted to 11 of 16 offspring. On chromosome 11 at p15.4, an NPL value of 5.49 ($p = 0.002$) was obtained at markers D11S2362 and D11S1760. This locus is shared by 10 offspring and was inherited from the grandfather. It is present in all individuals in branch A of the family. However, only two individuals in branch B, IDs 11 and 13, also carry this haplotype. The high NPL value of this locus was contributed only by branch A, resulting in a biased total NPL value for the entire family.

DISCUSSION

A Dutch family with 17 celiac disease patients in four generations is presented in this paper. Although the family was already known with celiac disease since 1978, only five patients were diagnosed by 2001. Intensive screening of the other family members resulted in the identification of 12 additional patients. Before diagnosis, most of these 12 patients did not report complaints. But after only a few weeks of adhering to a GFD, most of them reported a dramatic decrease of fatigue and disappearance of mild abdominal pain and diarrhea. All these individuals stated that their quality of life had increased significantly and they are willing to adhere to the GFD in the future. This emphasizes once again the importance of screening for celiac disease in family members of patients, even when no obvious complaints are reported.

The family presented here was originally collected for our affected sibpair study (11) but was excluded because of the exceptionally large number of affected individuals. This family alone would have provided 46 sibpairs, which would have seriously biased our results. One explanation for the high percentage of affected offspring in the second generation could be consanguinity between the grandparents, resulting in transmission of the same disease locus by both grandparents. The grandparents (IDs 01 and 02) originated from two small neighboring villages in the north of the Netherlands, but genealogical investigation did not reveal consanguinity within the last six generations.

The apparent dominant disease transmission through four generations suggested a Mendelian cause of the disease in this family. Surprisingly however, bilinear transmission of two possible non-HLA disease loci, at 9p21-13 and 6q25.3 respectively, was detected, and neither of these loci was shared by all patients. Sixteen affected individuals share a haplotype of 2.1 cM at 9p21-13, which was transmitted by the affected grandmother (ID 02). Only one patient, ID 11, inherited the other maternal haplotype. There are no reasons to doubt her diagnosis, so her disease is probably caused by other susceptibility genes. In addition, a locus at 6q25.3 was inherited from the grandfather (ID 01) and shared by 14 of 16 affected offspring. Transmission of a disease locus by the grandfather is not at all unlikely, as he presents with several characteristics compatible with celiac disease. The 6q25.3 locus was missed in the parametric analysis, because of the model that was used. This locus is present in ID 11, possibly compensating for the absence of the 9p21-13 locus. Several other loci with parametric lod scores $>1.0$ and nonparametric $p$-values $<0.05$ were also present, and some of these may add to the disease risk in some patients.

Recently we have completed the genomewide screen in affected Dutch MIII sibpairs (11) and identified major celiac disease loci on chromosome regions 19p13.1 and 6q21-22. However, no evidence for linkage to 9p21-13 or 6q25.3 was obtained in this study, implying that the major loci in this large family are not important loci in Dutch celiac disease patients in general. On the other hand, the two major loci identified in the affected sibpair study have no significant effect in this family. Further evidence for linkage of celiac disease to 9p21 was provided by two previous studies in Swedish/Norwegian ($p = 0.038$) (14) and Finnish affected sibpairs (lod $= 1.11$) (15), indicating that this locus may be a true, but not a major, risk factor for celiac disease. Linkage of celiac disease to 6q25.3 has not been reported previously, but linkage to this region has been observed several times in type 1 diabetes mellitus (16). Since type 1 diabetes mellitus is associated with celiac disease (17), this locus may harbor a shared susceptibility gene for these disorders.

As expected, linkage to the HLA-region on 6p21.3 was also present, with a maximum lod score of 2.33 at marker MIB. Although all affected individuals in the study family are DQ2-positive and this marker is informative in all but one transmission, the maximum attainable lod score was not reached.
Both grandparents are heterozygous DQ2 carriers, and ID 14 is homozygous DQ2. He has transmitted the grandfather’s DQ2 haplotype to his affected daughter (ID 32), thereby destroying the linkage. This situation illustrates the problems frequently encountered in linkage analysis in complex disorders. When risk alleles are common in the population, they can also be present in healthy spouses. In this family, DQ2 is also married-in by ID 18 (DQ2 heterozygote) and ID 40 (DQ2/DQ8).

Even within this one family, celiac disease appears to be genetically heterogeneous and of multigenic origin. Surprisingly, the major loci identified in our affected sibpair study do not play a significant role in this family. Nevertheless, this large family may be important for understanding the disease pathogenesis. Although the loci involved in this family may not be of major importance for celiac disease in general, they may play a role in a small proportion of celiac disease patients. Hence, these loci may be hard to detect by an affected sibpair approach because of their small relative risk.

ACKNOWLEDGMENTS

The authors are grateful to all members of the study family, who participated with great enthusiasm and willingness in this study. We wish to thank Winny van der Hoeven for her help in collecting blood samples and clinical information from the family members. Jackie Senior is acknowledged for critically reading the manuscript and Eric Hennekam for the genealogical investigation. The study was financially supported by a grant from the Dutch Digestive Diseases Foundation (WS 97-44). This research work was carried out in the Department of Biomedical Genetics.

This paper is dedicated to the memory of Lodewijk Sandkuijl. He participated in this project from the beginning and his contribution to the design of the study and analysis of the data was of great importance to us. He died unexpectedly, shortly before the genomewide screen was completed.

Reprint requests and correspondence: Cisca Wijmenga, Ph.D., Complex Genetics Group, Department of Biomedical Genetics, University Medical Center Utrecht, Stratumweg 2.117, Universiteitsweg 100, 3584 CG Utrecht, The Netherlands.

Received July 21, 2003; accepted October 22, 2003.

REFERENCES