

ARTICLE

Linkage of autosomal-dominant common variable immunodeficiency to chromosome 4q

Anemone Finck¹, Jos WM Van der Meer², Alejandro A Schäffer³, Jessica Pfannstiel¹, Claire Fieschi⁴, Alessandro Plebani⁵, A David B Webster⁶, Lennart Hammarström⁷ and Bodo Grimbacher^{*,1}

¹Department of Medicine, Division of Rheumatology and Clinical Immunology, University Hospital Freiburg, Freiburg, Germany; ²Department of Medicine and Nijmegen University, Center for Infectious Diseases, Nijmegen, The Netherlands; ³National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, USA; ⁴Hôpital Saint-Louis, Center Hayem, Paris, France; ⁵Department of Pediatrics and Institute of Medicine 'Angelo Novicelli', University of Brescia, Brescia, Italy; ⁶Department of Clinical Immunology, Royal Free Hospital, London, UK; ⁷Department of Medicine, Division of Clinical Immunology, Karolinska Institute at Huddinge, Stockholm, Sweden

The phenotype of common variable immunodeficiency (CVID) is characterized by recurrent infections owing to hypogammaglobulinemia, with deficiency in immunoglobulin (Ig)G and at least one of IgA or IgM. Family studies have shown a genetic association between CVID and selective IgA deficiency (IgAD), the latter being a milder disorder compatible with normal health. Approximately 20–25% of CVID cases are familial, if one includes families with at least one case of CVID and one of IgAD. Nijenhuis *et al* described a five-generation family with six cases of CVID, five cases of IgAD, and three cases of dysgammaglobulinemia. We conducted a genome-wide scan on this family seeking genetic linkage. One interval on chromosome 4q gives a peak multipoint LOD score of 2.70 using a strict model that treats only the CVID patients and one obligate carrier with dysgammaglobulinemia as affected. Extending the definition of likely affected to include IgAD boosts the peak multipoint LOD score to 3.38. The linkage interval spans at least from D4S2361 to D4S1572. We extended our study to a collection of 32 families with at least one CVID case and a second case of either CVID or IgAD. We used the same dominant penetrance model and genotyped and analyzed nine markers on 4q. The 32 families have a peak multipoint LOD score under heterogeneity of 0.96 between markers D4S423 and D4S1572 within the suggested linkage interval of the first family, and an estimated proportion of linked families (α) of 0.32, supporting the existence of a disease-causing gene for autosomal-dominant CVID/IgAD on chromosome 4q.

European Journal of Human Genetics (2006) 14, 867–875. doi:10.1038/sj.ejhg.5201634; published online 26 April 2006

Keywords: common variable immunodeficiency; recurrent infections; IgA deficiency; linkage analysis; primary immunodeficiency disorder; chromosome 4

*Correspondence: Dr B Grimbacher, Division of Rheumatology and Clinical Immunology, University Hospital Freiburg, Hugstetterstr. 55, Freiburg 79106, Germany.

Tel: +49 761 270 3696; Fax: 49 761 270 3531;

E-mail: grimbacher@medizin.ukl.uni-freiburg.de

Received 5 December 2005; revised 2 March 2006; accepted 16 March 2006; published online 26 April 2006

Introduction

Common variable immunodeficiency (CVID, OMIM 240500) is the most frequent symptomatic primary immunodeficiency. It is characterized by recurrent infections and deficiencies of immunoglobulin (Ig)A and IgG and, in half of the patients, IgM. The phenotype also includes autoimmune disorders in about 25% of patients, a

similar fraction suffering from gastrointestinal diseases. A smaller proportion have granulomatous disease, and there is also an increased risk of malignancies.^{1,2} The standard treatment for CVID is Ig substitution. The incidence of CVID is estimated to be between 1/25 000 and 1/66 000,^{3,4} being much less common than selective IgA deficiency (IgAD), which has an incidence between 1/600 and 1/800. Patients are given a diagnosis of CVID when there is no known cause for the hypogammaglobulinemia. Thus, the etiology is probably diverse and likely to encompass both monogenic and polygenic disorders, and to be influenced by environmental factors, including chronic infections.

Genetic studies of both CVID and IgAD show that between 20–25% of cases are familial,⁵ with a predominance of autosomal-dominant over recessive inheritance.^{5–7} In addition, CVID and IgAD commonly occur in different members of the same family and occasionally IgAD progresses to CVID.^{8–11} Most previous genetic studies of CVID or IgAD have concentrated on the HLA region on chromosome 6. Some of these studies used a case–control design,^{12–16} while others used a genetic linkage or haplotype sharing analysis.^{5,17,18} Vořechovský *et al*,¹⁸ designated the HLA susceptibility locus *IGAD1*.¹⁹ A previous genome-wide study in our laboratory on a three-generation German CVID family in 2002 revealed linkage to chromosome 5p,²⁰ but the disease-causing gene still remains elusive.

Candidate gene approaches have recently been successful in the identification of the molecular cause of some patients with CVID. Salzer *et al*²¹ and Castigli *et al*²² found that approximately 10% of CVID patients have either heterozygous or homozygous mutations of the *TNFRSF13B* gene, which encodes TACI. One patient has been reported with a homozygous mutation of *TNFRSF13C*, which encodes the BAFF-R (Warnatz K *et al*. XIth Meeting of the European Society for Immunodeficiency, Versailles, 2004;

abstract #B72). Nine CVID patients share an identical homozygous deletion in the *ICOS* gene,²³ and at least four patients have been identified with a homozygous mutation in *CD19*.²⁴ In addition, a few patients originally diagnosed with CVID have later been shown to be affected by X-linked lymphoproliferative syndrome (OMIM 308240) caused by mutations in *SH2D1A*^{25,26} or X-linked agammaglobulinemia (OMIM 307200) caused by mutations in *BTK*.^{27–29}

In this paper, we present a genetic linkage study in the largest CVID family published to date.³⁰ The family, NL1, exhibits autosomal-dominant inheritance with reduced penetrance and cases of CVID (6 individuals), IgAD (5 individuals), and dysgammaglobulinemia (three individuals). We show that the phenotype of CVID or IgAD in this family is linked to a locus on chromosome 4q. In an attempt to replicate this finding, we also present linkage analysis on 32 multiplex families with at least one CVID case, referred to below as the 'EU cohort'. The EU families come from a larger family collection that was ascertained for IgAD and previously used to map the *IGAD1* locus on chromosome 6.^{18,19}

Patients, materials, and methods

Patients and diagnostic measurements

The five-generation NL1 family with 54 individuals was described and illustrated in reference³⁰ (Figure 1). Of the 54 individuals, six adults have CVID and five of these are alive and had previously donated blood, which could be used for genotyping. There are an additional eight individuals with a heterogeneous assortment of dysgammaglobulinemias (Table 1); five of these have IgAD (defined as a serum IgA concentration of <0.9 g/l), while three have IgA concentrations in the normal range. One of the five IgAD individuals is both the son and father of CVID patients,

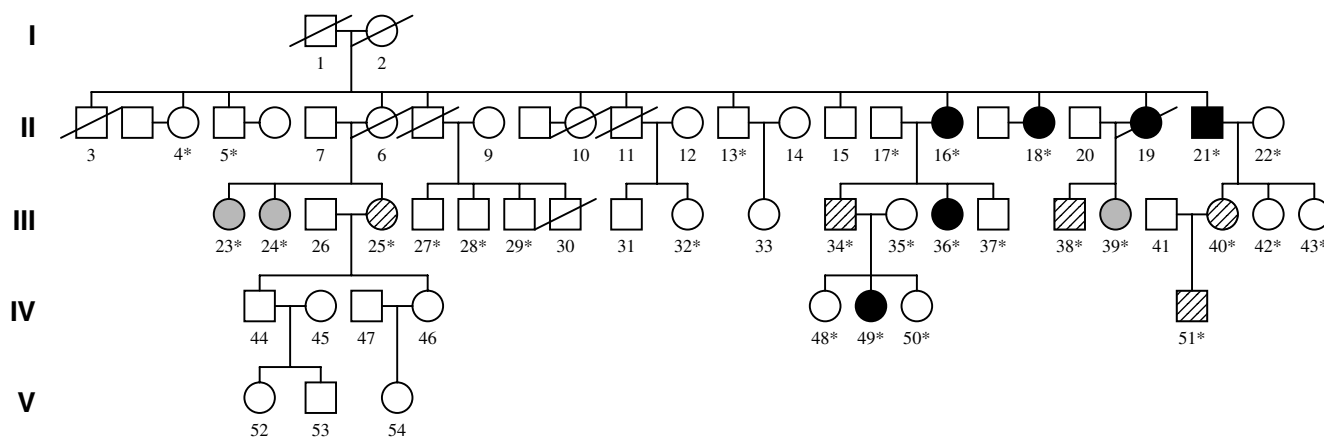


Figure 1 Pedigree of NL1. Filled symbols illustrate CVID patients and gray symbols individuals with dysgammaglobulinemia and IgAD. Circles represent female individuals, and squares represent male. A crossing line indicates deceased individuals. Individuals 25, 34, 38, 40 and 51 are IgA deficient. Genotyped individuals have an asterisk next to their number.

Table 1 Immunoglobulin levels [g/l] in family NL1

Ind#	Ind# (Nienhuis <i>et al</i>)	IgG (5.7–14.0)	IgA (0.9–3.0)	IgM (0.4–1.6)	γ -1 (4.9–11.4)	γ -2 (1.5–6.4)	γ -3 (0.2–1.1)	γ -4 (0.08–1.4)
4	II.2	8.51	1.78	0.66	4.88	4.02	0.77	0.15
5	II.4	8.2	1.66	0.66	4.75	3.24	0.26	0.71
13	II.14	9.09	1.99	1.04	7.42	2.33	0.41	0.19
15	II.16	12.53	2.85	0.6	10.42	2.28	0.45	0.3
16	II.17	2.55	0.24	0.13	—	—	—	—
18	II.19	2.49	0.13	0.53	1.98	<0.3	0.3	<0.08
19	II.21	0.85	0.23	0.38	0.6	<0.25	0.11	<0.04
21	II.23	9.45 ^a	0.91	0.16	6.77 ^a	3.29 ^a	0.54 ^a	0.25 ^a
23	III.1	6.54	1.7	0.46	5.82	1.19	0.35	0.2
24	III.2	7.16	1.54	0.77	5.34	0.83	0.32	0.41
25	III.3	5.4	0.93	0.4	5.01	0.61	0.14	<0.04
27	III.5	12.3	4.16	1.3	6.9	5.19	0.65	1.04
28	III.6	9.09	1.54	0.7	5.73	3.49	0.79	0.28
29	III.7	9.58	1.32	0.58	5.99	3.74	1.07	0.21
32	III.10	8.66	0.9	1.68	6.84	2.33	0.86	0.08
34	III.12	8.18	0.76	0.21	5.47	1.21	0.3	0.14
36	III.14	1.81	0.06	0.48	—	—	—	—
37	III.15	10.88	2.2	1.28	7.55	3.9	0.75	0.14
38	III.16	5.94	0.78	0.46	4.22	0.99	0.4	<0.09
39	III.17	7.09	1.15	1.63	6.9	0.48	0.38	<0.04
40	III.18	5.9	0.57	0.73	4.95	1.19	0.37	<0.04
42	III.20	7.2	0.92	0.68	6.25	1.16	0.56	0.19
43	III.21	10.79	1.12	1.64	5.99	5.06	0.47	0.29
48	IV.5	7.9	0.31	0.6	5.1	1.2	0.25	0.15
49	IV.6	3.64	0.15	0.21	2.64	0.23	0.18	<0.04
51	IV.8	7.88	0.31	0.35	5.66	2.13	0.63	0.36
53	V.2	6	0.3	0.9	4.2	0.5	0.37	<0.04
54	V.3	6	0.59	0.7	4.2	0.8	0.34	0.14

γ 1–4 are the four subclasses of IgG.

^aValue obtained under Ig replacement therapy.

so if the hypothesis of dominant inheritance suggested in reference³⁰ is correct, he is an obligate carrier of the CVID-associated mutation.

Vořechovský *et al*^{5,18,19} described 101 families with multiple cases of immunodeficiency, where 43 families had at least one case of CVID. Among the 43 families, 34 had sufficient sample material available for the genotyping. Two families were found to have distinct mutations in the candidate gene *TNFRSF13B*, which encodes TACI,²¹ leaving 32 EU families for this study.

There is no consensus in the literature on what level of serum IgA is compatible with a diagnosis of CVID/IgAD. Some patients in a large CVID cohort reported by Cunningham-Rundles and Bodian³¹ had levels above 0.5 g/l. In the EU families, the diagnosis of IgAD required a level below 0.05 g/l, considered necessary to reduce uncertainty about affection status in small families. The reason why a higher threshold of 0.9 g/l was used for the NL1 family³⁰ was because the obligate carrier had an IgA concentration of 0.76 g/l and also had IgM deficiency. None of the five CVID patients in this family had an IgA concentration <0.06 g/l.

Informed written consent was obtained from each individual prior to participation under an internal ethics review board-approved clinical study protocol

(#239/99 for BG and 435/99 for LH). Twenty-eight of the 54 individuals provided samples for genotyping (indicated in Figure 1).

Genotyping

A total of 324 microsatellite markers were genotyped on the NL1 family for the genome-wide scan and fine mapping. A total of nine microsatellite markers on chromosome 4q, overlapping with the best markers in the NL1 family, were genotyped in the EU families. Markers for fine mapping were selected with the aid of the Marshfield map³² where the given intermarker distances were used in multipoint genetic linkage analysis. The EU families had been genotyped across the genome as described previously.¹⁸ However, the nearest usable markers from that data are D4S398 (72.5 cM from the 4p telomere) and D4S430 (126.1 cM), which appear to fall outside the linkage interval of the NL1 family.

Primers and other reagents for genotyping were purchased from Invitrogen Research Genetics (Karlsruhe, Germany), biomers.net (Ulm, Germany) and Qiagen (Hilden, Germany). The polymerase chain reactions (PCR) for genotyping were performed according to the protocols accompanying the reagents. The PCR products were sequenced on an ABI377 sequencer (PE Applied

Biosystems, Foster City, CA, USA) using the COLLECTION and ANALYSIS software. Integer allele lengths were assigned in a semiautomatic manner using the GENOTYPER (PE Applied Biosystems) software package.

Genetic linkage analysis

Genetic linkage analysis of both the NL1 and EU families was carried out with a model-based approach assuming dominant inheritance in all analyses, and variable penetrance in some. One and two-marker LOD scores were computed using the FASTLINK software package.^{33–35} Four-marker multipoint LOD scores for the NL1 family were computed using VITESSE.³⁶ Full (all nine markers) multipoint LOD scores for the EU families were computed with GENEHUNTER.^{37,38} The possibility of locus heterogeneity for the EU families was evaluated with HOMOG³⁹ for single-marker LOD scores and with GENEHUNTER for the full multipoint LOD scores. LOD scores achieved under different models of affection status in the NL1 family were assessed with the simulation software FASTSLINK.^{34,40,41} By generating and analyzing all linked replicates, we computed the highest LOD score achievable. By simulating with all unlinked replicates, we estimated empirical *P*-values for the true LOD scores.

All analyses shown here used a disease allele frequency of 0.001. For the NL1 family, we used all equal marker allele frequencies due to the small sample size. For the EU families, marker allele frequencies were estimated using the downfreq program.⁴² All analyses assigned individuals to either a penetrance class with no uncertainty (encoded in LINKAGE notation as 0.0, 1.0, 1.0) or considerable uncertainty (0.05, 0.75, 0.75), where the latter class represents a 5% phenocopy rate and 75% penetrance. We used a phenocopy rate much larger than the population incidence to allow for substantial locus heterogeneity, since this has been established for CVID.^{18,20–23}

In the NL1 family, we show results on two different assignments for the seven hypogammaglobulinemic individuals who are not obligate carriers: (A) all unknown (0) status, and (B) IgAD individuals affected in the equivocal class and non-IgAD individuals unknown. For the EU families, all phenotyped individuals who were neither married in nor CVID-affected were assigned to the equivocal class and assigned affection status 1 (unaffected) if there was no evidence of IgAD, and affection status 2 (affected) if they had IgAD.

The extension from model A to model B was carried out because if the NL family maps to the same genetic locus as some of the EU families, which were ascertained for IgAD, then deficiency of IgA ought to be considered as affected. Unlike the NL1 family, the EU families do not include healthy individuals identified as having only IgG or IgM deficiency and normal IgA. Therefore, we assigned the unknown status (0) to the three hypogammaglo-

bulinemic adults in the NL1 family who have normal IgA concentrations.

Candidate gene sequencing

Candidate genes on chromosome 4 were evaluated by sequencing either the coding regions of the genes on genomic DNA (for *NFKB1*, *SCYE*, *CASP6*, and *DAPPI*) or the respective cDNA (*BANK1*). All primers were sought with the aid of the Primer Select software (PE Applied Biosystems); sequences are available upon request. RNA was reverse transcribed into cDNA using ImProm-II Reverse Transcription System (Promega Corporation, Madison, WI, USA), cDNA was amplified by PCR and subsequently sequenced with the amplification primers. After gel electrophoresis on an ABI Prism™ 377 DNA Sequencer, the data were analyzed by the DNA Sequencing Analysis software, version 3.4 (PE Applied Biosystems) and Sequencer™, version 3.4.1 (Gene Codes Corporation, Ann Arbor, MI, USA).

Results

Genetic linkage analysis of the NL1 family

We initially tried an affection status assignment where every individual with any form of hypogammaglobulinemia was considered affected and any individual without hypogammaglobulinemia was considered unaffected. This approach yielded no promising loci for two reasons. First, different individuals are deficient in nonoverlapping subsets of Ig subtypes, so the simple definition that hypogammaglobulinemic = affected is internally inconsistent. Second, the broad definition does not follow autosomal-dominant inheritance, while the narrow definition of CVID does follow dominant inheritance, except for one obligate carrier. Therefore, the analyses shown here use stricter criteria for affected status. CVID-affected individuals and the one obligate carrier were always assigned in a class with no uncertainty as affected (2) and married-in individuals were always assigned as unaffected (1), also in a penetrance class with no uncertainty. Unphenotyped individuals used solely to connect the families were assigned an unknown (0) affection status. In the NL1 family, and for the analyses shown here, an individual who had no signs of CVID or IgAD was assigned an unaffected (1) status only if that individual was the full sibling of a CVID-affected individual, or otherwise was assigned an unknown affection status (0). This reduces power, but greatly reduces the risk of confounding due to polygenic effects; it makes the analysis close to an affected-only model, which is often preferred for complex diseases such as CVID.

Single-marker analysis with a cautious model in which only CVID-affected individuals and one obligate carrier are affected, and the unaffected siblings are put in the equivocal penetrance class shows promising single-marker scores along a broad interval on chromosome 4 (Table 2).

Table 2 Markers in the linkage region on chromosome 4 and single marker LOD scores under a cautious model where only CVID patients and one obligate carrier are treated as affected

Marker	Genetic (cM)	Sequence (Mb)	0.00	0.01	0.03	0.05	0.10	0.15
D4S401	60.2	46.1	-0.03	-0.03	-0.03	-0.03	-0.02	-0.02
D4S3248	72.5	59.9	-0.07	-0.07	-0.06	-0.05	-0.04	-0.03
D4S3243	88.0	81.3	0.94	0.93	0.90	0.88	0.80	0.73
D4S2361	93.5	85.4	1.95	1.91	1.83	1.75	1.55	1.33
D4S1534	95.1	86.7	0.75	0.73	0.70	0.67	0.59	0.51
D4S423	100.8	92.8	2.16	2.12	2.04	1.95	1.73	1.49
D4S1572	108.0	104.1	1.15	1.12	1.05	0.99	0.82	0.65
D4S2623	114.0	111.2	0.58	0.58	0.56	0.54	0.49	0.43
D4S427	124.5	121.8	1.72	1.69	1.63	1.57	1.41	1.23

The six rightmost column headings are recombination fractions between the putative disease gene and a marker.

Using the consecutive markers D4S1534, D4S423, D4S1572, and D4S2623 in a multipoint analysis yields a LOD score of 2.30, which is the maximum achievable for these affection and penetrance assignments.

To get a better idea of the most likely linkage interval, we then moved to a stricter model with full penetrance, but no additional assignments of affected or unaffected status over the cautious model. The stricter model yields the LOD scores shown in Table 3. The four-point LOD score is 2.71, again the maximum achievable.

The EU data set available for replication was ascertained for IgAD, so we then extended the cautious model in the NL1 family by assigning affected status to the four adult individuals, besides the obligate carrier, who had IgAD but not CVID. These individuals were assigned to the equivocal (0.05, 0.75, 0.75) penetrance class, just as the unaffected siblings were in the cautious model. The single-marker LOD scores for the extended model are shown in Table 4. Marker D4S423 has a peak single-marker score of 3.25 at recombination fraction $\theta=0$, indicating that the marker alleles segregate perfectly with affected status. The four-marker LOD score peaks at 3.38 with a nearly flat curve across all four markers. The multipoint LOD score of 3.38 is again the maximum achievable for these affection status and penetrance assignments. Using FASTSLINK, we generated and analyzed 3000 *unlinked* replicates of the pedigree with the same locus specifications as for D4S423. The highest LOD score among the unlinked replicates is 2.67, well below the observed score of 3.25. Thus, the observed score is significant at $P < 0.001$.³⁹

Based on the three models, we conclude that in the NL1 family, the phenotype 'CVID or IgAD' is linked to a region on chromosome 4q. The linkage interval extends at least from marker D4S2361 at 85.4Mb from the top of chromosome 4 through marker D4S1572 at 104.1Mb (Figure 2). It might extend further, if the strict model treats the unaffected siblings of the CVID-affected individuals with too much certainty. The single-marker LOD scores vary quite a bit due to variations in the informativeness of markers in the NL1 pedigree. However, the multipoint LOD scores show no significant

Table 3 Markers in the linkage region on chromosome 4 and single marker LOD scores under a strict model where only CVID patients and one obligate carrier are treated as affected

Marker	0.00	0.01	0.03	0.05	0.10	0.15
D4S401	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
D4S3248	-0.33	-0.33	-0.32	-0.31	-0.26	-0.20
D4S3243	$-\infty$	-0.30	0.13	0.30	0.39	0.49
D4S2361	2.36	2.31	2.23	2.14	1.90	1.66
D4S1534	0.82	0.80	0.76	0.73	0.69	0.55
D4S423	2.57	2.52	2.43	2.33	2.23	1.82
D4S1572	1.56	1.52	1.44	1.37	1.17	0.97
D4S2623	$-\infty$	-0.64	-0.20	-0.03	0.16	0.22
D4S427	$-\infty$	0.66	1.05	1.18	1.24	1.18

This differs from the first model (Table 2) in that unaffected siblings of the affected individuals cannot carry the disease associated allele (i.e. full penetrance). The six rightmost column headings are recombination fractions between the putative disease gene and a marker.

Table 4 Markers in the linkage region on chromosome 4 and single marker LOD scores under an extension of the cautious model analyzed in Table 2

Marker	0.00	0.01	0.03	0.05	0.10	0.15
D4S401	-0.02	-0.02	-0.02	-0.02	-0.01	-0.01
D4S3248	0.54	0.53	0.52	0.50	0.46	0.42
D4S3243	1.55	1.53	1.49	1.44	1.39	1.16
D4S2361	2.98	2.93	2.82	2.70	2.59	2.11
D4S1534	1.29	1.27	1.22	1.17	1.12	0.92
D4S423	3.25	3.19	3.07	2.95	2.83	2.32
D4S1572	1.04	1.05	1.06	1.05	1.03	0.85
D4S2623	0.59	0.59	0.57	0.55	0.53	0.43
D4S427	0.62	0.66	0.74	0.79	0.82	0.83

The extension is that four adults with IgAD are treated as affected with the same equivocal penetrance function (0.05, 0.75, 0.75) used for the unaffected siblings of CVID-affected individuals. The six rightmost column headings are recombination fractions between the putative disease gene and a marker.

preference for any disease gene located within the (D4S2361,D4S1572) interval. In fact, in the extended model, the peak LOD score for D4S1572 occurs at $\theta=0.03$ (and not 0.00), suggesting that the CVID-associated gene lies above D4S1572. However, we cannot

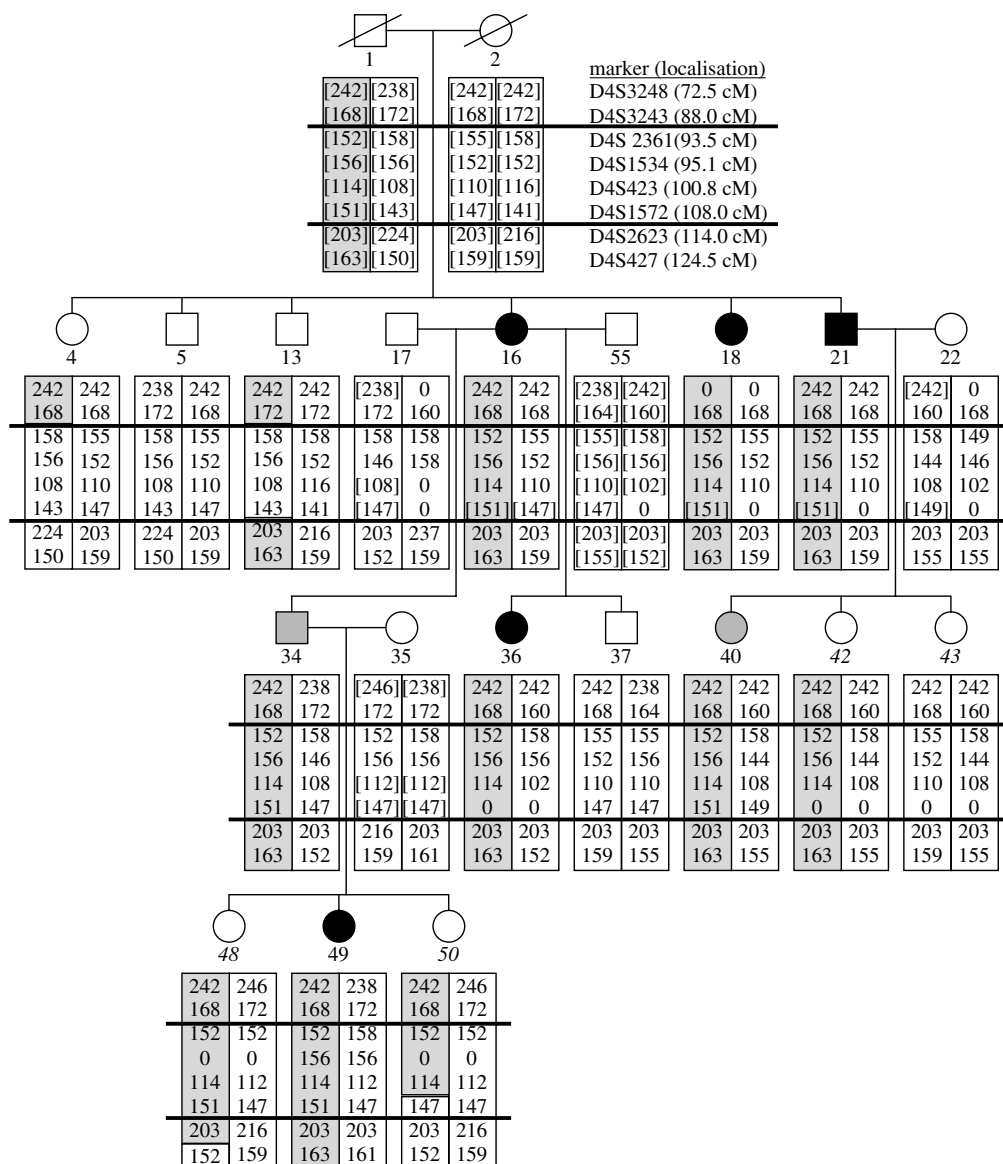


Figure 2 Haplotype analysis including the linkage region of chromosome 4. The numbers in the box indicate the length in bases of the microsatellite's allele. Alleles in brackets were inferred. The gray shading illustrates the disease-associated alleles/haplotype. The black lines indicate the position of the limiting crossovers under a strict model (Table 3) where unaffected siblings cannot carry the disease allele. Individuals appearing in italics were assigned unknown (0) disease status in the linkage analysis under the strict model (Table 3).

declare D4S1572 as a definitive linkage interval boundary based on any accepted syllogisms for inference in genetic linkage analysis.³⁹

Supporting evidence from EU families

An important goal in genetic studies of complex diseases such as CVID is replication of any statistical finding in additional data sets. Therefore, we studied the 32 EU families using nine markers in, or near, the linkage interval of the NL1 family; five of the markers are shared in both parts of the study. We used the same penetrance function and method of analysis, but this is not a pure

replication attempt as the EU families were ascertained for low IgA. Peak single-marker LOD scores are shown in Table 5. Eight of the nine markers achieve positive scores, and the remaining marker (D4S2623) achieves a neutral score of 0. Two of the markers, D4S423 and D4S1572, achieve peak scores ≥ 1.0 , and these are among the perfectly segregating markers for the NL1 family. Since there is known locus heterogeneity for CVID based on the four causative genes found so far, and the peak scores occur at recombination fractions significantly above 0, we tested for locus heterogeneity within the EU family collection. Using HOMOG, we estimated

Table 5 Peak total single-marker scores under homozygosity for EU families and the recombination fraction (θ) at which the peak scores occur, for nine markers in or near the linkage interval of the NL1 family

Marker	Peak score	θ
D4S405	0.26	0.29
D4S3000	0.38	0.29
D4S392	0.48	0.24
D4S2922	0.25	0.32
D4S1534	0.29	0.31
D4S423	1.65	0.15
D4S1572	1.00	0.22
D4S2623	0.00	0.43
D4S427	0.36	0.32

that for D4S423 the LOD score under heterogeneity (HLOD) at $\theta=0$ is 1.25, and the estimated proportion of linked families (α) is 0.48. For D4S1572, the HLOD at $\theta=0$ is 1.03 with $\alpha=0.34$. For marker D4S1572, the LOD score under heterogeneity is slightly higher than the peak (over all recombination fractions) LOD score under heterogeneity.

Finally, we used GENEHUNTER to estimate a full (all nine markers) LOD score under heterogeneity, which peaks at 0.96, with an estimated 32% of the families linked. The peak occurs between markers D4S423 and D4S1572. GENEHUNTER also computes a model-free NPL score, which peaks at 1.73 ($P<0.03$). The GENEHUNTER-reported average (over the 32 families) information content is 0.85 at the location of the peak NPL score. In sum, analysis of the EU families provides support for the chromosome 4q locus suggested by the NL1 family.

Our study design was to genotype and analyze family NL1 first, followed by targeted genotyping of the EU cohort in the region(s) where NL1 showed evidence of linkage. Even though the families were ascertained using different clinical criteria, one can combine the data sets for markers they share in common, and analyze the 33 families together. As the marker allele frequencies are estimated jointly, the results may be different from summing LOD scores for the NL1 family (Table 4) and the EU cohort (Table 5). In particular, this is one way to confirm that using uniform marker allele frequencies in the NL1 analysis does not inflate LOD scores significantly. We carried out a combined analysis for four of the shared markers and yielded peak single markers under *homogeneity* as follows: D4S1534 (score = 0.76; $\theta=0.26$), D4S423 (score = 3.98; $\theta=0.12$), D4S1572 (score = 1.92; $\theta=0.20$), and D4S427 (score = 0.73; $\theta=0.27$). Thus, the peak scores for D4S423 and D4S1572 increase substantially over their peak scores in NL1 alone (Table 4). As these are the two markers genotyped in both data sets that appear to be in the minimal linkage interval for NL1, the combined results provide additional evidence in favor of linkage to this region of chromosome 4q.

Functional candidate genes in the linkage interval

Using NCBI's MapViewer (<http://www.ncbi.nlm.nih.gov/mapview>) and associated hyperlinks, we identified five genes that are located within the linkage interval with known functions that make them plausible candidates to be mutated in CVID. As the etiology of CVID includes abnormalities in the number of B and T cells, cytokine production, apoptosis, and other functions,⁴³ the list should by no means be considered exhaustive. *DAPPI* is a candidate because of its role in B-cell signaling and because *DAPPI*^{-/-} mice have a deficiency in IgG3.⁴⁴ *BANK1* is a candidate because of its role in B cell response to antigens.⁴⁵ *NF κ B* is a candidate because of its prominent role in regulating immune responses.⁴⁶ *SCYE1* is a candidate because of its role in inflammatory response and apoptosis in T cells.⁴⁷ *CASP6* is a candidate because it induces apoptosis in response to infection by *Streptococcus pneumoniae*.⁴⁸ We sequenced all exons of *CASP6*, *DAPPI*, *NF κ B1*, and *SCYE1* on genomic DNA, and the cDNA of *BANK1* and *NF κ B1* in at least one CVID patient from the NL1 family. However, no mutations were found in these genes/transcripts.

Discussion

Up until 2002, genetic linkage studies of CVID and/or IgAD concentrated on the HLA region on chromosome 6 using a variety of study designs.¹²⁻¹⁹ Based on the positive results of Vořechovský *et al*,¹⁸ the HLA region was designated as the susceptibility locus *IGAD1*, and fine-mapped in a follow-up study.⁴⁹ There is substantial disagreement in the cited studies regarding where in the HLA region a putative disease-causing gene is located, if there is one, and whether it is a susceptibility locus for CVID or IgAD or both.⁵⁰⁻⁵²

In 2003, we reported on an autosomal genome-wide linkage scan of three families with multiple cases of CVID, IgAD, and dysgammaglobulinemia, exhibiting autosomal-dominant inheritance and reported linkage of the disease phenotype to the telomeric region of chromosome 5p.²⁰ The positive linkage evidence in this study came primarily from one large family. The 5p locus is not replicated in the EU cohort (data not shown) and clinical follow-up suggests that some members of the large 5p family have mycobacterial tuberculosis, an infection that is atypical for CVID.

Subsequent research has given strong support to the hypothesis that small numbers of families originally classified as multiplex CVID families have detectable monogenic immunodeficiencies, which are difficult to identify *a priori* by immunological assays alone. Genetic defects identified so far in patients with CVID include mutations in *ICOS* on chromosome 2q,²³ *CD19* on chromosome 16p,²⁴ *TNFRSF13C* encoding the BAFF receptor on chromosome 22q (Warnatz K *et al*. *XIth ESID Meeting*,

Versailles 2004; Abstract #B72), and *TNFRSF13B* encoding TACI on chromosome 17p.^{21,22} With the possible exception of mutations in *TNFRSF13B*, which show autosomal-dominant transmission, these defects all display pure autosomal-recessive traits and heterozygous individuals appear to be healthy.

In the studies listed in the previous paragraph, positional reasoning was either not used at all or only used to assess whether marker and phenotype data in a specific family could be consistent with a mutation in a nearby gene. In contrast, for the NL1 family upon which we report in this study, we started with a positional approach. As a consequence, candidate gene sequencing can restrict attention to approximately 1% of the human genome. Thus, the discovery of a disease-associated genetic locus in multiplex CVID/IgAD families with an autosomal-dominant trait contributes to the search of disease genes causing hypogammaglobulinemia.

CVID is a complex disease. The previous identification of at least four distinct genes mutated in CVID establishes that there is considerable locus heterogeneity. A few families in these studies have nonpenetrant individuals, strongly suggesting a polygenic effect creating more complexity. The interaction between the genome of the human host and the pathogens that cause the infections in CVID adds another layer of complexity. However, genetic studies may gradually change CVID from one common diagnosis of exclusion to numerous rarer diagnoses of inclusion (eg, ICOS deficiency, CD19 deficiency, TACI deficiency, BAFFR deficiency). Since the age of onset of CVID is often in adulthood, and the diagnosis often delayed due to the insidious onset of symptoms,² a genetic understanding of the etiology offers hope for earlier diagnosis and intervention.

The autosomal-dominant inheritance of CVID/IgAD in the largest pedigree published to date will most likely be explained by a mutation on chromosome 4q. A replication study in a collection of 32 smaller CVID/IgAD families suggests that some of these patients may also have a mutation in the same gene. It will be difficult to identify the gene solely by positional reasoning (ie genetic linkage or linkage disequilibrium) because the linkage interval is wide and other large CVID/IgAD families are hard to find. One approach we are currently pursuing is to determine by gene expression analysis which of the transcripts located within the linkage interval are differentially expressed in affected members of family NL1 when compared to healthy family members.

Acknowledgements

We would like to acknowledge the help of Dr Tom Nijenhuis, Dr Corrie MR Weemaes, Dr Ina Klasen, and Dr Anna Simon in conducting the clinical sample collection and the support of Dr Ulrich Salzer in the laboratory. We thank an anonymous reviewer for very helpful suggestions that improved the manuscript. We thank the families

and their physicians for their participation in this research study supported by Grant SFB620/C2 to BG from the Deutsche Forschungsgemeinschaft, by a grant from the Swedish research council to LH, by the Intramural Program of the National Institutes of Health, by the Fondazione Camillo Golgi, Brescia, Italy, and by EU-Projects QLRT-2001-01536 (IMPAD) and SP23-CT-2005-006411 (EURO-Policy-PID).

References

- Cunningham-Rundles C: Common variable immunodeficiency. *Curr Allergy Asthma Rep* 2001; **1**: 421–429.
- Webster ADB: Common variable immunodeficiency. *Immunol Allergy Clin* 2001; **21**: 1–22.
- IUIS: Primary immunodeficiency diseases. Report of an IUIS Scientific Committee. International Union of Immunological Societies. *Clin Exp Immunol* 1999; **118** (Suppl 1): 1–28.
- Hammarström L, Vořechovský I, Webster D: Selective IgA deficiency (SIgAD) and common variable immunodeficiency (CVID). *Clin Exp Immunol* 2000; **120**: 225–231.
- Vořechovský I, Zetterquist H, Paganelli R *et al*: Family and linkage study of selective IgA deficiency and common variable immunodeficiency. *Clin Immunol Immunopathol* 1995; **77**: 185–192.
- Wilton AN, Cobain TJ, Dawkins RL: Family studies of IgA deficiency. *Immunogenetics* 1985; **21**: 333–342.
- Ashman RE, Schaffer FM, Kemp JD *et al*: Genetic and immunologic analysis of a family containing five patients with common variable immune deficiency or selective IgA deficiency. *J Clin Immunol* 1992; **12**: 406–414.
- Ishizaka A, Nakanishi M, Yamada S, Sakiyama Y, Matsumoto S: Development of hypogammaglobulinaemia in a patient with common variable immunodeficiency. *Eur J Pediatr* 1989; **149**: 175–176.
- Español T, Catala M, Hernandez M, Caragol I, Bertran JM: Development of a common variable immunodeficiency in IgA-deficient patients. *Clin Immunol Immunopathol* 1996; **80**: 333–335.
- Johnson ML, Keeton LG, Zhu ZB, Volanakis JE, Cooper MD, Schroeder Jr HW: Age-related changes in serum immunoglobulins in patients with familial IgA deficiency and common variable immunodeficiency (CVID). *Clin Exp Immunol* 1997; **108**: 477–483.
- Carvalho Neves Forte W, Ferreira De Carvalho Jr F, Damaceno N, Vidal Perez F, Gonzales Lopes C, Mastrotti RA: evolution of IgA deficiency to IgG subclass deficiency and common variable immunodeficiency. *Allergol Immunopathol* 2000; **28**: 18–20.
- Schaffer FM, Palermos J, Zhu ZB, Barger BO, Cooper MD, Volanakis JE: Individuals with IgA deficiency and common variable immunodeficiency share polymorphisms of major histocompatibility complex class III genes. *Proc Natl Acad Sci USA* 1989; **86**: 8015–8019.
- Olerup O, Smith CIE, Hammarström L: Different amino acids at position 57 of the HLA-DQ beta chain associated with susceptibility and resistance to IgA deficiency. *Nature* 1990; **347**: 289–290.
- Olerup O, Smith CI, Björkander J, Hammarström L: Shared HLA class II-associated genetic susceptibility and resistance, related to the HLA-DQB1 gene, in IgA deficiency and common variable immunodeficiency. *Proc Natl Acad Sci USA* 1992; **89**: 10653–10657.
- Volanakis JE, Zhu ZB, Schaffer FM *et al*: Major histocompatibility complex class III genes and susceptibility to immunoglobulin A deficiency and common variable immunodeficiency. *J Clin Invest* 1992; **89**: 1914–1922.
- Cucca F, Zhu Z-B, Khanna A *et al*: Evaluation of IgA deficiency in Sardinians indicates a susceptibility gene is encoded within the HLA class III region. *Clin Exp Immunol* 1998; **111**: 76–80.
- Schroeder Jr HW, Zhu ZB, March RE *et al*: Susceptibility locus for IgA deficiency and common variable immunodeficiency in the HLA-DR3, -B8, -A1 haplotypes. *Mol Med* 1998; **4**: 72–86.

- 18 Vořechovský I, Webster ADB, Plebani A, Hammarström L: Genetic linkage of IgA deficiency to the major histocompatibility complex: evidence for allele segregation distortion, parent-of-origin penetrance differences, and the role of anti-IgA antibodies in disease predisposition. *Am J Hum Genet* 1999; **64**: 1096–1109.
- 19 Vořechovský I, Cullen M, Carrington M, Hammarström L, Webster ADB: Fine mapping of IGAD1 in IgA deficiency and common variable immunodeficiency: identification and characterization of haplotypes shared by affected members of 101 multiple-case families. *J Immunol* 2000; **164**: 4408–4416.
- 20 Braig DU, Schäffer AA, Glocker E *et al*: Linkage of autosomal common variable immunodeficiency to chromosome 5p and evidence for locus heterogeneity. *Hum Genet* 2003; **112**: 369–378.
- 21 Salzer U, Chapel HM, Webster ADB *et al*: Mutations in *TNFRSF13B*, which encodes TACI, are associated with common variable immunodeficiency in humans. *Nature Genet* 2005; **37**: 820–828.
- 22 Castigli E, Wilson SA, Garibyan L *et al*: TACI is mutant in common variable immunodeficiency and IgA deficiency. *Nat Genet* 2005; **37**: 829–834.
- 23 Grimbacher B, Hutloff A, Schlesier M *et al*: Homozygous loss of ICOS is associated with adult-onset common variable immunodeficiency. *Nat Immunol* 2003; **4**: 261–268.
- 24 van Zelm MC, Reisl I, van der Burg M *et al*: Novel antibody deficiency in patients with CD19 gene defects. *N Engl J Med* 2006 (in press).
- 25 Morra M, Silander O, Calpe S *et al*: Alterations of the X-linked lymphoproliferative disease gene SH2D1A in common variable immunodeficiency syndrome. *Blood* 2001; **98**: 1321–1325.
- 26 Nistala K, Gilmour KC, Cranston T *et al*: X-linked lymphoproliferative disease: three atypical cases. *Clin Exp Immunol* 2001; **126**: 126–130.
- 27 Bruton OC: Agammaglobulinemia. *Pediatrics* 1952; **9**: 722.
- 28 Saffran DC, Parolin O, Fitch-Hilgenberg ME *et al*: A point mutation in the SH2 domain of Bruton's tyrosine kinase in atypical X-linked agammaglobulinemia. *New Engl J Med* 1994; **330**: 1488–1491.
- 29 Smith CIE, Islam KB, Vořechovský I *et al*: X-linked agammaglobulinemia and other immunoglobulin deficiencies. *Immunol Rev* 1994; **138**: 159–183.
- 30 Nijenhuis T, Klasen I, Weemaes CM, Preijers F, de Vries E, van der Meer JW: Common variable immunodeficiency (CVID) in a family: an autosomal dominant mode of inheritance. *Neth J Med* 2001; **59**: 134–139.
- 31 Cunningham-Rundles C, Bodian C: Common variable immunodeficiency: clinical and immunological features of 248 patients. *Clin Immunol* 1999; **92**: 34–48.
- 32 Broman KW, Murray JC, Sheffield VC, White RL, Weber JL: Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am J Hum Genet* 1998; **63**: 861–869.
- 33 Lathrop GM, Lalouel J-M, Julier C, Ott J: Strategies for multilocus analysis in humans. *Proc Natl Acad Sci USA* 1984; **81**: 3443–3446.
- 34 Cottingham Jr RW, Idury RM, Schäffer AA: Faster sequential genetic linkage computations. *Am J Hum Genet* 1993; **53**: 252–263.
- 35 Schäffer AA, Gupta SK, Shriram K, Cottingham Jr RW: Avoiding recomputation in linkage analysis. *Hum Hered* 1994; **44**: 225–237.
- 36 O'Connell JR, Weeks DE: The VITESSE algorithm for rapid exact multilocus linkage analysis via genotype set-recoding and fuzzy inheritance. *Nat Genet* 1995; **11**: 402–408.
- 37 Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; **58**: 1347–1363.
- 38 Kruglyak L, Lander ES: Faster multipoint linkage analysis using Fourier transforms. *J Comp Biol* 1998; **5**: 1–7.
- 39 Ott J: *Analysis of Human Genetic Linkage*, revised ed. Baltimore: Johns Hopkins University Press, 1991.
- 40 Ott J: Computer-simulation methods in human linkage analysis. *Proc Natl Acad Sci USA* 1989; **86**: 4175–4178.
- 41 Weeks DE, Ott J, Lathrop GM: SLINK: a general simulation program for linkage analysis. *Am J Hum Genet* 1990; **47**: A204.
- 42 Terwilliger JD: A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am J Hum Genet* 1995; **56**: 777–787.
- 43 Di Renzo M, Pasqui AL, Auteri A: Common variable immunodeficiency: a review. *Clin Exp Med* 2004; **3**: 211–217.
- 44 Fournier E, Isakoff SJ, Ko K *et al*: The B cell SH2/PH domain-containing adaptor Bam32/DAPP1 is required for T cell-independent II antigen responses. *Curr Biol* 2003; **13**: 1858–1866.
- 45 Yokoyama K, Su I, Tezuka T *et al*: BANK regulates BCR-induced calcium mobilization by promoting tyrosine phosphorylation of IP₃ receptor. *EMBO J* 2002; **21**: 83–92.
- 46 Karin M, Ben-Neriah Y: Phosphorylation meets ubiquitination: the control of NF κ B activity. *Ann Rev Immunol* 2000; **18**: 621–663.
- 47 Murray JC, Heng YM, Symonds P *et al*: Endothelial monocyte-activating polypeptide-II (EMAP-II): a novel inducer of lymphocyte apoptosis. *J Leukocyte Biol* 2004; **75**: 772–776.
- 48 Schmeck B, Gross R, N'Guessan PD *et al*: *Streptococcus pneumoniae*-induced caspase 6-dependent apoptosis in lung epithelium. *Infect Immun* 2004; **72**: 4940–4947.
- 49 Kralovicova J, Hammarström L, Plebani A *et al*: Fine-scale mapping at IGAD1 and genome-wide genetic linkage analysis implicate HLA-DQ/DR as a major susceptibility locus in selective IgA deficiency and common variable immunodeficiency. *J Immunol* 2003; **170**: 2765–2775.
- 50 De la Concha EG, Fernandez-Arquero M, Gual L *et al*: MHC susceptibility genes to IgA deficiency are located in different regions on different HLA haplotypes. *J Immunol* 2002; **169**: 4637–4643.
- 51 Schroeder Jr HW, Schroeder III HW, Sheikh SM: The complex genetics of common variable immunodeficiency. *J Invest Med* 2004; **52**: 90–103.
- 52 Grimbacher B, Schaffer AA, Peter HH: The genetics of hypogammaglobulinemia. *Curr Allergy Asthma Rep* 2004; **4**: 349–358.