LETTER TO JMG

A genome screen of families at high risk for Hodgkin lymphoma: evidence for a susceptibility gene on chromosome 4

L R Goldin, M L McMaster, M Ter-Minassian, S Saddlemire, B Harmsen, G Lalonde, M A Tucker

J Med Genet 2005;**42**:595–601. doi: 10.1136/jmg.2004.027433

Hodgkin's disease was recently designated Hodgkin lymphoma (HL) in the World Health Organization Classification.[1] The National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) population based registries estimate that 7900 new cases are diagnosed annually in the USA.[2] Clues to its aetiology have been suggested by the bimodal age distribution; higher risks in males, in people with higher socioeconomic status, and in smaller families; and occurrence of Epstein-Barr virus in HL tumour cells.[3] The importance of genetic factors is indicated by reports of multiply affected families from case series,[4–6] a twin study,[7] a case–control study,[8] and population registry studies carried out in Utah,[9] Denmark,[10] Israel,[11] and Sweden.[12–14] We recently analysed data from registries in Sweden and Denmark and found significant familial aggregation of HL and other lymphoproliferative tumours.[15] The relative risk for HL among first degree relatives of cases compared with controls was 3.1. Relative risks were higher in males compared with females, and in siblings of cases compared with parents and offspring. Relatives of earlier onset cases were at higher risk for HL and for all lymphoproliferative tumours and were also at higher risk for developing early onset tumours themselves. These findings are consistent with those seen from earlier case series studies but have the advantage of being from large, population based samples.

It is not known whether or how extrinsic risk factors interact with genetic susceptibility. Identifying inherited susceptibility genes is an important step towards defining the pathway(s) leading to development of HL and understanding its complex aetiology. There have been many studies of somatic mutations in HL tumour cells, but although there are associations with HLA types, specific germline genes causing susceptibility have not yet been identified. Early studies of HLA Class I alleles in familial HL showed increased haplotype sharing among affected sibling pairs.[16 17] We have previously conducted studies of HLA Class II loci in 16 high risk HL families and found that alleles reported to be associated in case–control studies (such as DRB1*1501 and DQB1*0602) were also associated with familial HL using a family based analytical approach.[18] There have been no comprehensive searches of the genome for HL genes, largely due to the difficulty in assembling informative samples. Even though this tumour is strongly familial, the proportion of cases with a family history is small, and affected families typically have very few cases.

We studied 44 informative high risk HL families and applied a whole genome search using densely spaced microsatellite markers in order to localise susceptibility genes.

Key points

- Hodgkin lymphoma (HL) has a strong familial component but no genes have yet been identified.
- We performed a genomewide linkage screen in 44 high risk HL families with a total of 254 individuals with DNA samples. Among these families, there were 95 HL cases and four cases of non-Hodgkin lymphoma (NHL) who were informative for linkage. The cases were characterised by a young age at diagnosis and an even gender ratio. In two-thirds of the families, the cases were siblings or cousins.
- We genotyped 1058 microsatellite markers with an average spacing of 3.5 cM, analysed the data using both non-parametric and parametric linkage analysis, and computed both two point and multi-point linkage statistics.
- The strongest linkage finding was on chromosome 4p near the marker D4S394. The lod score calculated by Genehunter Plus was 2.6 (nominal p = 0.0002) when both HL and NHL individuals were considered affected. The mean identity by descent sharing among 35 affected sibling pairs was 72% in this region (nominal p = 0.00007).
- The results are consistent with recessive inheritance. Other locations suggestive of linkage were found on chromosomes 2 and 11. The number of independent regions identified is more than expected by chance, although no one region met genomewide significance levels.
- These linkage findings represent the first step towards identifying one or more loci leading to susceptibility to HL and understanding its complex aetiology.

METHODS

Ascertainment of HL pedigrees

The Genetic Epidemiology Branch (Division of Cancer Epidemiology and Genetics) has been recruiting families with two or more living cases of HL since 1970. This study was approved by an institutional review board, and informed consent was obtained on all subjects in this report. At the NIH clinical centre or on field trips, we evaluated all available affected individuals and first degree relatives of those affected, and obtained biospecimens. We also obtained original pathology material and reports for all HL and NHL cases where possible, and these were reviewed by the National Cancer Institute Laboratory of Pathology. Of the families investigated, 44 were judged to be informative for linkage studies, based on the number of available DNA samples (total 254) from affected and unaffected individuals. Sixteen of these families had been included in an earlier study of linkage and association with the HLA region.[18]

Table 1 shows the distribution of the number of affected individuals per family and the relationships among the HL cases. A total of 106 individuals in the families have been diagnosed with HL. Of these, DNA samples were available for 89, and genotypes for an additional six cases could be inferred from other family members. Of eight cases of NHL, four had either DNA samples or inferable genotypes. The level of diagnostic certainty for the HL cases was high, with 85% confirmed by either an outside pathology report, a slide reviewed at the NCI Laboratory of Pathology, or both. All cases were considered ''affected'' for linkage analysis. The mean age at diagnosis of HL was 26.8 years, which is much lower than that in the population, where the median age at diagnosis is 37 years.[2] Over 90% of our cases would be considered as having onset in childhood or young adulthood (earlier than 45 years of age) There was no difference in age at diagnosis among families with only two HL cases compared with those with more than two cases. Among those cases who could be classified into subtypes (75% of the total), there was a predominance (80%) of the nodular sclerosis (NS) subtype, with nearly all of the remaining having the mixed cellularity subtype, consistent with the young age distribution. There was a slight female predominance. As can be seen in table 1, two thirds of the families had cases among siblings and/or cousins; the remainder showed parent–offspring configurations with or without siblings or other relatives.

Genotyping

DNA was extracted from cryopreserved lymphocytes using standard methods. Genotyping was conducted under contract to deCODE Genetics using their screening set of 1058 microsatellite markers containing markers from the ABI linkage marker (version 2) screening and intercalating sets, and 500 custom made markers with known allele size distributions. Marker positions were obtained from the deCODE genetic map.[19] PCR reactions were set up in multiplex reactions with fluorescently labelled primer pairs selected to amplify highly informative two, three, and four microsatellite loci. Following PCR amplification, DNA samples were loaded into ABI 3730 capillary sequencers. In each 96 well DNA plate, 93 DNA samples and three CEPH controls (family 1347-2) were run. Alleles were automatically classified using deCODE Allele Caller software,[20] which provides consistently >99.7% accuracy of genotyping calls compared with manual procedures. The samples were barcoded and tracked at each step, and profiled with sufficient markers for unique identification. Sample identities were checked for accuracy based on the pedigree structures in order to identify sample duplications and exchanges. The extensive use of robotics and automation at all steps in the process provided a high degree of reliability, and reduced sample handling errors. A multistage data analysis approach was used to minimise errors in genotyping. After initial genotype identification was made, analyses were conducted to detect non-Mendelian transmission of genotypes from parents to offspring. These errors were then checked by re-analysing the results from the ABI sequencer or by re-typing the samples.

Table 1 Description of HL families

	No. of families
Family types	
2 HL cases	29
3 HL cases	12
4 HL cases	3
HL families with at least 1 NHL* case	8
Relationships among HL cases	
Siblings	22
Siblings + parent	5
Siblings + cousin or avuncular	6
Parent–offspring	8
Parent–offspring + cousin or avuncular	1
Cousins	2

*5/8 NHL cases diagnosed with diffuse large cell lymphoma.

Power analysis

Because the true genetic model for HL is not known, we computed power assuming a rare gene with either dominant or recessive inheritance and heterogeneity. We estimated the power to detect linkage using the program SLINK.[21 22] We conducted simulations assuming both dominant and recessive inheritance models and penetrances of both 50% and 80% for the at risk genotypes and 0.1% for the normal genotype (a total of four models) Allele frequencies were set at values that kept the lifetime risk constant at 0.24 as estimated from SEER data.[2] We assumed close linkage ($\theta = 0.001$) of the disease locus to a marker locus with eight alleles, which is reasonable given the dense spacing and high information content of the real genotypes. We generated 200 replicates under each model to compute the average lod scores and power of detecting lod scores of 1–3.

Linkage analysis

The genotype data were stored in a Microsoft Access database. Formatting changes needed for specific programs were made using MEGA2.[23] The genotype data were first checked for Mendelian consistency using the program PEDCHECK.[24] The RECODE program was used to prepare the data files for analysis and to estimate allele frequencies from all founders in the pedigrees. We checked for the presence of additional genotype errors using the mistyping option of Simwalk2 (version 2.89),[25–27] and eliminated genotypes that had probabilities of ⩾0.25 of being errors. In total, only a very small number of genotypes (<0.5%) was eliminated because of either Mendelian inconsistencies or high mistyping probability.

We first screened the 1058 markers using two point analyses with the MLINK program from the FASTLINK package.[28–31] We calculated lod scores assuming the same models as described above for the power simulations. In addition, we assumed that penetrance increased with age, using age incidence rates in the population to construct liability classes. Multipoint analyses were conducted using Genehunter[32] to compute both parametric lod scores, assuming heterogeneity (Hlod, with α = proportion of linked families) and non-parametric linkage (NPL) scores (z scores). Genehunter Plus lod scores were also calculated, because this method has been shown to give less conservative estimates of p values than does the original Genehunter method.[33] As there were only a few NHL cases in this familial cohort, only individuals with HL were classified as affected for initial linkage analyses across the genome, and all other individuals were considered unaffected. Regions of the genome with nominal p values ⩽0.01 by any analysis method were followed up with additional analyses, including broadening the affection status to include NHL and calculation of mean IBD sharing in affected sibling pairs (ASPs) using the program Sibpal in SAGE (version 4.5).[34]

RESULTS

Power of linkage detection

The simulations showed that if ⩾75% of the families were linked to the same locus, then the probability of obtaining a lod score of 3.0 or more was at least 75%. The power to find a recessive gene was always higher than for a dominant gene; this is not surprising, as two thirds of the families have

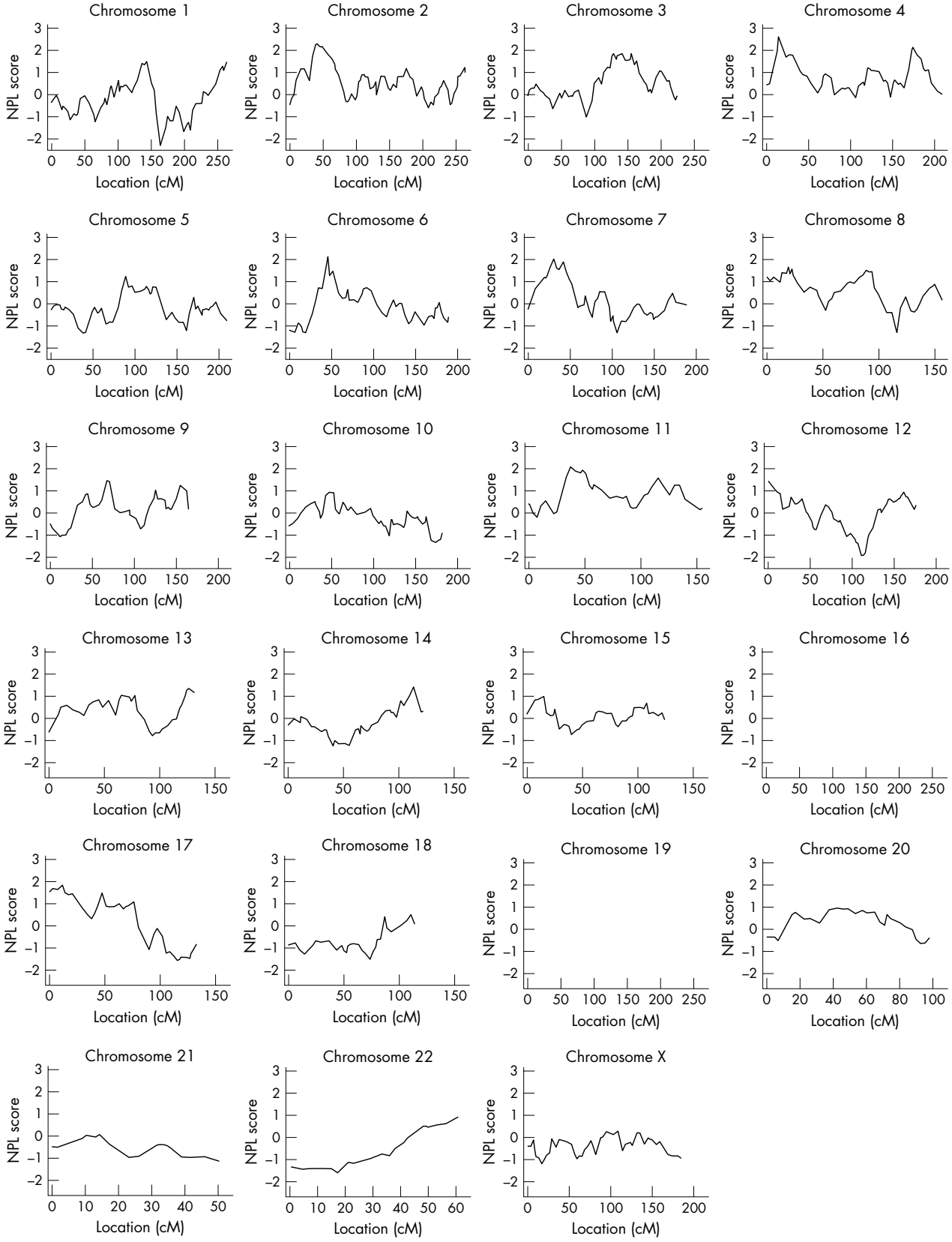

Figure 1 Multipoint non-parametric linkage (NPL) statistics calculated by Genehunter for each chromosome assuming only HL cases affected.

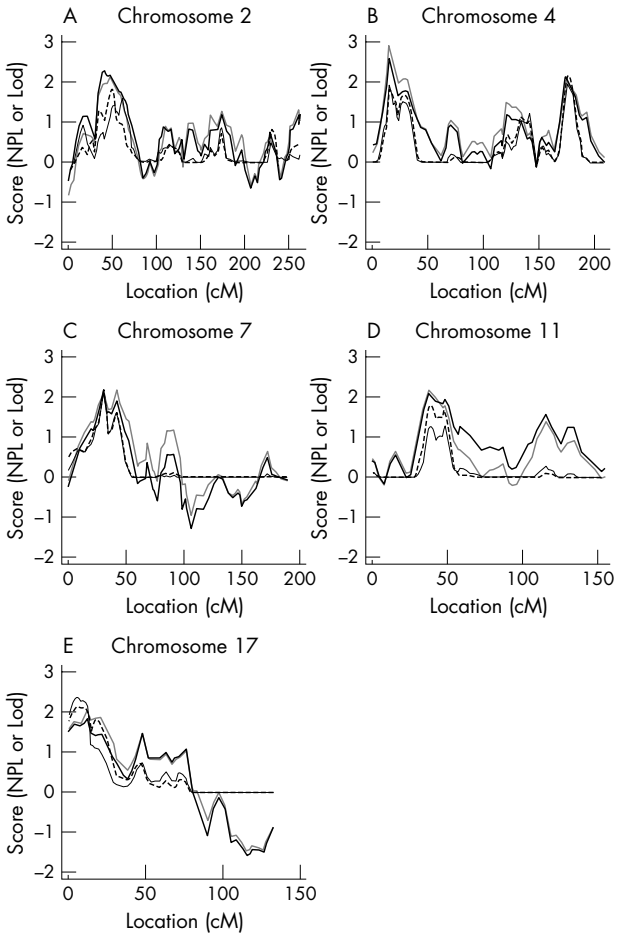

Figure 2 (A–E) Multipoint NPL and heterogeneity lod scores (Hlod) for chromosomes with suggestive results. Thick black line, NPL (HL affected);thin black line, Hlod (HL affected); thick grey line, NPL (HL or NHL affected); broken line, Hlod (HL or NHL affected). Model assumed for Hlod score was recessive with maximum penetrance of 50% for chromosomes 2, 4, 7, and 11, and dominant with maximum penetrance of 50% for chromosome 17.

affected sibling or cousin configurations and are thus consistent with recessive inheritance. If only half of the families were linked to a single locus, then the power to detect linkage was modest (expected lod score = 2.3 under recessive inheritance, 1.9 under dominant inheritance). If only 25% of the families were linked, we would have minimal power to detect a susceptibility gene.

Linkage results

Two point lod scores revealed several regions of the genome with evidence for linkage to HL. The strongest findings were in regions on chromosomes 2, 3, and 4. Table 2 shows loci that had two point lod scores ⩾2.0 under any one of the four inheritance models tested. These regions, on chromosomes 2, 3, and 4, showed clusters of consecutive markers with positive scores. Another region on chromosome 4 and a region on chromosome 11 also showed clusters of consecutive loci with positive scores, with maximum scores between 1.5 and 2.0 (not shown). Multipoint NPL statistics of all of the chromosomes as calculated by Genehunter are shown in fig. 1. Parametric lod scores with and without heterogeneity were also calculated but are not shown. The densely spaced markers resulted in a high level of informativeness, with information content calculated by Genehunter averaging 0.75 throughout the genome. Fig 2 (A–E) shows regions on chromosomes 2, 4, 7, 11, and 17 where either multipoint NPL or Hlod scores had nominal p values <0.01. These figures show NPL and parametric statistics for both narrow (HL only) and broad (including NHL) affection status models. The parametric lod scores were calculated assuming heterogeneity using the inheritance model (dominant or recessive) that gave the highest two point lod scores. As seen in table 2, one marker on chromosome 3 had a two point lod score of 4.0 under dominant inheritance. The flanking markers had lods of 1.7 and 1.0 (not shown), but multipoint statistics in this region were substantially lower (highest NPL score was <2.0 and highest Hlod score was 1.4) than the two point results; thus no additional graphs are shown. Tables 3 and 4 summarise the linkage statistics, locations, and marker names for these six regions.

The strongest evidence for linkage occurred on chromosome 4 where the peak NPL or Hlod was found at 14 cM, flanked by markers D4S2935 and D4S394 (fig 2B). Strong evidence for linkage was seen under both affection status models. Under the broader model (in which NHL cases counted as affected), the peak NPL score was 2.9 and peak Genehunter Plus lod score was 2.6 ($p = 0.0002$), which is strongly suggestive of linkage (table 3). The locations of the peak linkage scores were consistent among all analyses performed. The Hlod score was highest under the recessive model and the proportion of families linked was estimated at 43%. Table 4 shows that the mean IBD sharing among ASPs as calculated by Sibpal (>70% and highly significant) gave results consistent with the other methods.

Fig 2 (A–E) and tables 3 and 4 show five other locations that had positive non-parametric linkage or Hlod scores. For each region, the positive findings from Genehunter were supported by increased mean IBD sharing among ASPs (table 4). The second most significant region was on chromosome 2 (Genehunter Plus lod score was 2.4, $p = 0.0004$), although the location varied from 41 to 62 cM (fig 2A). On chromosome 11, there was a peak at location 37 to 39 cM (fig 2D), depending on the model, which had a maximum Genehunter Plus lod of 2.2 (rounded from 2.18) and $p = 0.0007$. Regions on 4q, 7, and 17 also showed positive results. There was a second positive peak on chromosome 4 at location 173–176 cM (fig 2B), a peak on chromosome 7 (fig 2C), and one on chromosome 17, in which the peak was at the p telomeric region (fig 2E). There was also a modest signal (NPL score of ~2.0) on chromosome 6 at the marker

Table 2 Summary of two point lods ⩾2.0

Chrom	Locus	Location (cM)	Recessive 50% penetrance lod (theta)	Dominant 50% penetrance lod (theta)
2	D2S2168	49.4	2.52 (0.10)	1.41 (0.10)
3	D3S1558	128.9	2.52 (0.10)	4.03 (0.01)
4	D4S2935	13.3	2.09 (0.05)	0.91 (0.10)
4	D4S394	15.3	2.39 (0.10)	0.15 (0.25)

Chrom, chromosome number.

Table 3 Linkage signals from multipoint analysis

	Affected HL only			Affected HL or NHL		
Chrom.	Statistic	Location (cM)	Markers	Statistic	Location (cM)	Markers
2						
NPL	2.20	40	D2S2346	2.11	47	D2S305–D2S2168
GHP lod	2.44	41	D2S2346	2.10	45	D2S305–D2S2168
Hlod (α)	1.72 (0.48)	61	D2S2230	1.81 (0.46)	50	D2S2168
4p						
NPL	2.59	14	D4S2935–D4S394	2.91	14	D4S2935–D4S394
GHP lod	2.51	14	D4S2935–D4S394	2.63	14	D4S2935–D4S394
Hlod (α)	2.15 (0.50)	14	D4S2935–D4S394	1.78 (0.45)	15	D4S2935–D4S394
4q						
NPL	2.10	174	D4S3030	2.05	174	D4S3030
GHP lod	1.37	174	D4S3030	1.27	174	D4S3030
Hlod (α)	2.20 (0.46)	176	D4S3030–D4S415	2.20 (0.43)	176	D4S3030–D4S415
7						
NPL	2.10	30	D7S2557–D7S2495	2.10	30	D7S2557–D7S2495
GHP lod	1.33	30	D7S2557–D7S2495	1.35	41	D7S2463–D7S516
Hlod	2.43 (0.45)	30	D7S2557–D7S2495	2.27 (0.43)	30	D7S2557–D7S2495
11						
NPL	2.09	37	D11S928–D11S1359	2.18	37	D11S928–D11S1359
GHP lod	2.18	38	D11S1359	2.16	39	D11S1359
Hlod	1.34 (0.43)	38	D11S1359	1.78 (0.43)	39	D11S1359
17						
NPL	1.83	12	D17S1876	2.03	12	D17S1876
GHP lod	0.90	6	D17S1529–D17S831	1.06	12	D17S1876
Hlod (α)	2.37 (0.64)	6	D17S1529–D17S831	2.14 (0.53)	5	D17S1529–D17S831

Chrom, chromosome number; GHP, Genehunter Plus.

Table 4 Maximum IBD sharing among affected sibling pairs (and p values) in six regions suggestive of linkage

	Affected HL only (32 affected sibling pairs)			Affected HL or NHL (35 affected sibling pairs)		
Chrom.	Marker (cM)	IBD	p	Marker (cM)	IBD	p
2	D2S305 (43)	0.67	0.00029	D2S2168 (49)	0.69	0.00021
4p	D4S394 (15)	0.73	0.00002	D4S394 (15)	0.72	0.000073
4q	D4S3030 (175)	0.66	0.007	D4S3030 (175)	0.63	0.017
7	D7S2495 (34)	0.66	0.007	D7S2495 (34)	0.64	0.0099
11	D11S1359 (39)	0.67	0.004	D11S1359 (39)	0.68	0.0013
17	D17S960 (21)	0.64	0.005	D17S960 (21)	0.66	0.001

Chrom, chromosome number.

D6S1571, which is close to the HLA region (fig 1); however, no other markers on this chromosome gave positive results. This is consistent with a small HLA effect in HL and with previous data showing an association with HLA types within a subset of these families.[18]

DISCUSSION

Our unique sample of high risk HL families has allowed us to conduct a genomewide scan using a dense set of markers. The strongest finding in this study was on chromosome 4p. The inheritance is likely to be recessive, given that the highest Hlod scores were found assuming a recessive model and affected sibling pairs had a mean IBD >70%. The sharing in ASPs was highly significant although the sample size was not large. The likelihood of recessive inheritance is also supported by previous studies showing higher risks in siblings.[4 15] The strength of linkage to this region increased slightly when the few individuals with NHL were considered affected. This is consistent with our population study showing that both HL and NHL were found more frequently in relatives of HL cases.[15] In fact, the cases of NHL in our families had a much earlier age of onset (mean 49 years) than the population (median age of diagnosis is 67 years[2]). This location on chromosome 4p is a high priority for follow up with additional families. The regions on chromosomes 2 and 11, with Genehunter Plus lod scores of 2.4 and 2.2, are also strong candidates for containing HL susceptibility genes. The three other findings on chromosomes 4q, 7, and 17 have lower significance levels but warrant further follow up. The promising finding on chromosome 3 based on two point lod scores did not hold up under multipoint analyses. Consistent with other data, we also have evidence from this study that the HLA region may play a role in familial risk (fig 1). It is possible that applying more complex modelling of gene effects (such as multilocus models) to the genomewide linkage data would lead to more definitive detection of a susceptibility gene or genes.

There is some disagreement about how to interpret significance levels when conducting a genomewide scan.[35] Wiltshire *et al*[36] suggested an approach for evaluating the significance of linkage findings. They pointed out that for complex diseases, several genes may be involved and therefore it is less likely that any single region will reach a high level of significance. They propose counting the number of independent regions of linkage detected and comparing this to chance expectations. They pointed out that the lod score thresholds for "significant" (lod = 3.6) or "suggestive" (lod = 2.2) linkage, as defined by Lander and Kruglyak,[37]

are often too stringent, as the thresholds for any one study depend on sample size, marker density, and marker informativeness. Using simulations, Wiltshire *et al*[36] found that for a 5 cM scan with 100 ASP families, a Genehunter Plus lod score of 1.78 would predict one linkage finding in a genomewide scan by chance and a more stringent lod score threshold of 2.2 would predict 0.37 linkage findings by chance. Our three strongest findings on chromosomes 4p, 2p, and 11p meet the more stringent threshold of 2.2 (table 3). Thus, we have identified more regions than expected by chance, which strongly suggests that there are one or more true loci causing susceptibility to HL among the locations identified.

There are several regions of the genome where recurrent cytogenetic changes are found in HL cells,[38 39] including amplifications of regions of 2p, 4p, 4q, and 9p, and deletions of chromosome 6q25.[40] The regions we identified by linkage do not appear to overlap with the cytogenetic regions. For example, duplication of the c-*REL-BCL11A* region on 2p was reported in HL cells, but these loci are about 30–50 cM centromeric from the peak we found. Similarly, amplifications on 4p may involve the fibroblast growth factor 3 gene,[39] which is about 10 Mb from the peak we identified. The linkage peak on the telomere of chromosome 17 is near the *p53* gene, but somatic mutations of this gene are not frequently found in HL cells,[41] and one study found no germline mutations in familial HL cases from our sample.[42] Locations determined from linkage analysis of complex diseases are imprecise owing to uncertainty about the underlying model, so it is possible that one or more of the regions identified by cytogenetic studies overlaps with regions we identified by linkage.

There are some limitations to this study. The highly selected families in our sample are not representative of HL in the population. Consistent with clinical descriptions of familial HL in the literature,[43] referrals to our group are mostly families with cases that have early onset and NS subtype. There are a few later onset cases in these families, but even these are found within families having early onset in other members. In terms of histological subtype, we also find mixed cellularity cases in the same families with NS subtype. Thus, it is not possible for us to analyse linkage to subgroups based on age at onset or histological characteristics of the tumour. In addition, most of our families have HL in siblings and/or cousins, which makes it difficult to detect a dominant susceptibility locus. Future studies applied to more families and a broader representation of clinical types will lead to more robust conclusions about the effects of susceptibility genes, genetic heterogeneity of HL, and the range of phenotypic expression of specific susceptibility genes.

The findings presented here are the first step in the discovery of germ line susceptibility gene(s) and delineation of the pathways involved in development of HL. Even though these susceptibility loci are being discovered in high risk families, they may also play a role in development of sporadic HL. Defining these pathways and determining their interactions with environmental factors may lead to more effective treatment and prevention, which could have a great impact on patients, many of who are young and lose years of life/productivity to disease or treatment related morbidity.

ACKNOWLEDGEMENTS

We acknowledge the contributions of A Goldstein, L Harty, A Lin, J Whitehouse, and G Shaw for family ascertainment, and of E Jaffe for pathology review. The genotyping was conducted under NCI contract N02-CP-01108 with a subcontract BRC-1108-35 to deCODE Genetics. Some of the results of this paper were obtained by using the program package SAGE, which is supported by a US Public Health Service Resource Grant (RR03655) from the National Center for Research Resources.

Authors' affiliations

L R Goldin, M L McMaster, M Ter-Minassian, S Saddlemire, M A Tucker, Genetic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA
B Harmsen, G Lalonde, Westat Research Inc., Rockville, MD, USA

Competing interests: none declared

Correspondence to: Dr L R Goldin, Genetic Epidemiology Branch, DCEG, NCI, 6120 Executive Blvd, 7236 MSC, Bethesda, MD 20892-7236, USA; goldinl@mail.nih.gov

Received 22 September 2004
Revised 4 November 2004
Accepted 4 November 2004

REFERENCES

1 **Jaffe ES**, Harris NL, Stein H, Vardiman JW, eds. *World Health Organization classification of tumours. Pathology and genetics of tumours of haematopoietic and lymphoid tissues*. Lyon, France: IARC Press, 2001.
2 **Ries LAG**, Eisner MP, Kosary CL, Hankey BF, Miller BA, Clegg L, Mariotto A, Feuer EJ, Edwards BK, eds. *SEER cancer statistics review, 1975–2001*. Bethesda, MD: National Cancer Institute, 2004. Available at: http://seer.cancer.gov/csr/1975_2001/ Accessed September 21, 2004.
3 **Mueller NE**. Hodgkin's Disease. In: Schottenfield D, Fraumeni JF Jr, eds. *Cancer epidemiology and prevention*, 2nd ed. New York: Oxford University Press, 1996:893–919.
4 **Grufferman S**, Cole P, Smith PG, Lukes RJ. Hodgkin's disease in siblings. *New Engl J Med* 1977;**296**:248–50.
5 **Haim N**, Cohen Y, Robinson E. Malignant lymphoma in first-degree blood relatives. *Cancer* 1982;**49**:2197–2200.
6 **Kerzin-Storrar L**, Faed MJW, MacGillivray JB, Smith PG. Incidence of familial Hodgkin's disease. *Br J Cancer* 1983;**47**:707–12.
7 **Mack TM**, Cozen W, Shigbata DK, Weiss LM, Nathwani BN, Hernandez AM, Taylor CR, Hamilton AS, Deapen DM, Rappaport EB. Concordance for Hodgkin's Disease in identical twins suggesting genetic susceptibility to the young-adult form of the disease. *N Eng J Med* 1995;**332**:413–18.
8 **Bernard SM**, Cartwright RA, Darwin CM, Richards IDG, Roberts B, O'Brien C, Bird CC. Hodgkin's disease: Case control epidemiological study in Yorkshire. *Br J Cancer* 1987;**55**:85–90.
9 **Goldgar DE**, Easton DF, Cannon-Albright LA, Skolnick MH. Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. *J Natl Cancer Inst* 1994;**86**:1600–8.
10 **Westergaard T**, Melbye M, Pedersen JB, Frisch M, Olsen JH, Andersen PK. Birth order, sibship size and risk of Hodgkin's disease in children and young adults: A population-based study of 31 million person-years. *Int J Cancer* 1997;**2**:977–81.
11 **Paltiel O**, Schmit T, Adler B, Rachmilevitz EA, Polliak A, Cohen A, Haim N, Ben Shachar M, Epelbaum R, Barchana M, Cohen R, Ben Yehuda D. The incidence of lymphoma in first-degree relatives of patients with Hodgkin disease and non-hodgkin lymphoma. *Cancer* 2000;**88**:2357–66.
12 **Shugart YY**, Hemminki K, Vaittinen P, Kingman A, Dong C. A genetic study of Hodgkin's lymphoma: an estimate of heritability and anticipation based on the familial cancer database in Sweden. *Hum Genet* 2000;**106**:553–6.
13 **Lindelof B**, Eklund G. Analysis of hereditary component of cancer by use of a familial index by site. *Lancet* 2001;**358**:1695–8.
14 **Hemminki K**, Li X, Czene. Familial risk of cancer: data for clinical counseling and cancer genetics. *Int J Cancer* 2004;**108**:109–14.
15 **Goldin LR**, Pfeiffer RM, Gridley G, Gail MH, Li X, Mellemkjaer L, Olsen JH, Hemminki K, Linet MS. Familial aggregation of Hodgkin lymphoma and related tumors. *Cancer* 2004;**100**:1902–8.
16 **Greene MH**, McKeen EA, Li FP, Blattner WA, Fraumeni JF Jr. HLA antigens in familial Hodgkin's disease. *Int J Cancer* 1979;**23**:777–80.
17 **Hors J**, Dausset J. HLA and susceptibility to Hodgkin's disease. *Immunol Rev* 1983;**70**:167–92.
18 **Harty LC**, Lin AY, Goldstein AM, Jaffe ES, Carringon M, Tucker MA, Modi WS. HLA-DR, HLA-DQ, and TAP genes in familial Hodgkin disease. *Blood* 2002;**99**:690–3.
19 **Kong A**, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K. A high-resolution recombination map of the human genome. *Nat Genet* 2002;**31**:241–7.
20 **Fjalldal JG**, Benedikstsson K, Sigurdsson J, Ellingssen LM. Automated genotyping: combining neural networks and decision trees to perform robust allele calling. *Proc Int Joint Conf Neural Networks* 2001:A1–6.
21 **Weeks DE**, Ott J, Lathrop GM. SLINK: a general simulation program for linkage analysis. *Am J Hum Genet* 1990;**47**:A204.
22 **Ott J**. Computer-simulation methods in human linkage analysis. *Proc Natl Acad Sci USA* 1989;**86**:4175–8.

23 **Mukhopadhyay N**, Almasy L, Schroeder M, Mulvihill WP, Weeks DE. Mega2, a data-handling program for facilitating genetic linkage and association analyses. *Am J Hum Genet* 1999;**65**:A436.
24 **O'Connell JR**, Weeks DE. PedCheck: A program for identifying genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998;**63**:259–66.
25 **Sobel E**, Lange K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker sharing statistics. *Am J Hum Genet* 1996;**58**:1323–37.
26 **Sobel E**, Sengul H, Weeks DE. Multipoint estimation of identity-by-descent probabilities at arbitrary positions among marker loci on general pedigrees. *Hum Hered* 2001;**52**:121–31.
27 **Sobel E**, Papp JC, Lange K. Detection and integration of genotyping errors in statistical genetics. *Am J Hum Genet* 2002;**70**:496–508.
28 **Lathrop GM**, Lalouel JM, Julier C, Ott J. Strategies for multilocus analysis in humans. *PNAS* 1984;**81**:3443–6.
29 **Lathrop GM**, Lalouel JM. Easy calculations of lod scores and genetic risks on small computers. *Am J Hum Genet* 1984;**36**:460–5.
30 **Cottingham Jr RW**, Idury RM, Schäaffer AA. Faster sequential genetic linkage computations. *Am J Hum Genet* 1993:252–63.
31 **Schäffer AA**, Gupta SK, Shriram K, Cottingham RW Jr. Avoiding recomputation in linkage analysis. *Hum Hered* 1994;**44**:225–37.
32 **Kruglyak L**, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996;**58**:1347–63.
33 **Kong A**, Cox NJ. Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet* 1997;**61**:1179–88.
34 **Statistical Solutions Ltd**. *SAGE: statistical analysis for genetic epidemiology*. Cork: Statistical Solutions Ltd, 2003.
35 **Witte J**, Elston RC, Schork NJ. Genetic dissection of complex traits. *Nature Genet* 1996;**12**:355–6.
36 **Wiltshire S**, Cardon LR, McCarthy MI. Evaluating the results of genomewide linkage scans of complex traits by locus counting. *Am J Hum Genet* 2002;**71**:1175–82.
37 **Lander E**, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995;**11**:241–7.
38 **Joos S**, Kupper M, Ohl S, von Bonin F, Mechtersheimer G, Bentz M, Marynen P, Moller P, Pfreundschuh M, Trumper L, Lichter P. Genomic imbalances including amplification of the tyrosine kinase gene *JAK2* in CD30+ Hodgkin cells. *Cancer Res* 2002;**60**:549–52.
39 **Joos S**, Granzow M, Holtgreve-Grez H, Siebert R, Harder L, Martin-Subero JI, Wolf J, Adamowicz M, Barth TF, Lichter P, Jauch A. Hodgkin's lymphoma cell lines are characterized by frequent aberrations on chromosomes 2p and 9p including *REL* and *JAK2*. *Int J Cancer* 2003;**103**:489–95.
40 **Re D**, Starostik P, Massoudi N, Staratschek-Jox A, Dries V, Thomas RK, Diehl V, Wolf J. Allelic losses on chromosome 6q25 in Hodgkin and Reed Sternberg cells. *Cancer Res* 2003;**63**:2606–9.
41 **Maggio EM**, Stekelenburg E, Van den Berg A, Poppema S. TP53 gene mutations in Hodgkin lymphoma are infrequent and not associated with absence of Epstein-Barr virus. *Int J Cancer* 2001;**94**:60–6.
42 **Weintraub M**, Lin AY, Franklin J, Tucker MA, Magrath IT, Bhatia KG. Absence of germline p53 mutations in familial lymphoma. *Oncogene* 1996;**12**:687–91.
43 **Ferraris AM**, Racchi O, Repezzi D, Gaetani GF, Boffetta P. Familial Hodgkin's disease: a disease of young adulthood? *Ann Hematol* 1997;**74**:131–4.

SHORT REPORT

A novel susceptibility locus at 2p24 for generalised epilepsy with febrile seizures plus

D Audenaert, L Claes, K G Claeys, L Deprez, T Van Dyck, D Goossens, J Del-Favero, Paesschen, C Van Broeckhoven, P De Jonghe

J Med Genet 2005;**42**:947–952. doi: 10.1136/jmg.2005.031393

Generalised epilepsy with febrile seizures plus (GEFS+) is a clinically and genetically heterogeneous epilepsy syndrome. Using positional cloning strategies, mutations in *SCN1B*, *SCN1A*, and *GABRG2* have been identified as genetic causes of GEFS+. In the present study, we describe a large four generation family with GEFS+ in which we performed a 10 cM density genome-wide scan. We obtained conclusive evidence for a novel GEFS+ locus on chromosome 2p24 with a maximum two point logarithm of the odds (LOD) score of 4.22 for marker D2S305 at zero recombination. Fine mapping and haplotype segregation analysis in this family delineated a candidate region of 3.24 cM, corresponding to a physical distance of 4.2 Mb. Linkage to 2p24 was confirmed ($p=0.007$) in a collection of 50 nuclear and multiplex families with febrile seizures and epilepsy. Transmission disequilibrium testing and association studies provided further evidence ($p<0.05$) that 2p24 is a susceptibility locus for febrile seizures and epilepsy. Furthermore, we could reduce the candidate region to a 2.14 cM interval, localised between D2S1360 and D2S2342, based upon an ancestral haplotype. Identification of the disease gene at this locus will contribute to a better understanding of the complex genetic aetiology of febrile seizures and epilepsy.

Febrile seizures (MIM 121210) represent the most common childhood convulsion disorder affecting 2–5% of children between the ages of 6 months and 5 years. Seizures are usually generalised tonic-clonic and of brief duration. They are by definition associated with fever and occur without evidence of intracranial infections or other definable causes. Typical febrile seizures remit spontaneously before the age of 5. Twin and family studies have shown that febrile seizures have a heritable component of about 70%.[1 2] Most studies support a polygenic or multifactorial mode of inheritance.[1–3] However, in some rare families, segregation of the febrile seizure phenotype is compatible with a monogenic inheritance model.[2 3]

Although infantile febrile seizures are mostly benign, patients have an increased risk for epilepsy later in life. Generalised epilepsy with febrile seizures plus (GEFS+, MIM 604233) is a familial epilepsy syndrome that links febrile seizures with epilepsy.[4] It encompasses a continuum of phenotypes with mild and severe forms of epilepsy. Most affected individuals present with typical febrile seizures or atypical febrile seizures, that is, febrile seizures that persist beyond 5 years of age. In addition, patients can have afebrile seizures including generalised tonic-clonic, absence, myoclonic, or atonic seizures. The most severe forms of epilepsy in the GEFS+ spectrum are myoclonic-astatic epilepsy and severe myoclonic epilepsy of infancy. GEFS+ has an autosomal dominant inheritance pattern with reduced penetrance.

To date, eight chromosomal loci have been reported for febrile seizures and GEFS+: FEB1 at 8q13–q21, FEB2 at 19p13.3, FEB3 at 2q23–q24, FEB4 at 5q14–q15, FEB5 at 6q22–q24, GEFS+1 at 19q13.1, GEFS+2 at 2q21–q33, and GEFS+3 at 5q31.1–q33.1.[5–13] Mutations in the β1 subunit (*SCN1B*) and the α1 subunit (*SCN1A*) of the voltage gated sodium channel, and mutations in the γ2 subunit of the $GABA_A$ receptor (*GABRG2*), have been identified in GEFS+ families that are, respectively, linked to the GEFS+1, GEFS+2, and GEFS+3 loci.[10 12–14]

In this report, we describe a large non-consanguineous family with febrile seizures and epilepsy (fig 1) that was compatible with GEFS+. The family comprised 11 patients in four successive generations and one affected spouse and was from Flanders, the Dutch speaking region of Belgium. All individuals were clinically assessed by means of a structured interview. For all patients, information was obtained about seizure type, duration, age at onset, frequency, and association with fever. The parents or the older siblings of the patients primarily provided clinical information, and all spouses were interviewed for familial occurrence of seizures to exclude possible bilineal inheritance. The detailed clinical characteristics of all patients are shown in table 1.

In summary, of the 11 patients, eight had seizures associated with fever, but none developed epilepsy later in life. The febrile seizures were generalised tonic-clonic. Most seizures were brief, but in two individuals (III.4 and III.10) they lasted for more than 30 min. The age at onset was very similar, ranging between 6 months and 2.5 years. The number of seizures varied between one and three except for individual II.5 who experienced 23 seizures associated with fever. Three patients had epileptic seizures without a history of febrile seizures. The epilepsy phenotype in those individuals was variable. Patient III.5 presented with generalised tonic seizures, while individual III.6 had only absences. For individual I.1, no detailed clinical data could be obtained.

Before initiating a genome-wide scan, we performed a simulation study using SLINK to estimate the power of this family.[15 16] Two point logarithm of the odds (LOD) scores were calculated with MLINK software (version 5.1) from the LINKAGE package.[17] All individuals with one or more febrile or epileptic seizures were considered affected. We assumed an autosomal dominant inheritance pattern with reduced disease penetrance and a disease frequency of 0.0001. We set the disease penetrance at 90% based upon calculations according to the method of Johnson *et al*.[3] Since the frequency of febrile seizures in the population is 2–5%, we incorporated a 3% phenocopy rate. The data were analysed

Abbreviations: GEFS+, generalised epilepsy with febrile seizures plus; LOD, logarithm of the odds; NPL, non-parametric linkage; TDT, transmission disequilibrium test

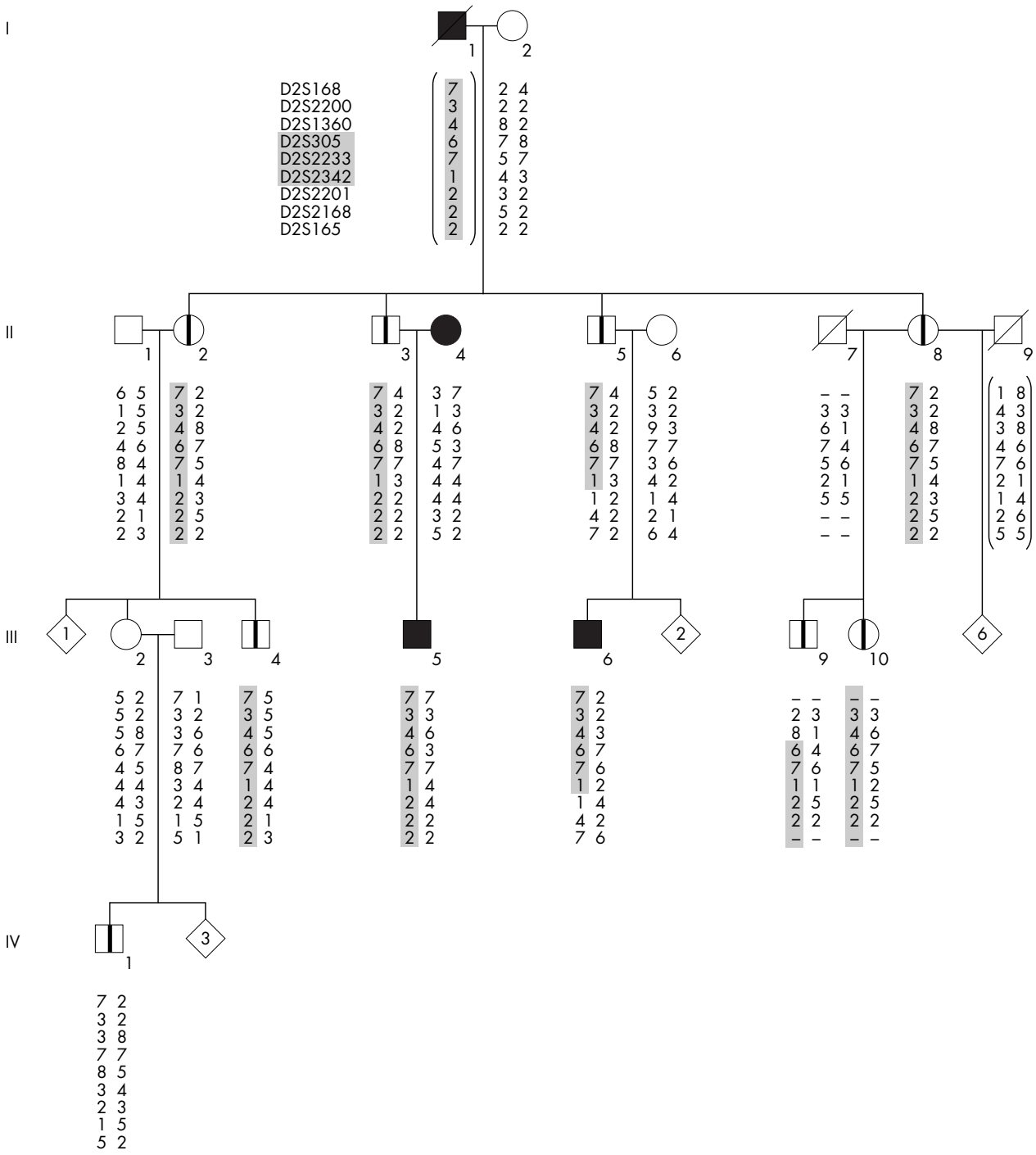

Figure 1 Pedigree of the four generation Belgian-Dutch family with GEFS+. Closed box: man with epileptic seizures; closed circle: woman with epileptic seizures; vertical bar in empty box: man with febrile seizures; vertical bar in empty circle: woman with febrile seizures; forward slash: deceased. Haplotypes of chromosome 2p markers are shown under each symbol, a dash (–) indicates that genotypes were not obtained for the respective marker. To maintain confidentiality, the number of at risk individuals that were genotyped in the linkage analysis is indicated in a diamond. The disease haplotype is shaded. Haplotypes between brackets were inferred.

assuming equifrequent marker alleles and gender average recombination rates. A maximum two point LOD score of 5.07 was generated, using an eight allele marker with equal allele frequencies.

We obtained peripheral blood from all participating individuals and extracted genomic DNA using standard methods. All participants or their legal representative signed a written informed consent form, and the Medical Ethical Committee of the University of Antwerp approved this study.

We performed a 10 cM genome-wide scan with the ABI Prism Linkage Mapping Set MD10 (Applied Biosystems, Foster City, CA) and genotyped 29 family members with 382 microsatellite markers distributed over the human genome. After amplification, PCR products were pooled and size fractionated on an ABI 3700 automated sequencer (Applied Biosystems). We used ABI Prism GeneScan and Genotyper version 3.7 software (Applied Biosystems) to determine allele sizes. For fine mapping, we analysed 16 additional

Table 1 Clinical data of patients in the four generation GEFS+ family

Patient	Gender	Age (years)	FS	ES	Onset age (months)	Seizure type	Duration (min)	No. of seizures
I.1	M	–	No	Yes	–	–	–	–
II.2	F	63	Yes	No	17	GTCS	2	3
II.3	M	62	Yes	No	18	GTCS	3–5	2
II.5	M	55	Yes	No	18	GTCS	5	23
II.8	F	57	Yes	No	17–24	GTCS	–	1
III.4	M	35	Yes	No	6	GTCS	2–30	3
III.5	M	37	No	Yes	9 years	GTS	1–2	3
III.6	M	24	No	Yes	9	Absences	5	1
III.9	M	16	Yes	No	30	GTCS	5	1
III.10	F	15	Yes	No	18	GTCS	>45	1
IV.1	M	14	Yes	No	15	GTCS	<1	2

ES, epileptic seizures; FS, febrile seizures; GTCS, generalised tonic-clonic seizures; GTS, generalised tonic seizures.

microsatellite markers selected from the Marshfield genetic map. Primer pairs for each marker were chosen with a proprietary algorithm implemented in the Multiplexer program (Goossens *et al*, unpublished data). For each marker, we calculated two point LOD scores and obtained a maximum LOD score of 4.22 for marker D2S305 at zero recombination (table 2), assuming the same genetic model as in the simulation study.

None of the previously reported loci showed evidence of linkage, and no other markers from the genome-wide panel gave a LOD score above 3 (data not shown). After analysis of recombination events in patients, we delineated a candidate region of 3.24 cM, localised between markers D2S1360 and D2S2201 (fig 1), corresponding to an estimated physical distance of 4.2 Mb. The telomeric boundary was defined by the recombinant haplotype in II.5 and his son III.6. A recombination in III.9 delimited the centromeric boundary. All patients except IV.1 shared the haplotype segregating with the disease. His unaffected mother also did not inherit the disease haplotype, indicating that individual IV.1 was likely a phenocopy and that his febrile seizures were unrelated to the genetic aetiology in the family. This is not unusual in GEFS+ families, and in part reflects the high frequency of febrile seizures (2–5%) in the general population. It is also of interest that none of the unaffected at risk family members carried the disease haplotype, indicating that the GEFS+ phenotype is fully penetrant in this pedigree. This is rather exceptional as GEFS+ or febrile seizures usually show a disease penetrance varying between 60% and 90%.

After establishing linkage at chromosome 2p24, an additional group of 50 families with febrile seizures and/or epilepsy (table 3) was analysed. All families were of Belgian-Dutch origin. The group comprised a total of 291 individuals of whom 142 were patients. The proband of each family presented with seizures that were associated with fever. For each proband, at least one of the parents was available and provided clinical information about seizure frequency and association with fever. Parents were interviewed to identify a possible positive history for febrile seizures in other family members. If other patients in the family were available, the patients and/or their parents were contacted and interviewed. Additional patients were considered affected if they had at least one unprovoked febrile and/or

Table 3 Characteristics of the febrile seizures and/or the epilepsy families

	2 generation (1 affected offspring)	2 generation (>1 affected offspring)	3 or more generation (≥2 patients)
No. of families	11	8	31
No. of individuals analysed	31	36	224
No. of patients	21	20	101
Phenotype of patients			
FS	17	15	87
FS+ES	4	1	6
ES	0	5	8

ES, epileptic seizures; FS, febrile seizures.

Table 4 Non-parametric linkage analysis

Marker	Distance (cM)	NPL score	p value
D2S1360	38.33	2.25	0.011
D2S305	38.87	2.42	0.007
D2S2233	39.93	2.43	0.007
D2S2342	40.47	2.16	0.013
D2S2201	41.57	1.79	0.032

NPL, non-parametric linkage. Positions of the markers were obtained from the Marshfield genetic map.

Table 2 Two point LOD scores for markers at 2p24

		LOD scores at recombination fraction θ of						
Marker	Distance (cM)	0.00	0.01	0.05	0.10	0.20	0.30	0.40
D2S168*	27.06	1.17	1.29	1.56	1.66	1.53	1.16	0.63
D2S2200	27.60	1.17	1.26	1.43	1.46	1.28	0.93	0.49
D2S1360	38.33	3.10	3.15	3.18	3.04	2.51	1.78	0.94
D2S305*	38.87	4.22	4.15	3.87	3.51	2.73	1.87	0.96
D2S2233	39.93	3.03	2.99	2.79	2.52	1.90	1.23	0.55
D2S2342	40.47	3.57	3.51	3.25	2.92	2.23	1.52	0.77
D2S2201	41.57	1.65	2.56	2.98	2.94	2.49	1.81	0.98
D2S2168	45.30	1.04	1.95	2.36	2.32	1.92	1.33	0.67
D2S165*	47.43	−1.04	−0.10	0.45	0.59	0.58	0.44	0.24

Poitions of the markers were obtained from the Marshfield genetic map.
*Microsatellite marker from the 10 cM genome-wide scan panel.

Table 5 Allelic association analysis for five markers at 2p24

Marker	$p_{overall}$	Alleles	Proband		Control		
			n	%	n	%	p
D2S1360	0.656	2	10	10	22	11.2	0.748
		3	6	6	16	8.2	0.502
		4	26	26	61	31.1	0.360
		5	10	10	19	9.7	0.933
		6	9	9	15	7.7	0.688
		7	10	10	19	9.7	0.933
		8	7	7	12	6.1	0.771
		9	6	6	13	6.6	0.834
		Other	16		19		
D2S305	0.262	2	10	10	23	11.7	0.654
		3	9	9	18	9.2	0.959
		4	12	12	35	17.9	0.192
		6	46	46	67	34.2	0.048*
		7	12	12	22	11.2	0.843
		Other	11		31		
D2S2233	0.048*	2	10	10	36	18.4	0.060
		4	13	13	43	21.9	0.063
		5	9	9	18	9.2	0.959
		6	13	13	26	13.3	0.949
		7	32	32	42	21.4	0.047*
		8	12	12	11	5.6	0.052
		9	7	7	14	7.1	0.964
		Other	4		6		
D2S2342	0.897	2	46	46	87	44.4	0.792
		5	15	15	33	16.8	0.685
		6	27	27	48	24.5	0.639
		Other	12		28		
D2S2201	0.476	1	6	6	9	4.6	0.601
		5	8	8	17	8.7	0.844
		7	6	6	15	7.7	0.600
		8	64	64	112	57.1	0.256
		9	7	7	12	6.1	0.771
		Other	9		31		

Only alleles with an observed frequency >5% in the patients were included in the analysis.
*p<0.05, statistically significant.

afebrile seizure. Mutations in *SCN1A*, *SCN1B*, and *GABRG2* were excluded by mutation analysis of genomic DNA of each proband[18] (Audenaert *et al*, unpublished data).

To examine linkage to 2p24, we genotyped these families for three markers localised within the candidate region and two flanking markers, spanning a region of 3.24 cM. We calculated multipoint non-parametric linkage (NPL Z_{all}) scores and associated p values, using the total stat command in the Genehunter program (http://linkage.rockefeller.edu/soft/gh). Eleven families consisted of only one affected

Table 6 Two marker sliding window analysis

Haplotype markers	$p_{overall}$	Alleles	Proband		Control		
			n	%	n	%	p
D2S1360–D2S305	0.681	H-2-6	5	5	7	3.8	0.556
		H-4-6	11	11	17	9.2	0.517
		H-4-7	5	5	7	3.8	0.556
		Other	79		165		
D2S305–D2S2233	0.008*	H-4-7	4	4	7	3.8	0.854
		H-6-6	5	5	10	5.4	0.970
		H-6-7	17	17	12	6.5	0.003*
		H-6-8	4	4	2	1.1	0.085
		Other	70		165		
D2S2233–D2S2342	0.114	H-4-2	5	5	12	6.5	0.695
		H-6-2	9	9	13	7.1	0.463
		H-7-2	15	15	18	9.8	0.133
		H-7-5	4	4	3	1.6	0.186
		H-7-6	8	8	11	6.0	0.428
		H-8-6	4	4	3	1.6	0.186
		H-9-2	4	4	6	3.3	0.672
		Other	51		130		
D2S2342–D2S2201	0.775	H-2-8	25	25	42	22.8	0.487
		H-2-9	5	5	2	1.1	0.033*
		H-5-8	7	7	18	9.8	0.523
		H-6-8	16	16	27	14.7	0.608
		Other	47		107		

Only haplotype alleles with an observed frequency >4% in the patients were included in the analysis.
*p<0.05, statistically significant.

offspring and were omitted from the NPL analysis. A maximum multipoint NPL Z_{all} score of 2.43 ($p = 0.007$) was obtained at D2S2233 (table 4). The information content of the markers was higher than 80% throughout the 3.24 cM region. These results indicated that the locus at 2p24 is involved in the genetic aetiology underlying the susceptibility for febrile seizures and epilepsy in these families.

To confirm this finding, we performed for each of the five markers at 2p24 a transmission disequilibrium test (TDT) with the probands of the 50 families, using the Genehunter tdt command. This showed that allele 6 at D2S305 was transmitted significantly more to the patients compared to other alleles at D25305 ($p = 0.022$; allele 6 was transmitted 29 times and not transmitted 14 times).

As a next step, we performed an association study and compared the allelic distribution between probands of the 50 families and 98 Belgian-Dutch population control individuals. Overall p values were calculated using CLUMP (http://linkage.rockefeller.edu/soft/clump.html). The overall allelic distribution of the markers was not significantly different between patients and controls, except for D2S2233 (table 5; $p_{overall} = 0.048$). Using the χ^2 statistic, we compared allele frequencies between patients and controls. We found significant allelic association with febrile seizures at D2S305 and D2S2233 (table 5): allele 6 at D2S305 ($p_{allele6} = 0.048$; OR 1.64; 95% CI 1.00–2.68) and allele 7 at D2S2233 ($p_{allele7} = 0.047$; OR 1.73; 95% CI 1.00–2.96).

These findings were further supported using sliding window analysis of two marker haplotypes. To obtain observed haplotypes, we genotyped the 50 probands with their parents and 49 triads of Belgian-Dutch population control individuals. Genehunter generated the most likely haplotypes of patients and control individuals. Using CLUMP, we compared the overall allelic distribution of the two marker haplotypes and found a highly significant overall p value at D2S305–D2S2233 (table 6; $p_{overall} = 0.008$). We calculated the contribution of each haplotype to the overall p value using the χ^2 statistic, and found that H-6-7 at D2S305–D2S2233 was significantly overrepresented in patients (table 6; $p_{H\text{-}6\text{-}7} = 0.003$; OR 2.94; 95% CI 1.34–6.43). These results demonstrated that a common haplotype at chromosome 2p24, H-6-7, contributed to developing febrile seizures in the Belgian-Dutch population. Furthermore, the disease haplotype in the large linkage family also contained the H-6-7 allele at D2S305–D2S2233 (fig 1). This allowed us to further refine the candidate region at 2p24 to a 2.14 cM interval localised between D2S1360 and D2S2342, based upon ancestral recombination events. A p value less than 0.05 was also obtained for H-2-9 at D2S2342–D2S2201. However, the frequency of this haplotype allele in the control group (table 6; 1.1%) was too low to provide conclusive evidence.

The chromosome 2p24 locus for febrile seizures contains 14 known and six putative genes. So far, the three genes involved in GEFS+, that is, *SCN1B*, *SCN1A*, and *GABRG2*, encode subunits of voltage gated or ligand gated ion channels. Therefore, one of the most attractive candidate genes in this locus is *KCNS3* (voltage gated potassium channel, subfamily S, member 3). It encodes a subunit of the voltage gated potassium channel of the delayed rectifier type and is abundantly expressed in the brain.[19] Further characterisation of *KCNS3* and other genes at the chromosome 2p24 locus might lead to identification of the actual susceptibility alleles in the Belgian-Dutch population and the causal variant in our linkage family.

ACKNOWLEDGEMENTS

We are grateful to the patients and their relatives for their cooperation and participation in this study. We thank Iris Smouts for collecting blood samples and the people at the VIB Genetic Service Facility (http://www.vibgeneticservicefacility.be) for the genetic analyses.

ELECTRONIC-DATABASE INFORMATION

The following websites have been mentioned in this report: Center for Medical Genetics, Marshfield Medical Research Foundation, http://research.marshfieldclinic.org/genetics/; Laboratory of Statistical Genetics, Rockefeller University, http://linkage.rockefeller.edu (for Genehunter and CLUMP); Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for febrile seizures, and generalised epilepsy with febrile seizures plus); VIB Genetic Service Facility, http://www.vibgeneticservicefacility.be.

Authors' affiliations
D Audenaert, L Claes, K G Claeys, L Deprez, T Van Dyck, D Goossens, J Del-Favero, C Van Broeckhoven, P De Jonghe, Department of Molecular Genetics, Flanders Interuniversity Institute for Biotechnology, University of Antwerp, Antwerp, Belgium
K G Claeys, P De Jonghe, Division of Neurology, University Hospital of Antwerp, Antwerp, Belgium
W Van Paesschen, Division of Neurology, University Hospital of Leuven, Leuven, Belgium

Financial support was received from the Fund for Scientific Research Flanders (FWO-F), the Queen Elisabeth Medical Foundation, and the Interuniversity Attraction Poles (IUAP) program P5/19 of the Federal Science Policy Office, Belgium. DA is a PhD fellow of the FWO-F and LD of the Institute for Science and Technology (IWT), Belgium

Competing interests: none declared

All participants or their legal representative signed a written informed consent form, and the Medical Ethical Committee of the University of Antwerp approved this study

Correspondence to: Professor Peter De Jonghe, Department of Molecular Genetics (VIB8), Neurogenetics Research Group, University of Antwerp (UA), Universiteitsplein 1, 2610 Antwerp, Belgium; peter.dejonghe@ua.ac.be

Revised version received 22 March 2005
Accepted for publication 28 March 2005
Published Online First 12 April 2005

REFERENCES

1 **Tsuboi T**. Genetic analysis of febrile convulsions: twin and family studies. *Hum Genet* 1987;**75**:7–14.
2 **Rich SS**, Annegers JF, Hauser WA, Anderson VE. Complex segregation analysis of febrile convulsions. *Am J Hum Genet* 1987;**41**:249–57.
3 **Johnson WG**, Kugler SL, Stenroos ES, Meulener MC, Rangwalla I, Johnson TW, Mandelbaum DE. Pedigree analysis in families with febrile seizures. *Am J Med Genet* 1996;**61**:345–52.
4 **Scheffer IE**, Berkovic SF. Generalized epilepsy with febrile seizures plus. A genetic disorder with heterogeneous clinical phenotypes. *Brain* 1997;**120**:479–90.
5 **Wallace RH**, Berkovic SF, Howell RA, Sutherland GR, Mulley JC. Suggestion of a major gene for familial febrile convulsions mapping to 8q13–21. *J Med Genet* 1996;**33**:308–12.
6 **Johnson EW**, Dubovsky J, Rich SS, O'Donovan CA, Orr HT, Anderson VE, Gil-Nagel A, Ahmann P, Dokken CG, Schneider DT, Weber JL. Evidence for a novel gene for familial febrile convulsions, FEB2, linked to chromosome 19p in an extended family from the Midwest. *Hum Mol Genet* 1998;**7**:63–7.
7 **Peiffer A**, Thompson J, Charlier C, Otterud B, Varvil T, Pappas C, Barnitz C, Gruenthal K, Kuhn R, Leppert M. A locus for febrile seizures (FEB3) maps to chromosome 2q23–24. *Ann Neurol* 1999;**46**:671–8.
8 **Nakayama J**, Hamano K, Iwasaki N, Nakahara S, Horigome Y, Saitoh H, Aoki T, Maki T, Kikuchi M, Migita T, Ohto T, Yokouchi Y, Tanaka R, Hasegawa M, Matsui A, Hamaguchi H, Arinami T. Significant evidence for linkage of febrile seizures to chromosome 5q14–q15. *Hum Mol Genet* 2000;**9**:87–91.
9 **Nabbout R**, Prud'homme JF, Herman A, Feingold J, Brice A, Dulac O, LeGuern E. A locus for simple pure febrile seizures maps to chromosome 6q22–q24. *Brain* 2002;**125**:2668–80.
10 **Wallace RH**, Wang DW, Singh R, Scheffer IE, George AL Jr, Phillips HA, Saar K, Reis A, Johnson EW, Sutherland GR, Berkovic SF, Mulley JC. Febrile seizures and generalized epilepsy associated with a mutation in the Na+-channel beta1 subunit gene SCN1B. *Nat Genet* 1998;**19**:366–70.

11 **Baulac S**, Gourfinkel-An I, Picard F, Rosenberg-Bourgin M, Prud'homme JF, Baulac M, Brice A, LeGuern E. A second locus for familial generalized epilepsy with febrile seizures plus maps to chromosome 2q21–q33. *Am J Hum Genet* 1999;**65**:1078–85.
12 **Baulac S**, Huberfeld G, Gourfinkel-An I, Mitropoulou G, Beranger A, Prud'homme JF, Baulac M, Brice A, Bruzzone R, LeGuern E. First genetic evidence of GABA(A) receptor dysfunction in epilepsy: a mutation in the gamma2-subunit gene. *Nat Genet* 2001;**28**:46–8.
13 **Wallace RH**, Marini C, Petrou S, Harkin LA, Bowser DN, Panchal RG, Williams DA, Sutherland GR, Mulley JC, Scheffer IE, Berkovic SF. Mutant GABA(A) receptor gamma2-subunit in childhood absence epilepsy and febrile seizures. *Nat Genet* 2001;**28**:49–52.
14 **Escayg A**, MacDonald BT, Meisler MH, Baulac S, Huberfeld G, An-Gourfinkel I, Brice A, LeGuern E, Moulard B, Chaigne D, Buresi C, Malafosse A. Mutations of SCN1A, encoding a neuronal sodium channel, in two families with GEFS+2. *Nat Genet* 2000;**24**:343–5.
15 **Ott J**. Computer-simulation methods in human linkage analysis. *Proc Natl Acad Sci U S A* 1989;**86**:4175–8.
16 **Weeks DE**, Ott J, Lathrop GM. SLINK: a general simulation program for linkage analysis. *Am J Hum Genet* 1990;**47**:A204.
17 **Cottingham RW Jr**, Idury RM, Schaffer AA. Faster sequential genetic linkage computations. *Am J Hum Genet* 1993;**53**:252–63.
18 **Audenaert D**, Claes L, Ceulemans B, Löfgren A, Van Broeckhoven C, De Jonghe P. A deletion in SCN1B is associated with febrile seizures and early-onset absence epilepsy. *Neurology* 2003;**61**:854–6.
19 **Stocker M**, Kerschensteiner D. Cloning and tissue distribution of two new potassium channel alpha-subunits from rat brain. *Biochem Biophys Res Commun* 1998;**248**:927–34.

CORRECTION

doi: 10.1136/jmg.2005.27433corr1

In the Letter to JMG titled, a genome screen of families at high risk for Hodgkin lymphoma: evidence for a susceptibility gene on chromosome 4 (*J Med Genet* 2005;**42**: 595–601) figure 1 was incorrect. The curves are missing from Chromosomes 16 and 19. A full corrected figure 1 is available on the JMG website at http://www.jmedgenet.com/supplemental.