UNIVERSITY OF LONDON THESIS

Degree _Pho_   Year _2005_   Name of Author _PINEDA -TRUJILLO, N.E_

COPYRIGHT
This is a thesis accepted for a Higher Degree of the University of London. It is an unpublished typescript and the copyright is held by the author. All persons consulting the thesis must read and abide by the Copyright Declaration below.

COPYRIGHT DECLARATION
I recognise that the copyright of the above-described thesis rests with the author and that no quotation from it or information derived from it may be published without the prior written consent of the author.

LOANS
Theses may not be lent to individuals, but the Senate House Library may lend a copy to approved libraries within the United Kingdom, for consultation solely on the premises of those libraries. Application should be made to: Inter-Library Loans, Senate House Library, Senate House, Malet Street, London WC1E 7HU.

REPRODUCTION
University of London theses may not be reproduced without explicit written permission from the Senate House Library. Enquiries should be addressed to the Theses Section of the Library. Regulations concerning reproduction vary according to the date of acceptance of the thesis and are listed below as guidelines.

A.   Before 1962. Permission granted only upon the prior written consent of the author. (The Senate House Library will provide addresses where possible).

B.   1962 - 1974. In many cases the author has agreed to permit copying upon completion of a Copyright Declaration.

C.   1975 - 1988. Most theses may be copied upon completion of a Copyright Declaration.

D.   1989 onwards. Most theses may be copied.

This thesis comes within category D.

☑   This copy has been deposited in the Library of _UCL_

☐   This copy has been deposited in the Senate House Library, Senate House, Malet Street, London WC1E 7HU.

Linkage Analysis of Mendelian Forms of Complex Disorders in a South American Population

By

Nicolas Guillermo Pineda-Trujillo

Submitted for the degree of Doctor of Philosophy

University College London, UCL

September 2005

The Galton Laboratory
Wolfson House
4 Stephenson Way
London NW1 2HE

UMI Number: U593114

UMI

Dissertation Publishing

ProQuest

To my Mother and to the memory of my Father, who passed away during the time I was in London doing my PhD.

ACKNOWLEDGMENTS

Linkage Analysis of Mendelian forms of Complex Disorders in a South American Population

Abstract

Genetic analysis in Mendelian forms of complex disorders may help increase our understanding of the biology of these disorders. Genes identified in Mendelian forms could also be relevant for disease susceptibility of the more common forms of the disease. I have studied a collection of patient samples from Antioquia, Colombia, including Mendelian and sporadic forms of four well characterized diseases: Parkinson's [PD], Generalized epilepsy with Febrile Seizures Plus [GEFS+], Batten's disease and Type 1 Diabetes [T1D]. In addition, I studied two extended families with unusual disorders. Affected individuals in one family presented with episodic crisis of abdominal pain, possibly corresponding to Abdominal Epilepsy [AE]. Patients from the other family presented with a multiform movement disorder [MMD], at times diagnosed as Parkinson's or Huntington's disease. Brain pathology from an affected individual showed the presence of iron deposits as a key finding. I subsequently performed linkage analysis of these families. Initial analysis focused on testing linkage to known candidate genes or regions. If these proved negative or if no candidate gene/region was available, a genome scan was then conducted.

I have excluded all previously known genes/loci in two out of six families with GEFS+. The remaining four families were found linked to either *SCN1A* or *GABRG2* genes. In a family with Parkinson disease, a novel mutation was identified in exon 3 of the *PARK2* gene. It was not possible to establish whether a common chromosome was involved in the same mutation reported separately by both a Spanish group and we since those mutant chromosomes are presenting a very different haplotype in both populations. In Batten disease, the first mutation in *CLN5* gene outside Europe was found, which leads to a variant of the juvenile form rather than to a late infantile presentation of Neuronal Ceroid Lipofuscinosis. For T1D, a haplotype on chromosome 2 was found to be segregating with the disease in a large family with autoimmune T1D.

Regarding the two genome wide scans, in MMD, suggestion of linkage to two functionally candidate genes was found. These are the Neurogolobin (NGB) and the Ferritin heavy polypeptide (FTH1) genes. In Abdominal epilepsy, a locus was identified on chromosome 8 and fine mapping led to the identification of a critical region extending 1.3 Mb at 8q13.

LINKAGE ANALYSIS OF MENDELIAN FORMS OF COMPLEX

DISORDERS IN A SOUTH AMERICAN POPULATION

TABLE OF CONTENTS

Dedication

Acknowledgements

ABSTRACT

CHAPTER TWO

CHAPTER THREE

CHAPTER FOUR

CHARACTERIZATION OF A FAMILY WITH A MULTIFORM MOVEMENT

CHAPTER EIGHT

LINKAGE ANALYSIS OF TYPE 1 DIABETES MELLITUS IN AN EXTENDED

List of Figures

CHAPTER ONE

INTRODUCTION

This thesis describes studies aimed at identification of the genes involved in a variety of Mendelian forms of disorders which are more commonly of complex aetiology.

Series of large families from Antioquia-Colombia with Parkinson's disease, Generalized Epilepsy with Febrile Seizures Plus [GEFS+], Batten's disease, Type 1 Diabetes [T1D], Multiform Movement Disorder [MMD] and Abdominal epilepsy have been studied.

In the first five disorders, both linkage and association analysis, as a direct analysis of candidate genes/loci was carried out. After having excluded several candidate loci in MMD and having no clear candidate genes for Abdominal epilepsy, a genome wide scan was conducted in both disorders which involved typing, on average, one STR every six cM.

Genes identified in Mendelian forms of Complex disorders can help increase our understanding of the biology of these disorders and could also be relevant in common forms of the disease. For several inherited diseases, virtually everything that is known about the biological processes underlying the disease process derives from knowledge of the protein affected by the mutations that lead to the disease.

In this first section the background for the success to candidate gene and positional cloning approaches in Mendelian traits (and Mendelian-like) is described. Also, description of which could be the most appropriate approaches for the case of complex traits is provided.

1. Gene mapping of complex traits

Complex traits involve several loci contributing towards a phenotype, in addition to interaction between them. Also, environmental factors are expected to play a role. These factors in addition to irregular expressivity, incomplete penetrance and the existence of phenocopies complicate the situation [1].

As different loci participate in the phenotype, each probably with very low effect, linkage analysis studies have not been very successful in identifying these susceptibility loci. Therefore, association or Linkage disequilibrium (LD) studies have received more attention in the expectation that these may be more effective in identifying these small effects [2,3].

Despite the fact that no consensus has been adopted about what strategy is the best to identify complex disease genes, many searches have been launched to identify new markers to use. Thus, projects are underway looking for SNPs (Single Nucleotide Polymorphisms), in order to provide with a high density map of markers that then can be used in association or LD studies [4].

Two main methodologies are used for looking for genes involved in Complex traits: Map-based and gene (and sequence) based approaches. The map-based approach would require an estimate of 500,000 to 1,000,000 SNPs to be typed in individuals from European descent, and perhaps two-fold more for individuals from African populations (due to shorter haplotype-block and higher diversity) [5,6].

The alternative to the map-based approach is based on typing SNPs in coding or promoter regions. SNPs in coding sequence are easier to identify as disease-predisposing variants since it is possible to assess the predicted severity of the amino acid change. Some amino acid substitutions might have weak physiological effects that

3

added to other weak effects in other genes may result in the phenotype. The gene-based approach if applied to the whole genome will require 50,000 to 100,000 SPNs (Ten-fold less than in map-based) to be typed, in order to evaluate the estimated total number of genes.

Variants in regulatory regions have been less frequently identified (see Table 1.1) but it is expected that a more frequent involvement in disease aetiology to be identified, at least in complex traits.

Most SNPs show low allele frequencies. Thus, there are more SNPs with minor frequencies of 1% than SNPs with minor frequencies of 20% [7, 8]. This difference in frequency might be reflected in the relative risk associated to the SNP in relation to a disease phenotype. For example, for SNPs with a strong effect on function, early onset, severe diseases may be involved and be at a lower frequency range than a SNP involved in later onset and/or milder diseases [9]. This observation is in agreement with the degree of selection that may have acted on the disease-associated SNPs. A comprehensive analysis of genetic susceptibility needs to focus on the lower frequency range of potential disease-causing alleles (e.g. less than 1%) [9]. Importantly, these studies might be more efficient in certain populations that because of their particular history may have more homogeneous genetic components.

2. Linkage analysis in Mendelian traits

2.1 The beginning of a positional cloning study

The process of finding mutations that cause diseases where a biochemical defect is unknown is called positional cloning. The main steps for positional cloning involve

identifying families, collecting biological samples, genotyping markers, initial location by linkage, fine and physical mapping, and mutation identification [1].

Positionally cloning begins with linkage analysis. This method is applied in families presenting with a particular disease (or trait). It consists of testing for linkage between polymorphic markers and the disease. Clear definition of the phenotype is critical since it may happen that the signals of lod scores be obscured by misclassification of phenotypes, or even, in some cases, could result in failure of detection of linkage. When multiple families are analyzed, genetic heterogeneity, complex modes of inheritance and misdiagnosis regularly lead to poor lod scores. An appropriate approach to linkage when the initial results are not straightforward is to use a first data set to try to make an appropriate hypothesis and then to replicate the study on a new cohort using the predetermined additional phenotypic ideas derived from the first analysis [9].

2.2 The first successes of positional cloning

The early successes of positional cloning by linkage analysis included identification of the genes underlying chronic granulomatous disease, the X-linked muscular dystrophies, cystic fibrosis, Fanconi anemia, Ataxia talangietasia and Neurofibromatosis I (Reviewed in [9]). Of these, Cystic fibrosis was the one where linkage mapping alone led to the identification of the CFTR gene [10, 11]. These studies were conducted analysing RFLPs (Restriction Fragment Length Polymorphisms), which detected SNPs.

2.3 Simple inheritance

Most of the genes that have been identified by positional cloning were originally mapped in families presenting with inheritance patterns that clearly were consistent with an autosomal dominant, co-dominant, autosomal recessive, or X-linked mode. Since the disorders presenting with such a clear inheritance pattern are rare, the number of meioses that can be evaluated in such studies is limited, therefore, the resolution level is 1-10 cM. The standard procedure, after detecting initial linkage using a set of markers is to re-examine the same samples with markers spaced more closely in the region of interest. However, the limitation in the resolution remains in the number of meioses. And then, when the actual mutant gene is found, there is often evidence of more than one mutation leading to the same phenotype, incomplete penetrance, and different degrees of severity which complicates the situation.

2.4 Homozygosity mapping

Homozygosity mapping is used to map recessive traits and consists of evaluating affected relatives of known degree for chromosomal regions inherited in common from a common ancestor [12]. This method has the advantage that only a few individuals are required and that genetic heterogeneity is less likely to affect the analyses because each set of closely related individuals is very likely to carry the same mutation (autozygosity) [9]. This method is particularly suitable in populations with high rates of inbreeding. For example, from the five genes identified as causing Fanconi anemia, at least three were mapped by homozygosity mapping, taking advantage of the low probability of heterogeneity [13-15].

3. Genetic Markers.

Eukaryotic genomes show considerable DNA sequence variations (polymorphisms) within individuals and between populations. Two types of DNA polymorphisms are particularly well characterised: STRs and SNPs.

Short Tandem Repeats (STRs) or microsatellite repeats consist of 1-6 unit repeats [16]. One of their advantages is that they can be typed using the polymerase chain reaction (PCR). They have high mutation rate and many alleles, which can be easily detected in the population by means of denaturing sequencing gels. Polymorphism in STRs arise from mutational events such as unequal crossover, replication slippage or double-strand break repair [17-19]. Their heterozygosity reaches values greater than 0.7 and they are also very abundant in the genome; they are often found in non-coding regions of genes and often they are non-functional. However, there are associations of STRs in UTRs (UnTranslated Regions) and introns which suggest they play a role in gene expression and disease [20]. In addition, polymorphisms of tandem repeats within protein coding sequences are known to modulate disease risk and can result in coding changes leading to diseases such as myotonic dystrophy [21], Huntington's disease [22]. The most common repeat is the dinucleotide CA [23]. STRs are commonly found through both prokaryotic and eukaryotic genomes [24].

Single or Simple Nucleotide Polymorphisms (SNPs) consist of the substitution of one nucleotide by another or the addition or deletion of one or a few nucleotides. SNPs are more abundant than STRs showing a SNP every 1000 bases [25]. As usually only two

alleles are found in populations, the information content per SNP marker is lower than multiallelic STR markers and heterozygosities only up to 0.5 are possible.

For two SNP markers (composite marker) to provide similar information to one STR, several assumptions must be met. Amongst these, no association between the two linked SNPs, phase unequivocally determined, the number of haplotypes at the composite marker should correspond to the number of alleles at the STR and the haplotype frequencies should correspond to the allele frequencies of the STR [26].

4. Databases for genes identified

4.1 Disease associated genes

Databases such as Online Mendelian Inheritance in Man (OMIM, http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=omim/), Human Gene Mutation Database (HGMD, http://www.hgmd.cr.ac.uk/) and Locus Link (http://.www.ncbi.nlm.nih.gov/locuslink/) provide up to date information of the number of genes as wells as the phenotypes associated to them. Thus, HGMD in August 08th, 2005 reported 1863 disease associated genes. From these, the great majority correspond to the so called Mendelian traits, suggesting that only the "easy" ones (simple inheritance) have been found so far.

4.2 Type and frequency of mutations

HGMD has classified the mutations reported by type and frequency. By August 2005 over 47339 mutations in 1863 genes associated with human diseases had been

reported. Table 1.1 summarizes frequencies of mutations in human disease phenotypes. Mutations including in-frame amino acid substitutions or stop codons are the most frequent ones (~57%), followed by deletions (~17%) and splice site mutations (~9.5%).

CHANGE	NUMBER	PERCENTAGE (%)
Micro-lesions		
Missense/Nonsense	27099	57.24
Splicing	4480	9.46
Regulatory	560	1.18
Small deletions	7888	16.66
Small insertions	3145	6.64
Small indels	447	0.94
Gross lesions		
Repeat variations	117	0.25
Gross insertions & duplications	422	0.89
Complex rearrangements (including inversions)	557	1.17
Gross deletions	2624	5.57
Total	47339	100

Table 1.1. Relative frequency of types of mutations underlying disease phenotypes reported in HGMD website updated on 08/08, 2005.

5. Isolate populations

Isolate populations have been documented as very powerful for gene mapping of complex traits. Most cases of allelic association will result from linkage disequilibrium (LD). An important source of LD is recent admixture of two populations, and the length of the region of LD can be large when few generations are available to allow its decay [1]. However, there is debate about the advantage of isolate populations in identifying genes affecting complex traits [27], since LD in isolate and outbred populations has been shown to be similar in some studies.

Nevertheless, differences in LD between isolated and outbred populations have been documented, particularly for SNPs, in addition to their reduced genetic heterogeneity, which can significantly increase genotypic relative risk [28]. Thus, Finnish, Sardinians and Ashkenazi populations present a higher LD level for SNPs to a distance greater than 200 kb.

Complex traits are expected to be more homogeneous in isolate populations because of the small number of founders and genetic drift, increasing the power of linkage analysis. This is because fewer genes will be contributing towards a disease trait, and therefore the effect of each susceptibility gene will be easier to detect [29].

6. Antioquia (Colombia)

Colombia is located in the Northwest of South America. The country is politically divided in 32 departments (provinces) and a capital district, Bogotá. The great majority of Colombians are catholic and their language is Spanish.

One important department in Colombia, by its economical and political status, is Antioquia (Figure 1.1). This is located in Northwest Colombia between the central and western branches of the Andean mountains. It is estimated that Antioquia currently has 4,500,000 inhabitants. The metropolitan area of its capital, Medellin, has approximately 3,000,000 people. People originated in Antioquia are called "paisas".

It has been reported that Europeans founded the first non-Amerind settlement in what is present-day Colombia in the early 16th century [30] allowing the establishment of a "criollo" population.

Several lines of genetic evidence suggest that Antioquia is an isolate population. First, using classic markers (blood-groups), the ancestral ethnic component was estimated as 85% Caucasian and 15% Amerindian [31]. The genetic contribution by negroids was essentially zero. Second, strong admixture distortions in the gender vectors of the original founders of the paisa community was found by using DNA markers (STRs and SNPs). This study concluded that ~94% of the Y chromosome are European, 5% African and 1% are Amerind [32]. Similarly, they found that ~90% of the mtDNA (mitochondrial DNA) in Antioquia is of Amerindian origin. These results indicate a highly asymmetric pattern of mating in early Antioquia, involving mainly European men and local native women [32]. Further analysis of DNA markers in Antioquia, provided evidence of LD patterns suggested by the sex bias founding, which can be exploited for the identification and fine mapping of disease variants in complex disorders [33].

Figure 1.1. Map of Colombia showing Antioquia in green colour.
Map modified from http://www.crwflags.com/fotw/misc/co(.gif

In addition to the sources of LD in Antioquia, another special feature of this population

is the large number of children per family. This was characteristic until the last

generation that in conjunction with the homogeneity of the population makes it

interesting for gene mapping purposes. By using large families, two founder effects

have been described in Antioquia for two Mendelian forms of the complex disorders

Early-onset Alzheimer's [34, 35] and Juvenile Parkinson's [36] diseases. Both Mendelian

12

and complex traits could be assessed in this population by combining linkage analysis and LD methods.

7. Aims

My thesis aimed to make progress in the genetic characterization of several Mendelian forms of Complex disorders, including Parkinson's disease, Generalized Epilepsy with Febrile Seizures Plus [GEFS+], Batten's disease, Type 1 Diabetes [T1D], Multiform Movement Disorder [MMD] and Abdominal Epilepsy. All the samples analyzed here were collected from individuals from Antioquia-Colombia.

Specific aims

- To evaluate candidate genes/loci by linkage analysis

- To examine the best candidate loci by SSCP and/or sequence analysis

- In some cases to conduct a genome wide scan, to identify the chromosomal region where the mutant gene lies

- To undertake fine mapping of regions where novel loci are identified

CHAPTER TWO

METHODS

Several different methods have been used in developing this thesis work. The great majority of methods employed here are related to linkage mapping. Thus, two-point linkage analysis (including power simulations), multipoint analysis and haplotyping were the basis of my thesis. Here I present some basic concepts of the methods used for analysis of the families studied and the data collected in them. Common protocols among chapters are described in the final part of this chapter.

1. Methods of analysis, Basic Principles and implementation

1.1 Linkage Analysis

Gene linkage is a measure of the coinheritance of two or more loci in a family which occurs when the loci are in close proximity and there is little recombination between them. To better understand this concept, let us assume that a father transmits his paternal chromosome 1 to his child and that the mother transmits her paternal chromosome 1 to their child too. On this chromosome we are looking at two marker loci, M1 and M2. In Figure 2.1 it is easy to notice that the father has inherited from his mother alleles 2 and 3 at marker loci M1 and M2; this determines his phase as 12/23 as shown in Figure 2.1A, indicated by a vertical bar. Also, a similar analysis allows us to identify the mother's phase as 31/23.

·If a recombinational event occurs, as result of crossing over during the meiosis, the child will receive a combination of chromosome 1 portions from his grandparents through his parents.

M₁ 1|2 2/2 3|2 2/2

M_1 1|2   2/2   3|2   2/2
M_2 2|3   3/3   1|3   3/3
Father   Mother
M_1 1|2   3|2
M_2 2|3   1|3
Child
M_1 1|3
M_2 2|1
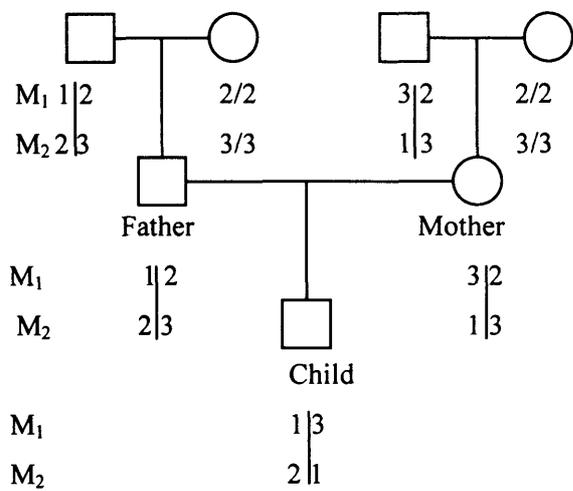
Figure 2.1A. Phase known in a three generations pedigree. Hypothetical data for two marker loci.

1|2   2/2   3|2   2/2
2|3   3/3   1|3   3/3
Father   Mother
1|2   3|2
2|3   1|3
Child
1|3
3|1
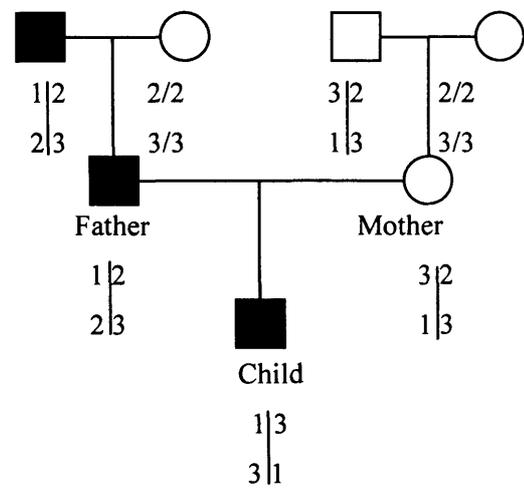
Figure 2.1B. Phase known pedigree, showing a recombinant child for M2. Filled symbols indicate individuals presenting with a certain dominant trait (disease).

In Figure 2.1B the child has received alleles 13 from his father, which represents a combination of both his paternal grandparents' chromosomes 1. Thus, the recombination happened in his father and the child is said to be a recombinant for marker locus 2 (M2) [37].

The rate of recombination (θ) is a function of the genetic distance. If k times a recombination event is observed between these two loci among n observations (meioses), it would provide that the estimated theta is

$$\hat{\theta} = \frac{k}{n}$$

16

For example, if we had observed that in Figure 2.1B there were four children instead of one and that only the one shown was actually a recombinant, it would produce an estimated theta

$$\hat{\theta} = \frac{1}{4} = 0.25$$

As this value is lower than 0.5 (free recombination), it indicates linkage between M1 and M2. Also, if we consider the same situation for M1 and the trait as for M1 and M2, it would indicate $\hat{\theta} = 0$, i.e. no recombination between M1 and the trait locus. As the father is fully informative i.e. doubly heterozygous (heterozygous for M1 and the trait locus) and as we have already identified his phase, he provides the maximum information for linkage.

The lod score method

While in animal systems the statistical approach to evaluate linkage usually involves chi-square, in humans more complicated calculations are required. This extra difficulty arises because of the variability of family structure, limitations in family size, and the limitation to determination of phase in doubly heterozygous parents.

The lod score method was developed by Morton (1955) [38] by integrating the probability test of Haldane and Smith [39] with the sequential sampling developed by Wald [40].

Two steps are required to calculate the logarithm of odds (lod). First, the likelihood of the pedigree is calculated considering the number of recombinants and

nonrecombinants. Then, the common logarithm is calculated for the likelihood ratio. Application of the common logarithm allows adding up the resulting scores across families [41].

Assuming that we have three non-recombinants and one recombinant (in our example), the probabilities for each one is calculated separately. The probability 1- θ for these three non-recombinants to happen independently is $(1-\theta)^3$. The probability for the recombinant is θ^1.

In a general form, it is:

$$L(\hat{\theta}) = \theta^R (1-\theta)^{NR}$$

R= Recombinant; NR= nonrecombinant

If this likelihood is divided by its value under θ =0.5, free recombination, the likelihood ratio is obtained, which is also called the odds ratio.

$$L*(\hat{\theta}) = L(\hat{\theta})\big/L(0.5)$$

In our example, it would be:

$$L*(\theta = 0.25) = \left[\theta^1(1-\theta)^3\right]\big/(0.5)^4$$

Then, the common logarithm of the odds is the lod score (Z).

$$Z(\hat{\theta}) = \log_{10} L(\theta)\big/L(0.5)$$

If a lod score ≥ 3 is found, it represents statistical evidence in favour of linkage. If a $Z\leq-2$ is found, it represents exclusion of that locus for either a Mendelian or complex trait [42].

Phase unknown

When it is not possible to establish phase (i.e. two-generations pedigree or both parents are doubly heterozygous for same alleles), it is not possible to count recombinants and nonrecombinants, since both situations have equal a priori probabilities. Then, estimation of the recombination fraction is done by the maximum likelihood method, which in linkage analysis is found numerically by varying the values of the parameter (θ) until an approximate maximum is found [37].

If phase was unknown, the actual likelihood ratio needs to be multiplied by 0.5.

$$Z(\hat{\theta}) = \log_{10}\left\{\frac{1}{2}\left[\frac{\theta^R (1-\theta)^{NR}}{(0.5)^R (0.5)^{NR}}\right] + \frac{1}{2}\left[\frac{\theta^{NR}(1-\theta)^R}{(0.5)^{NR}(0.5)^R}\right]\right\}$$

Multipoint analysis

An extension of the two-point analysis is the multipoint lod score analysis. This consists of analyzing a fixed map of markers and a trait locus. Thus, the order of the markers does matter now and the analysis consists of testing for linkage of the trait locus along the set of markers from one end through to the other. Multipoint analysis has certain advantages over two-point analysis. First, a locus that was originally uninformative can become informative through haplotype information. Second, in fine mapping, multipoint analysis is very useful to pinpoint a disease locus [1]. The greatest

utility of multilocus analysis is locating the linked disease, providing that the mode of inheritance is well known [43]. The multipoint lod score is defined as:

$$Z(x) = \log_{10}[L(x)/L(\infty)]$$

where L(x) is the likelihood that a disease locus is located at a distance X on a fixed map consisting of several markers. $L(\alpha)$ is the likelihood that the disease locus is not on the map ($\theta = 0.5$).

The above likelihood ratio often is expressed as a location score, which is a simple multiple of the lod score. Thus, $S(x)=Ln[L(x)/L(\alpha)]$ [44].

The maximum multipoint lod score possible is equal to the maximum possible two-point lod score when all meioses are informative [43].

Models in a linkage analysis

At its most elementary level, a model postulates the number of loci necessary to explain the phenotypes. Any Mendelian model deals with three different probabilities: Priors, penetrances and transmission probabilities [45]. Models contain variables called parameters, whose values are estimated based on observations. When the families under study show a clear mode of inheritance, the so called parametric methods can reliably be used in analyzing the data obtained. The most critical parameters are mode of inheritance, penetrance, allele frequencies and phenocopy rate. Among these, the most drastic one is the inheritance pattern. For example, if a particular trait segregates in an autosomal recessive fashion and mistakenly is assumed to be autosomal dominant, dramatical differences in the actual lod score would be observed. If the family shown in Figure 2.2, which presents a child affected by a recessive disorder, were analyzed assuming dominance, very negative lod scores would be obtained even

20

at low penetrance (50%) (Z= -1.44). Using recessive conditions (penetrance 100%) and same allele frequencies in this family the lod score was 1.703 at θ= 0 (Table 2.1). Therefore, if mode of inheritance is not clear, a non-parametric analysis method is instead recommended. An alternative is to analyze the data collected using different models, where the inheritance pattern, penetrance and phenocopy rate can be varied. Allele frequencies are relatively easier to estimate than the other parameters. Estimating gene frequencies will prevent large fluctuations in lod scores. Population frequency of each marker and disease allele is required for the computation of the likelihood [43].

If every founder in the pedigree has unambiguosly determined genotypes, then estimating the allele frequencies will have no effect on the calculation of the lod scores. Allele frequencies are normally estimated from a sample of a normal population; they also can be estimated from the founders of a large pedigree or set of pedigrees, taking them as an independent sample. Alternatively, the ILINK program of the LINKAGE package can also be used to estimate allele frequencies when some of the founders have not been typed [46].

Table 2.1 shows that the allele 1 frequency (at the marker locus) has more effect on the analysis than the frequency of the disease allele, especially when the disease allele frequency is smaller.
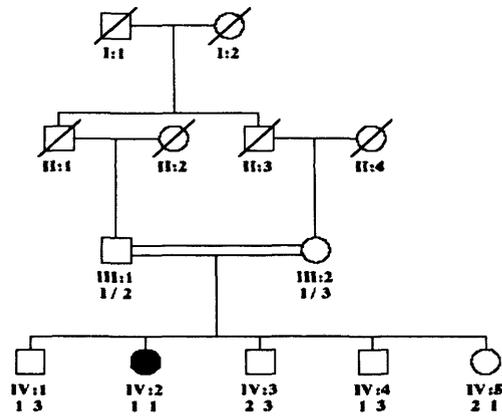
Figure 2.2. Family segregating a recessive trait. Genotypes are available for III and IV generations. All four founders are missing their genotypes.

Frequency of allele 1	Frequency of disease allele			
	0.5	0.1	0.01	0.0001
0.9	0.531	0.585	0.624	0.632
0.5	0.558	0.789	0.835	0.858
0.1	0.677	1.06	1.322	1.38
0.01	0.8	1.333	1.611	1.654
0.0001	0.828	1.38	1.66	1.703

Table 2.1. Maximum Lod scores obtained for family shown in Figure 2.2, at θ= 0.

This happens because the penetrances at the disease locus tells us more about about the disease locus genotypes of the untyped individuals than we know about their marker locus genotypes. The fact that we have less information for the genotypes at the marker locus of the untyped individuals makes the gene frequency particularly relevant [43].

. Software used to calculate Lod scores and haplotypes

The LINKAGE package [44] was mainly used to calculate two-point lod scores. This package is based in the pioneering algorithm of Elston and Stewart [45].

22

To calculate two-point lod scores MLINK from the package LINKAGE was used [44]. It was assumed that the allele frequencies were 1/n, where n= number of alleles.

To calculate multipoint lod scores in the Multiform Movement Disorder family, GENEHUNTER v 2.1 [47] was used. This software uses the Lander-Green algorithm improved by Kong and Cox [48]. The largest a pedigree that can be run in GENEHUNTER is 2N-F$\leq$16, where N= number of non-founders and F= number of founders [47].

Simwalk2 [49] was used to obtain both multipoint lod scores in the rest of families analyzed here as well as to find the most likely haplotypes for the entire set of families. Simwalk2 uses the Markov Chain Monte Carlo algorithm [49]. To prepare Simwalk2 input files, MEGA2v3.0_r4 was used [50].

Finally, haplotype of pedigrees were drawn using HaploPainter [51].

Checking for Mendelian inconsistencies

After collecting genotypes in family members, checking for Mendelian inconsistencies is always recommended. This should be done before analyzing the data. To evaluate whether alleles transmitted to children in every nuclear family in a pedigree are consistent with Mendel's laws four particular situations are tested: The alleles of a child and a parent should be compatible, the child should be compatible with each parent separately when both parents are considered simultaneously, there should be no more than four alleles in a sibship, and no more than two alleles in a sibship when two of them are homozygotes for two different alleles [52]. Pedcheck implements four

different algorithms to evaluate four different error levels [52], which can detect inconsistencies that other software for checking Mendelian inconsistencies do not detect, such as UNKNOWN [53] of the LINKAGE package [44]. UNKNOWN may fail detecting inconsistencies in very complex pedigrees [54].

1.2 Power simulations

When beginning a gene mapping study, one may wish to know beforehand which individuals in the pedigree (s) are those that eventually will provide the most information at the time of analzing the genotypes; also one may wish to know the power of the set of families to reach a statistically significant lod socre. These are questions that can be answered by means of a power simulation. A simulation is based only on the structure of the pedigree, the number of affected individuals and the informativeness of the markers. No biological samples have yet been collected.

Successive risk calculations are the principle of the simulations. They calculate the conditional probability distribution, $P(g/x)$, of the genotypes given the phenotypes in a pedigree with N people.

So, if $x=(x_1, x_2, ...x_N)$ represent the vector of phenotypes of the N inviduals and if $g=(g1, g2, ...gN)$ represent the vector of multilocus including phase information, then

$$P(g/x) = P(g_1/x)P(g_2/g_1,x)P(g_3/g_1,g_2,x)... \text{ [37,55]}.$$

Various factors influence the possibility of establishing linkage, such as penetrance, heterogeneity, and informativeness of markers.

Marker heterozygosity determines the frequency with which a parent is potentially informative (double heterozygote) [37]. Heterozygosity is the probability that an

24

individual is heterozygous for two alleles at a marker locus with given allele frequencies. It is defined as H, where $H = 1 - \sum p_i^2$, and p_i is the population frequency of the ith allele. For a gene with a alleles, H is largest with equally frequent alleles, $p_i = p_j = 1/a$. Then, heterozygosity takes the simple form, $H = 1 - 1/a$[37].

For example, a locus with four equally frequent alleles has a maximum heterozygosity of H= 0.75. From here, it is worth mentioning that the number of alleles does influence heterozygosity. For example, a SNP will only provide a maximum H= 0.5 when both its alleles are equally frequent. In contrast, since STRs normally show more than four alleles, it is common to find that their heterozygosity is greater than 0.75.

Simulations are run at a given theta ($\theta<0.5$), "true recombination fraction". Every pedigree in the simulation is replicated by a given number of times (one hundred to one thousand). In every replicate, the marker alleles are different but will segregate in a Mendelian manner. The so called ELOD is the expected lod score at that predetermined θ, and is calculated by dividing the sum of the lod scores by the number of replicates. This value determines the average expected lod score for that particular family if the model used in the simulation nearly corresponds to what will be obtained in the genotypes. The maximum lod score is the highest lod score obtained at any value of recombination fraction in the entire set of replicates. This corresponds to the highest potential lod score when all meioses are fully informative under the assumed model and allele frequencies.

The power refers to the percentage of replicates where a given lod score (e.g. Z≥3) was obtained at the true recombination fraction, per pedigree.

Power Simulation Software

There are at least three software packages that can be used in estimating the power of a set of pedigrees. They are SIMLINK [56], SIMULATE [57] and SLINK [55]. They all have common features and differences. For example, all three programs allow the simulation of a marker locus unlinked to a disease locus. SLINK and SIMLINK, however, allow the simulation of genetic marker data for a locus linked to a disease locus.

SLINK was used to estimate the power of the families studied here. All simulations were carried out in 100 replicas of the pedigrees. In each case, a marker locus with four alleles equally frequent was simulated.

2. Laboratory methods

2.1 Genome wide scans

Markers across all the genome were typed for both Multiform Movement Disorder and Abdominal Epilepsy.

Samples were prepared to 100 ng/µL in 130 µL making a total of 13000 ng of DNA and then were shipped to the Icelandic company deCODE for markers to be typed. Five hundred and fifty two markers were typed at an average density of six cM. Marker repeats consisted of 490 dinucleotides, 57 tetranucleotides and 5 trinucleotides. Their average heterozosity was 70.36. Appendix 1 shows a list of the markers typed by deCODE including their genetic position, cytogenetic band and heterozygosity.

2.2 Microsatellite marker typing

Different fluorescent labels were used to allow pooling and load them together onto a gel in an ABI-377 sequencer (Applied Biosystems).

STRs were amplified in an MJ Research (DYAD^{TM}) thermocycler, in a final volume of 10 μL using 30-100 ng of genomic DNA, 0.33μM of each primer and 1X of Qiagen HotStarTaq Master Mix Kit. Their PCR conditions were as follows: 94°15' followed by 10 cycles of 94°30" 55°30" and 72°30"; then 20 cycles of 89°15"55°15"72°30" and a final extension at 72° for 10 min. Once the fragments were run on an ABI-377 sequencer (Perkin-Elmer), the allelic sizes were determined using Genescan 2.1 and Genotyper 3.5 software (Applied Biosystems).

2.3 Single Strand Conformational Polymorphism analysis, SSCP

SSCPs were run in 10% non-denaturing polyacrylamide gels. This procedure involved five steps:

1) Preparation of the gel (10%). The acrylamide solution was prepared from a stock at 50% acrylamide-bisacrylamide (37.5:1) (BioRad), adding 5 mL of this solution to 1.25 mL 10X TBE (Tris-Boric Acid EDTA) buffer. Plus 18.75 mL of water and, to accelerate the polymerization reaction, 100 μL of ammonium persulphate (APS) and 40 μL of (TEMED) NNN'N'-Tetramethilehtilenedia mine (BDH) were added. The gel was prepared in a vertical electrophoresis device with spacers of 0.75 mm.

2) Each sample contained 3 μL of water, 4 μL Formamide, 1 μL of PCR product. Samples were denatured at 94°C for 5 minutes and then were chilled on ice.

3) Run electrophoresis was carried out using 0.5X TBE in the gel tank, and 300 V for 4 hours at room temperature.

4) Staining. Silver staining was performed using a standard protocol that is used in the lab: Fixation involved using 200 mL of 10% ethanol, 0.5 % acetic acid for 3 minutes. The solution was then poured off and this process repeated once again.

For staining: 200 mL of 0.1% $AgNO_3$ for 3 minutes (twice) was used.

The staining was developed using 200 mL of 1.5% NaOH, 20 mg $Na(BH_4)$ and 0.8 mL 37% formaldehyde. Bands appeared in 5-20 minutes.

Drying gels was done after rinsing the gels three times in distilled water and transferred them onto a piece of 3MM paper. Drying of the gels was carried out in a Gel Dryer Model 583 (Bio-Rad) for 2 hours.

2.4 Sequencing

Following a PCR purification step for removing primer excess, dNTPs, Taq and buffer using the QIAquick PCR purification kit (QIAGEN), sequence of bases was performed using the Big Dye Terminator v3.1 Cycle sequencing kit (Applied Biosystems). Sequence reactions were performed in an MJ Research (DYAD^{TM}) thermocycler in a final volume of 20 µL according to manufacturer specifications. After purification in columns from the QIAquick PCR purification kit (QIAGEN), samples were dissolved in 3 µL of loading buffer containing 2.5 µL formamide plus 0.5 µL of bromophenol blue and run in a polyacrylamide gen 6% on an ABI-377 (Perkin-Elmer).

CHAPTER THREE

PARK2 AND EARLY ONSET

PARKINSON DISEASE

1. INTRODUCTION

1.1 Generalities

Parkinsonism is a feature of several distinguishable entities such as Parkinson's disease, Lewy bodies disease (Dementia Lewy bodies) and Tauthopathies.

James Parkinson made the first description of Parkinson's disease (PD) in 1817 [58]. The disease is manifested as a movement alteration characterized by bradykinesia, tremor, and postural instability [58]. In general the response to the therapy (L-dopa) is good but "on-off" phenomena and induced treatment dyskinesia rapidly appear [59]. Classical (or late) onset PD presents at $\geq$ 50 years. Parkinsonism in the first decades of life was in the past considered to be distinct from PD; these cases are now considered as two groups: Early onset PD (EOPD) or Juvenile PD (JPD). EOPD presents age at onset <45 years, while JPD presents age at onset $\leq$ 25 years. PD affects over 1% of the general population older than 55 years [60].

Symptoms associated with PD appear due to the degeneration of the dopaminergic neurons in various regions of the central nervous system including the substantia nigra, the locus coeruleus and basal nuclei [61]. Microscopic Lewy bodies within these nuclei are the pathological hallmark of PD. Neuronal death is accompanied by the appearance of cytoplasmic Lewy body inclusions, which are composed of a range of proteins including α-synuclein, ubiquitin, neurofilaments and oxidized/nitrated proteins [62]

Genes recently identified play an important role in the mechanisms of Parkinson pathology. It is currently known that Parkin (encoded by *PARK2* gene) has a ligase function of proteins ubiquitinated and that mutations in it lead to loss of function reducing its ability to regulate degradation of substrate removal [63]. It has also been observed that an O-glycosylated form of α-synuclein is ubiquitinated by Parkin, suggesting that loss of Parkin function will result in α-synuclein accumulation [64]. One patient with parkin mutations has been reported as Lewy body positive [65]. In addition, overexpression of mutant α-synuclein produces an inhibition of proteosome-associated proteolytic activities [66].

It was believed until recently that environmental factors such as viral infections [67] and neurotoxins [68] were the main cause of Parkinsonism. However, the identification in recent years of mutations in six different genes (*SNCA* [69] *PARK2* [70], *UCHL1* [71], *PARK7* [72] *PINK1* [73] and *MAPT* [74]) and of at least six different loci linked to the disorder (*PARK3* [75], *PARK4* [76], *PARK8* [77], *PARK9* [78], *PARK10* [79] and *PARK11*[80]) project the importance of genetic factors (Table 3.1).

Considering their aetiology as well as their pattern of segregation, several forms of Parkinsonism can be identified. There are both a sporadic type (or idiopathic) and familial forms. Genetic risk factors probably are also involved in the idiopathic forms [81]. Sporadic cases have been reported as carriers of mutations in genes identified in some familial forms [82, 83] (Table 3.1), but no mutations have yet been associated with the classical form of the disease, which is described as a complex trait. More recently, two separate studies in families have found linkage to 1p32 (*PARK10*) [79, 84]. Because of the late age at onset of the patients involved in these two studies, it seems likely that

this locus is more frequently involved in the classical form of the disease than any of the other genes/loci previously identified.

Type	Chromosome	Protein	Gene/Locus	References
I. Sporadic	4q21.1	α-Synuclein	*SNCA/ PARK1*	85
	6q25.2	parkin	*PARK2*	70, 86, 87
	1p32	*?*	*PARK10*	79, 84
II. Familial				
Dominant	4q21.1	α-Synuclein	*SNCA/PARK1*	85
	4p	UCH-L1	*UCHL1/PARK5*	71
	2p13	?	*PARK3*	75
	4p14-16.3	?	*PARK4*	76
	12p11.2-q13.1	?	*PARK8*	77
	17q21.1	tau	*MAPT*	74, 88
	1p32	?	*PARK10*	79, 84
	2q36-37	?	*PARK11*	80
Recessive	6q25.2	parkin	*PARK2*	70, 86, 87
	1p35-36	Pink1	*PINK1/PARK6*	89
	1p35	dj1	*DJ1/PARK7*	72, 90
	1p36	*?*	*PARK9*	78

Table 3.1. Genetics of Parkinsonism

1.2 Autosomal Dominant Parkinson's Disease (AD-PD).

PARK1

In 1996, reports of AD-PD began to appear by Polymeropoulus and colleagues. Through a genome wide scan approach in a large Italian kindred and using 140 genetic markers, with an average density of 20 cM, they found that markers localizing in the region 4q21-23 were the only ones showing linkage to the disease with a maximum lod score (Zmax)= 6.0 at a recombination fraction (θ)= 0 for marker D4S2380 [85]. Through

saturation of markers in this region and having observed some recombinants between the genetic markers and the disease it could be established that the locus *PD1* (known today as *PARK1*) was located between markers D4S2371 and D4S2986 constituting a six cM interval. It was known at that time that there were four genes within this region which should be sequenced to search for disease causing mutations; these genes encoded for: Alcohol dehydrogenase, formaldehyde dehydrogenase, α-synuclein and UDP-N-acetyl glucosamine phosphotransferase[91]}. The four exons of *SNCA* were examined revealing a transition in its fourth exon, consisting of G to A in position 209 of its cDNA (c.209 G>A), which causes an aminoacid substitution of Ala for Thr in position 53 of the protein. This substitution could be tested simply by means of a screening test using the restriction enzyme Tsp45 I, which targets the mutant sequence. In addition to the Italian family, three Greek PD families carried this mutation. None of 314 control chromosomes, of Italian origin, carried the mutation [69].

The substitution Ala53Thr lies in a region of the protein where the secondary structure predicts an α-helix formation bounded by beta sheets. This mutation interrupts the α-helix and extends the beta sheet structure, which is involved in self-aggregation and could lead to the formation of amyloid plaques or Lewy bodies [92, 93]. A second mutation was found a year later in a German family, which consists of a nucleotide substitution of G for C in position 88 resulting in an Ala30Pro mutation [94]. More recently, another family was found to carry a triplication of a large region that contained the gene [95]. Last year a fourth mutation was reported in SNCA [96]. These findings further support the implication of α-synuclein in the pathogenesis of the disease. Interestingly, the Ala53Thr is not a conserved aminoacid among different species. In the alignment of human α-synuclein with its homolog from rat (*Rattus nervecus*), zebra finch (*Serinus canaria*), cow (*Bos Taurus*) and *Torpedo californica* it

was found that, in this particular position, only *Bos Taurus* had an Alanine in this residue. The others had either Threonine (Thr=T) or Asparagine (Asn=N).

PARK5

Another gene identified as responsible for AD-PD is the one encoding Ubiquitin carboxy-terminal hydrolase L1 (*PARK5*), in which a missense mutation (Ile93Met) has been reported in exon 4, leading to a partial loss (~50%) of the catalytic activity of this thiol-protease [71]. This change can be tested using BsmFI since the mutation (c. 277 C>G) creates a new site for the enzyme. None of 500 control chromosomes carried the mutation.

UCH-L1 comprises over 2% of the total brain protein [97], and belongs to a family of deubiquitinating enzymes. It is believed that it cuts the ubiquitin polymers breaking them down to monomers and hydrolyses bonds between ubiquitin molecules and small adducts such as glutathione and cellular amines [98]. Also, its presence has been reported in Lewy bodies [99].

These three features (abundance in brain, presence in Lewy bodies and being involved in the ubiquitin-dependent proteolytic pathway) show that UCH-L1 fits in the picture of PD pathogenesis. UCHL1 mutation contributes to the genetic aetiology of just a small number of patients with familial PD. Mutations in both *SNCA* and *UCHL1* present incomplete penetrance (~85%) [69, 100].

MAPT

A mutation in yet another gene was found in a family with frontotemporal dementia and Parkinsonism (FTDP-17). A splice-site mutation in *MAPT* causes the disorder by destabilizing a putative stem-loop structure involved in regulating the alternative

splicing of exon 10 [74]. In addition, Hutton and colleagues found missense *MAPT* mutations in families with FTDP-17 [88]. Even though this phenotype is very variable, these families show a phenotype that is very similar to idiopathic PD [101].

There are several AD-PD loci where no genes have yet been isolated. This is the case for *PARK3, PARK4, PARK8, PARK10* and *PARK11* [75-77]; Hicks, 2002; Pankratz, 2003; Li, 2002}.

1.3 Autosomal Recessive Parkinson Disease (AR-JP).

PARK2

In order to delineate the molecular mechanisms of EOPD, there was great interest in the *SOD2* gene, which encodes the manganese superoxide dismutase (MnSOD) [86], which is an intramitochondrial enzyme that acts on the superoxide anion originated in the mitochondrial respiratory chain [98]. The following three considerations suggested the presence of oxidative stress in the mitochondria of the brain of PD patients and a possible role of MnSOD modulating this process.

1) Studies on the brain indicated the alteration of function of the mitochondrial respiratory chain in PD [102-104].

2) That there could be a causal relation between the inhibition of the functions in the respiratory mitochondrial chain and the release of superoxide anions in the mitochondria [98].

.3) That the MnSOD is inducible by superoxide anions [105] and that its increased activity in the substantia nigra of PD brain had been demonstrated [106].

These hints led Matsumine and colleagues to sequence the five exons of the *SOD2* gene in 13 Japanese families as part of a functional candidate gene approach. Only one

family was found to present perfect segregation of the disease with a genetic variant. This variation was located in exon 2 and consisted of a transition of C to T leading to the aminoacid substitution Ala9Val. In light of this observation they decided to evaluate ten polymorphic markers (STR's) around *SOD2*, covering 17 cM in the chromosomal region 6q25.27 and tested for genetic linkage in those 13 families finding a Zmax= 7.26 and 7.71 for D6S305 (θ=0.03) and D6S253 (θ=0.02) loci, respectively [86]. These two markers were rather far from *SOD2*.

Simultaneously, Jones and colleagues were evaluating families from four different ethnic groups using the genome wide approach with an average density of 22 cM [87]. They found linkage to the same chromosomal region as Matsumine and colleagues, and further evaluation of additional markers in the region let them exclude the *SOD2* gene by observing two recombinant individuals, narrowing down the candidate region to 6.9 cM between markers D6S305 and D6S1599 [87].

At the same time another group was evaluating BAC's (Bacterial Artificial Chromosome) corresponding to the same region and by using exon trapping methodology and sequencing cDNA, they identified the *PARK2* gene, which has 12 exons and spans 1.5 Mb. Its product (Parkin) has 465 aminoacids with a moderate similarity to Ubiquitin at the amino terminus, and at the C-terminus end to a RING-finger motif [70]. Four families with deletions of exons were reported in this study. In one family there was a deletion of five exons (3-7) and another three unrelated families had deletions of exon 4. *PARK2* is expressed in various tissues but it is abundant in the brain including the substantia nigra [70]. Since then more families and sporadic cases have been reported as due to *PARK2* mutations. To date, 41 exon rearrangements including 27 different deletions, 13 duplications and 1 triplication have been reported [83, 100, 107-115] and 59 point mutations have also been reported [36, 83, 107, 108, 110, 113, 116-126]. It

is now clear that rather than deletions, point mutations are the commonest cause of Parkinson disease due to *PARK2* mutations. It has also been observed that the most frequent deletion involves exons 3-4 and that the most frequently reported point mutation is c.924 C>T in exon 7 (reviewed in [127]).

Genetic heterogeneity

Locus and allelic heterogeneity have been observed for both AD-PD and AR-PD.

For AR-JP the first evidence of locus heterogeneity arose when analyzing haplotypes in two branches of a Greek family with JPD. It was evident that patients in one of the branches carried *PARK2* mutations (Ex5-7 del) and patients from the other branch did not, nor did they share a haplotype or show linkage to this gene [100]. In addition, it has been reported that approximately 50% of the AR-JP families are due to mutations in *PARK2* [83] implying that the remaining families must have a different causal gene and consistently two other loci have already been identified on chromosome 1p (Table 3.1). These two loci are *PARK6* [89] and *PARK7* [90]. In the latter, another gene involved in the ubiquitin-proteasome pathway has recently been identified [72] (Table 3.1).

Most of AD-PD patients do not carry mutations in either *PARK1* or *PARK5* and it has more recently been shown that new loci are involved in this form of Parkinsonism [74-77].

1.4 Preliminary results

Even though different genes are involved in EOPD, the most frequently implicated is *PARK2* and, furthermore, there are several reports suggesting that even being a carrier of *PARK2* mutations is a risk factor for being affected by Parkinson disease [65, 83, 110, 113, 118, 128, 129]. Not all these reports evaluated the entire *PARK2* gene nor used methods

to test for dosage effects, such as deletions or duplications. They have mainly used SSCP and sequence, which might allow for some deletions or multiplications to go undetected. Also, all these studies lack familial structures that could allow haplotype reconstruction to identify sights of compound heterozygous in families with multiplex affected individuals.

Prior to the start of this thesis, two founder effects in Antioquia-Colombia for two *PARK2* mutations were reported in three multiplex families and one sporadic case with EOPD [36]. Patients from families PJF1 and PJF3 were homozygous for a G to A transition in exon 6 at position 736 (c.736 G>A). This mutation leads to the substitution of Cys for Tyr at position 212 of parkin. In addition, in family PJF2 a homozygous GT insertion between nucleotides 321 and 322 (321-322insGT) in exon 3 was detected in all affected individuals. The sporadic case was also found to be homozygous for this mutation [36]. In addition, three Colombian sporadic-like patients were shown to be carriers (heterozygous) of the c.736 G>A mutation (A.R.L unpublished data) and they were born in the same town (Peque-Antioquia) as family 3 (PJF3). Therefore, it was interesting to know whether those affected individuals originating in same small town (Peque) and carrying the same *PARK2* mutation are actually members of a same extended pedigree.

On the other hand, in 2002 a Spanish group reported, in their patients, the mutation c.736 G>A, previously reported in Antioquia [117] raising the question whether the observation of same mutation in Antioquia was due to recurrent mutation or to a common ancestor.

1.5 Aims

The focus of this chapter was to evaluate whether the three sporadic-like patients from Antioquia-Colombia are compound heterozygous involving a third mutation in the population or were just carriers. Also, given that the mutation c.736 G>A was separately reported in Spain [117] and Antioquia to evaluate whether this observation was due to a common ancestor or alternatively was due to a recurrent mutation. In addition, in order to examine the extent of the founder effect for *PARK2* in Antioquia, I performed the molecular characterization of 23 additional sporadic EOPD patients from the province.

2. Materials and Methods

2.1 Patients

Patients were recruited from the Servicio de Neurologia of the Universidad de Antioquia (NEUROCIENCIAS). 23 sporadic patients with early-onset PD from the provincial capital (Medellin) were included in this study. These patients have no known relationship to each other and originate in various parts of the province. To be able to contrast with the sporadic cases, 17 control trios from Antioquia were also included.

In addition three different sporadic-like cases and other four patients (from 6 nuclear families) were born in Peque. Also, two Colombian families (PJF1 and PJF3) [36] and one Spanish family [117] carrying the mutation c. 736 G>A were studied. In addition

two families where segregation of the 321-322insGT mutation was reported [36] were also studied.

2.2 Screening for the C212Y mutation

The missense mutation G736A, resulting in the Cys212Tyr substitution in *PARK2* was tested in all familial and sporadic cases using the forward primer sequence (C212Y) 5'TCTTTCAGGAATTTTTCTTTA<u>C</u>AT3', which includes a mismatch identified by the underlined base. This mismatch in an individual with the mutation will complete the core for the restriction enzyme NdeI. The reverse primer was Exon 6R (Table 3.2) (Kitada et al. 1998). PCRs were performed in an MJ Research (DYAD™) thermocycler in a final volume of 25 microliters with 30-100 ng/µL of genomic DNA, 0.33 µM of each primer, 1X of BioTaq buffer, 1.5 mM of Magnesium, 0.2 mM of each dNTP. PCR conditions included 94o 5' of initial denaturing followed by 10 cycles of 94°15", 48°15" and 72°30"; then 20 cycles of 89°15", 48°15" and 72°30", followed by a final elongation step of 72°10'. Digestion with the restriction enzyme NdeI was done overnight using 5-10 units per sample according to manufacturer specifications (New England Biolabs). Digested PCR fragments were examined in a 3.5% agarose gel stained with ethidium bromide.

2.3 Screening for the 321-322 insGT in Exon3

To type this mutation a fluorescence based approach was used. The forward primer for exon 3 was labelled with the dye 6-FAM in the 5' end and runs were performed on an ABI-377 (Applied Biosystems).

In order to appropriately compare carriers, homozygous and non-carrier controls were included. The expected normal size is 311 bp.

2.4 Microsatellite marker typing

Microsatellite markers typed are shown in Table 3.3. All eleven markers were evaluated in four Colombian families (PJF1, PJF2, PJF3, PJF4) and the Spanish family. In addition, 17 control trios and 23 sporadic cases were tested for 4 of these eleven markers (D6S1550, D6S253, D6S305, D6S980).

2.5 Haplotype analysis

It was performed for the four Colombian families and the Spanish family studied here. In addition, the control trios were analyzed to infer haplotypes in their children. Please refer to the methods chapter.

2.6 SSCP

SSCP was performed either to confirm the genotype for c.736>G or in the rest of exons. Please refer to methods chapter.

2.7 Sequencing

Sequencing was conducted on exons that showed a variant conformation on SSCP. Please refer to methods chapter.

2.8 Population allele-specific markers (PAS)

In order to examine the major component of the Peque population (i.e. Amerindian or European), six PAS were evaluated in 46 unaffected individuals from Peque (Table 3.4). These markers were MID93, MID472, MID1326, MID1327, MID1338 and MID1705. PCR conditions used the following touch down: $94°$ for 15 min, followed by ten cycles of $94°$ for 15sec, $63°(-0.5o$ each cycle) for 30 sec, $72°$ for 30sec; finally, twenty cycles of $89°$ for 15sec, $52-58°$ for 15sec, $72°$ for 30sec. A final extension step at $72°$ for 10 min, was added.

The PCRs were performed in a volume of 10 μL as described above in the microsatellite marker typing section. Admixture estimation was performed with admix [130].

2.9 Association analysis

Hardy-Weinberg equilibrium was tested using Genepop on the web (http://wbiomed.curtin.edu.au/genepop/). Association analysis for the four intragenic STRs was performed in a three-step procedure. First, allele frequencies were estimated in both cases and parents from the control group, using Genepop on the web (http://wbiomed.curtin.edu.au/genepop/). Then, a chi-square test was applied to determine whether a statistical difference of allele frequencies was observed between groups. Finally, evaluation of the haplotype frequencies was examined in Arlequin 2000 [131]. To do this analysis, haplotype definition was performed in the children from the control trios; then, most likely haplotypes in the cases were compared with those present in the children from the control trios.

Exon	Oligo F (5'-3')	Oligo R (5'-3')
1	GCGCGGCTGGCGCCGCTGCGCGCA	GCGGCGCAGAGAGGCTGTAC
2	ATGTTGCTATCACCATTTAAGGG	AGATTGGCAGCGCAGGCGGCATG
3	ACATGTCACTTTTGCTTCCCT	AGGCCATGCTCCATGCAGACTGC
4	ACAAGCTTTTAAAGAGTTTCTTGT	AGGCAATGTGTTAGTACACA
5	ACATGTCTTAAGGAGTACATTT	TCTCTAATTTCCTGGCAAACAGTG
6	AGAGATTGTTTACTGTGGAAACA	GAGTGATGCTATTTTTAGATCCT
7	TGCCTTTCCACACTGACAGGTACT	TCTGTTCTTCATTAGCATTAGAGA
8	TGATAGTCATAACTGTGTGTAAG	ACTGTCTCATTAGCGTCTATCTT
9	GGGTGAAATTTGCAGTCAGT	AATATAATCCCAGCCCATGTGCA
10	ATTGCCAAATGCAACCTMTGTC	TTGGAGGAATGAGTAGGGGATT
11	ACAGGGAACATAAACTCTGATCC	CAACACACCAGGCACCTTCAGA
12	GTTTGGGAATGCGTGTTTT	AGAATTAGAAAATGAAGGTAGACA

Table 3.2. Primer sequences used to amplify the *PARK2* exons [70].

MARKER	LABEL	ASR	LANE
D6S253	HEX	271-291	1
D6S255	FAM	163-175	1
D6S386	HEX	168-198	1
D6S1277	FAM	292-318	1
D6S1550	HEX	122-136	1
D6S1599	TET	131-155	1
D6S1581	FAM	215-229	1
D6S1579	HEX	147-163	2
D6S437	FAM	129-163	2
D6S980	FAM	292-324	2
D6S305	FAM	200-230	2

Table 3.3. Markers used to evaluate the PARK2 locus.
ASR= Average Size Range. LANE= Lane to be loaded onto the ABI-377.

3. Results

3.1 Haplotype analysis in the pedigrees

A schematic view of the markers evaluated in *PARK2* locus is shown in Figure 3.1.
Haplotypes for the pedigrees tested are shown in Figures 3.2-3.7.

Haplotype analysis revealed that affected individuals in PJF1 (Figure 3.2) were
homozygous for the same mutant chromosomes. The homozygosity region in this
family extends about 16 cM in a physical length of 6.9 Mb, from D6S1550 up to the
most telomeric marker evaluated, which was D6S386 (Figures 3.1& 3.2).

In contrast to PJF1, affected individuals in PJF3 (Figure 3.3) showed two mutant
chromosomes. One of these shared most of the haplotype found in PJF1. They shared
alleles from markers D6S1581 up to D6S1277. The second mutant chromosome was
identical to the first, from marker D6S1550 up to D6S1599, with the only exception
being marker D6S980 where there is an allele 11 instead of an allele 10. This interval
matches the region of homozygosity in this family, which contains the *PARK2* gene
(Figure 3.1). Extension of PJF3 revealed that the three sporadic-like cases from Peque
belong to other branches of the pedigree. In all, 14 more affected individuals were
identified in this extended pedigree (extended PJF3) distributed in six additional
nuclear families (Figure 3.4). One of these nuclear families showed a dominant-like
mode of inheritance (V:9, VI:1 and VI:2). Interestingly, it was found that the three
sporadic-like cases that were joined onto the extended pedigree (V:1, V:12, VI:2) were
carriers of the G736A mutation and, in addition, they all shared a common second

haplotype (Table 3.5). SSCPs revealed an altered conformation in exon 2 for at least one of these samples (Figure 3.8). Sequencing uncovered a substitution of G to A in position 212 of the *PARK2* cDNA. This change involves the 37th codon, where a Proline (Pro=P) is conserved since the codon CC<u>G</u> changes to a synonymous one, CC<u>A</u> encoding Pro again. The A allele was found in 21.15% of the controls from Peque, suggesting that it is not a causal disease variation.
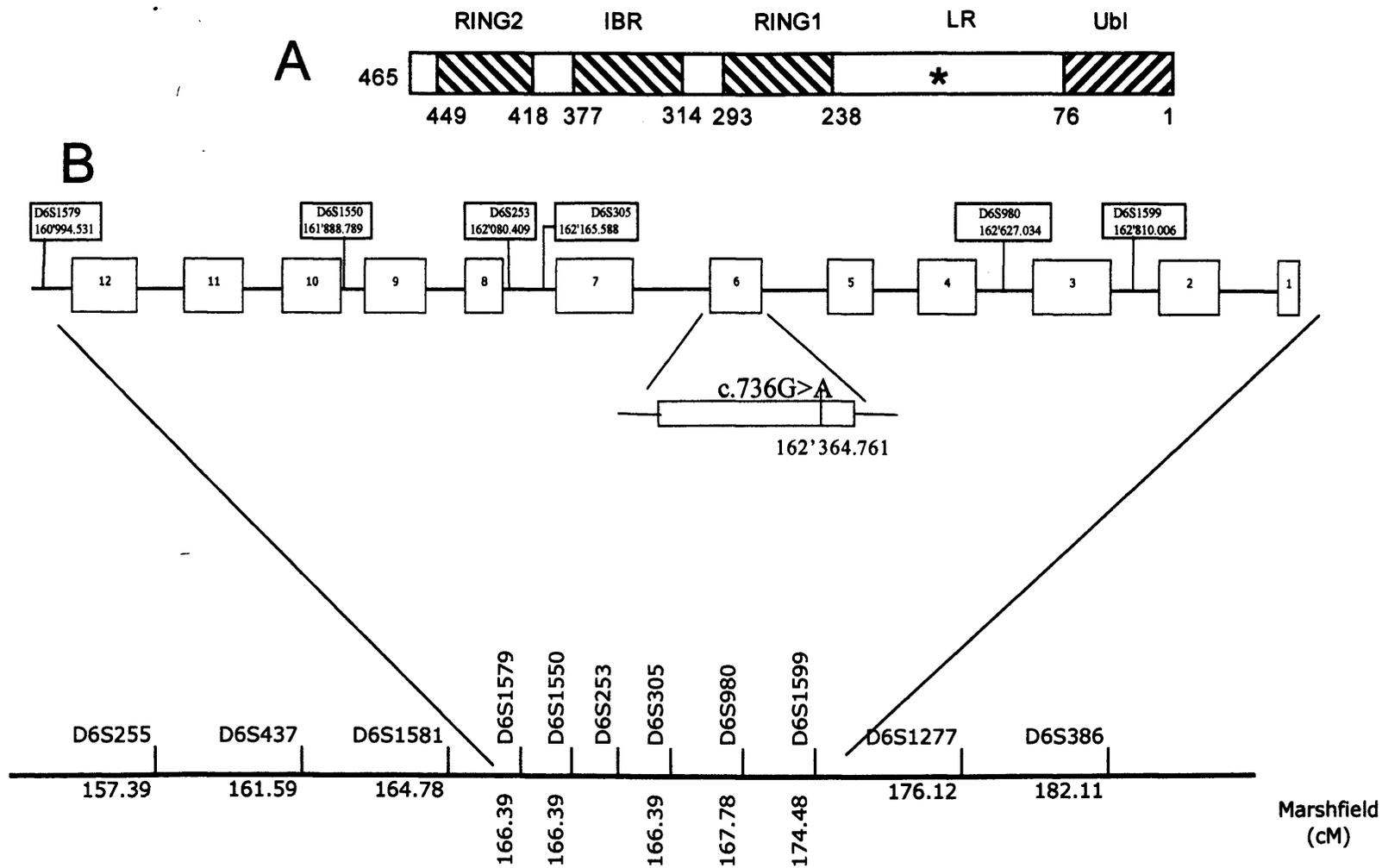
Figure 3.1. *PARK2* locus diagram. A, the structure of the protein Parkin. Above are shown its three parts (Ubl -Ubiquitin like-, LR -Linking region- and RING box). IBR -In between RINGs-, *= approximate location of Cys212Tyr. Underneath the amino acid numbering is shown. B, The map of the eleven microsatellites evaluated and their position respective to *PARK2* exons. The mutation c.736G>A, which results in the amino acid substitution Cys212Tyr, is indicated.

46

Figure 3.2 — PJF1 family (pedigree with haplotypes)

Markers and map positions (cM):

Marker	Position
D6S255	0.00
D6S437	4.20
D6S1581	7.40
D6S1579	9.00
D6S1550	9.90
D6S253	10.10
D6S305	10.20
C212Y	10.30
D6S980	10.40
D6S1599	17.09
D6S1277	18.73
D6S386	24.72

Figure 3.2. PJF1 family. Below each symbol is age(age at onset) and haplotypes for markers shown on the left. In C212Y, a number 2 indicates mutant allele. Boxed is the homozygosity region shared among affected individuals.

Figure 3.3 — PJF3 family (pedigree with haplotypes)

Markers and map positions (cM):

Marker	Position
D6S255	0.00
D6S437	4.20
D6S1581	7.40
D6S1579	9.00
D6S1550	9.90
D6S253	10.10
D6S305	10.20
C212Y	10.30
D6S980	10.40
D6S1599	17.09
D6S1277	18.73
D6S386	24.72

Figure 3.3. PJF3 family. Below each symbol is age(age at onset) and haplotypes for markers shown on the left. In C212Y, a number 2 indicates mutant allele. Boxed is the homozygosity region shared among affected individuals.

47

Figure 3.4. Extended PJF3. This large kindred includes PJF3 (framed) and three sporadic-like cases who joined the pedigree in different branches (V:1, V:12 and VI:2). These three cases are heterozygous for c.736 G>A in exon 6 and they also share a second haplotype that was associated with an exon 3 duplication

Further analysis by SSCP did not reveal new sequence variations. However, an exon 3 duplication [109] was found in individuals V:1 and VI:2 (Suzanne Lesage, personal communication), which was assumed to be present in individual V:12 as well.

Extension of the pedigree of the sporadic case reported in Pineda-Trujillo and colleagues (2001) established that there were no more affected individuals in this family. Haplotype and mutation carrier states were in agreement with the observed phenotypes. This family was named PJF4 (Figure 3.6). PJF2 (Figure 3.5) and PJF4 have the same *PARK2* mutation in exon 3, which is an insertion of two base pairs (321-322 insGT) [83]. Haplotype analysis revealed that the maximum homozygosity region extends from D6S255 to D6S1277, with the exception of locus D6S980 where individual II:3 (from PJF2) appears as heterozygous 4/5. Segregation of the mutation with the phenotype was perfect in these four families, with the only exception of individual II:3 in PJF3, who is a non-penetrant individual for mutation c.736 G>A.

3.2 Suggestion of a recurrent mutation involving c.736G>A mutation

Evaluation of haplotypes in the Spanish family carrying the c. 736 G>A mutation [117] showed that the two populations are carrying completely different chromosomes (Figure 3.7).

3.3 Sporadic cases

Allele frequencies for the four tested markers are shown in Figures 3.9-3.12 and haplotype frequencies are shown in Tables 3.6-3.7. Association analysis revealed that

markers D6S1550 and D6S980 are associated with the disease (Table 3.8). The associated alleles for D6S1550 were 126 and 134, which in Figure 3.9 correspond to alleles 3 and 7, respectively. Allele 294 of marker D6S980 is overrepresented in cases, in contrast to controls. In Figure 3.12, this allele corresponds to allele 9. Haplotype analysis showed that these alleles are present in the four most frequent haplotypes present in the cases (Table 3.7). Haplotypes in the 17 controls were unique (Table 3.6).

For three out of four markers there was no significant deviation from the frequencies expected under Hardy-Weinberg equilibrium. D6S1550 showed a significant deviation from Hardy-Weinberg equilibrium (Table 3.9), which might indicate that at least one of the assumptions is violated, e.g. random mating. None of the four markers showed significant intermarker LD either in the parents or in cases (Table 3.10), although the pair D6S1550-D6S980 shows a marginal P-value.

			Allele frequency for the shorter allele		
Marker name	Dye	Chromosome	Peque	European	Amerindian
MID93	TET	22	0.706	0.209	0.965
MID472	HEX	6	0.542	0.815	0.425
MID1326	FAM	X	0.333	0.701	0.015
MID1327	FAM	X	0.389	0.544	0.037
MID1338	HEX	X	0.083	0.704	0.398
MID1705	HEX	X	0.083	0.814	0.146

Table 3.4. MID markers used to evaluate the Amerindian versus European component in Peque population and frequencies for their shorter allele.
Frequencies in Europeans and Amerindians were taken from htt://research.marshfieldclinic.org. ˡ
ˈPeque= 0.72 Amerindian.
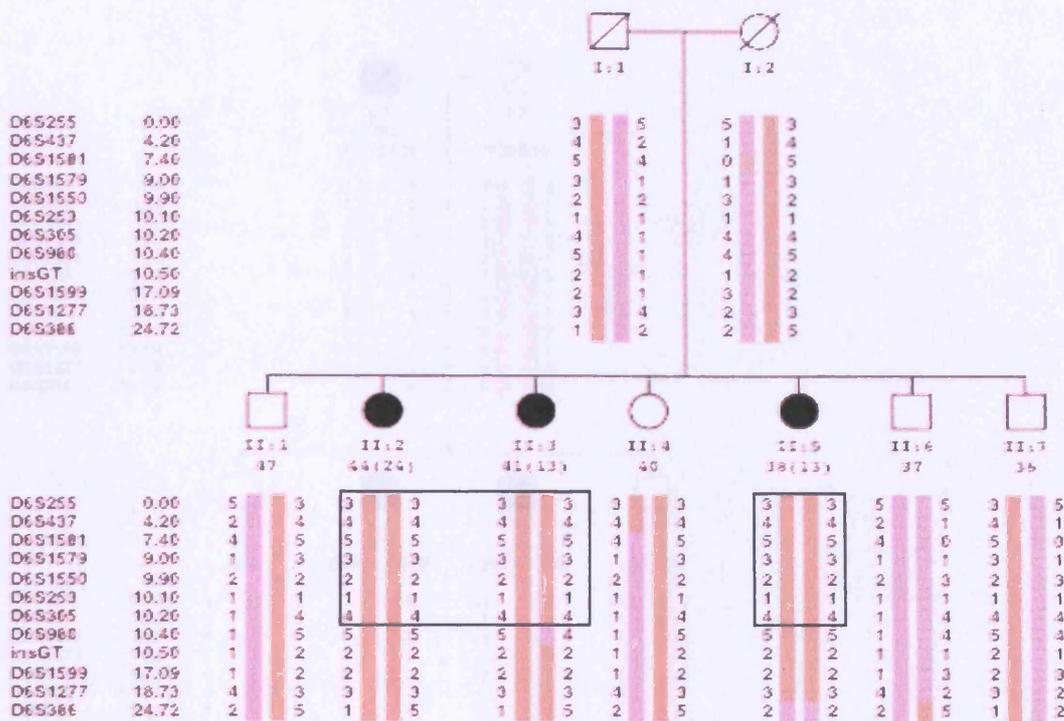
Figure 3.5. PJF2 family. Below each symbol is age(age at onset) and haplotypes for markers shown on the left. In insGT, a number 2 indicates mutant allele. Boxed is the homozygosity region shared among affected individuals.
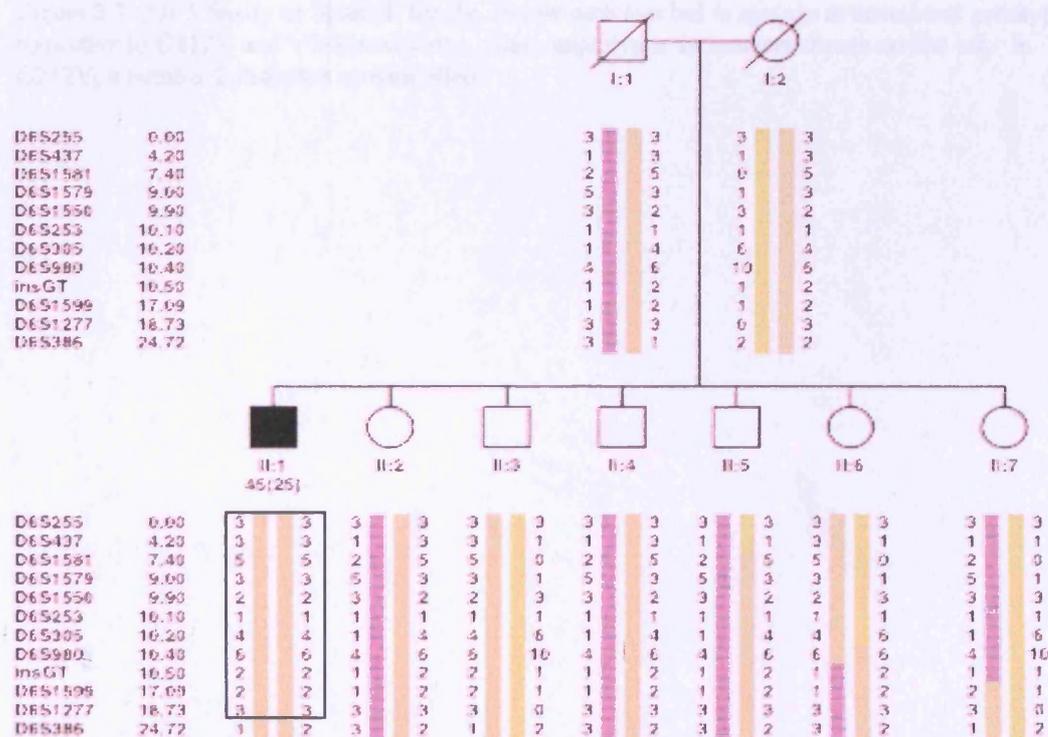
Figure 3.6. PJF4 family. Below each symbol is age(age at onset) and haplotypes for markers shown on the left. In insGT, a number 2 indicates mutant allele. Boxed is the homozygosity region in the affected individual.

51
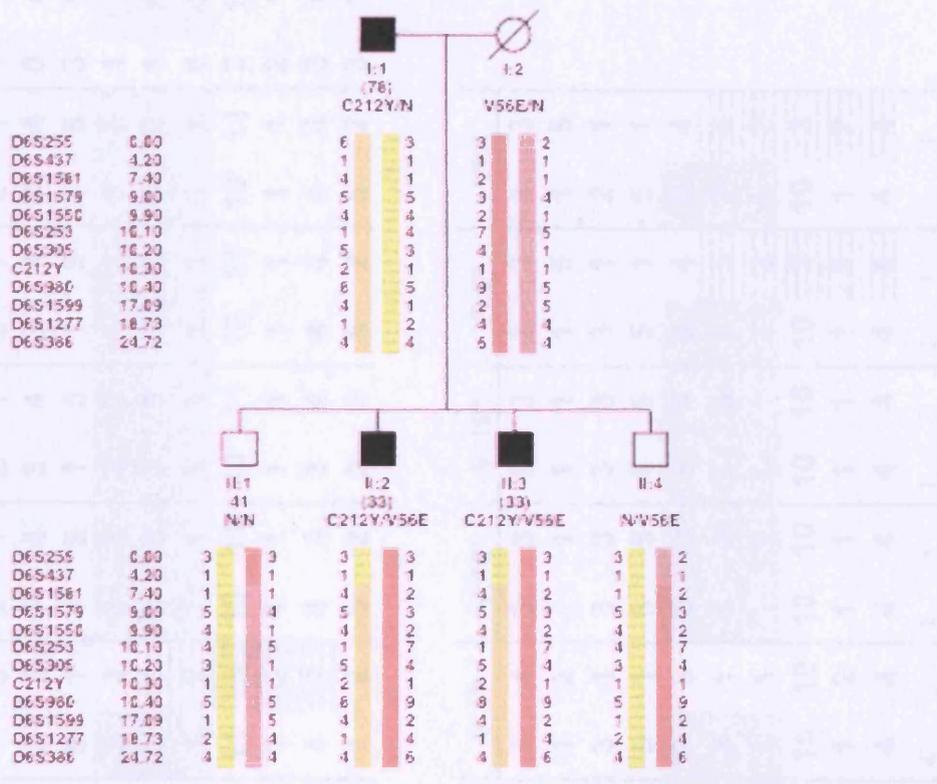
Figure 3.7. PJF5 family or Spanish family. Below each symbol is age(age at onset) and genotypes respective to C212Y and V56E mutations. Also, haplotypes for markers shown on the left. In C212Y, a number 2 indicates mutant allele.

marker\ID	IV:3		IV:8		IV:9		IV:10?		IV:11		V:1*		V:2*		V:3*		V:4		V:5*		V:6		V:11	
D6S255	3	3	3	3	3	3	3	3	3	3	3	3	4	3	4	3	4	3	3	3	3	3	3	3
D6S437	1	1	1	8	1	2	1	2	1	4	1	3	2	1	2	1	2	1	2	1	1	1	1	1
D6S1581	3	4	3	2	3	5	3	5	2	3	3	5	3	4	3	4	3	4	3	4	0	4	2	3
D6S1579	1	5	5	1	5	1	1	5	1	5	5	1	1	5	1	5	1	5	1	5	3	5	5	1
D6S1550	3	3	3	3	3	3	3	3	3	4	3	3	3	3	3	3	3	3	3	3	1	3	3	3
D6S253	6	3	3	4	3	2	3	2	3	1	3	5	3	3	3	3	3	3	3	3	1	3	3	4
D6S305	1	1	1	1	1	3	1	3	1	6	1	2	1	1	1	1	1	1	1	1	8	1	1	1
D6S980	4	11	10	4	10	7	11	7	10	12	11	3	10	11	10	11	10	11	10	11	3	11	10	4
D6S1599	1	1	1	2	1	2	1	2	1	1	1	2	1	1	1	1	1	1	1	1	2	1	1	1
D6S1277	4	4	4	4	4	3	3	4	1	4	4	4	4	4	3	4	4	4	4	4	3	4	4	4
D6S386	1	2	1	5	2	5	2	5	4	5	1	2	5	2	3	2	5	2	5	2	3	2	2	1

marker\ID	V:12*		V:13		V:14*		V:15		V:16*		V:17		V:18*		V:19*		VI:2*		VI:4	
D6S255	3	3	3	3	3	3	3	3	3	3	4	4	3	3	3	3	3	3	3	3
D6S437	1	3	1	1	1	1	8	1	1	1	1	2	1	1	1	1	1	5	1	5
D6S1581	3	5	5	3	3	3	2	3	3	3	3	1	3	3	3	3	3	1	2	1
D6S1579	5	1	1	5	5	5	1	5	5	5	5	1	5	5	5	5	5	1	5	1
D6S1550	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	3	4
D6S253	3	5	2	3	3	3	4	3	3	3	3	1	3	3	3	3	3	5	3	5
D6S305	1	2	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2
D6S980	10	3	7	10	10	10	4	10	10	10	10	10	10	10	10	10	10	3	10	3
D6S1599	1	2	2	1	1	1	2	1	1	1	1	2	1	1	1	1	1	2	1	2
D6S1277	4	4	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
D6S386	1	5	5	1	1	5	5	5	1	2	4	5	1	5	1	5	2	1	2	1

Table 3.5. Haplotypes for extended JPF3. ID's correspond to figure 4..
*= affected individuals; ? phenotype unknown.

| | CONTROLS | | | | |
Haplotype	Freq.	D6S1550	D6S253	D6S305	D6S980
1	1	128	271	230	312
2	1	128	271	218	316
3	1	128	271	222	294
4	1	128	283	218	290
5	1	128	271	224	298
6	1	120	271	218	290
7	1	128	271	230	296
8	1	136	271	218	292
9	1	136	271	218	296
10	1	128	271	228	296
11	1	128	287	228	296
12	1	136	271	204	284
13	1	136	281	218	296
14	1	128	271	222	298
15	1	128	287	204	296
16	1	130	271	234	292
17	1	128	271	204	296
18	1	128	271	218	288
19	1	126	271	226	288
20	1	130	287	230	296
21	1	126	271	224	286
22	1	130	271	228	290
23	1	126	287	204	276
24	1	130	271	230	290
25	1	126	287	224	284
26	1	130	271	218	288
27	1	126	287	222	296
28	1	134	271	204	296
29	1	128	271	204	292
30	1	128	271	234	288
31	1	134	271	218	290
32	1	126	285	218	298
33	1	134	281	224	290
34	1	126	271	222	302

Table 3.6. Haplotypes in 17 controls for D6S1550-D6S253-D6S305-D6S980

Haplotype	Freq.	CASES			
		D6S1550	D6S253	D6S305	D6S980
1	1	120	271	204	?
2	1	120	271	238	286
3	1	120	283	222	290
4	1	124	271	228	280
5	1	126	271	204	288
6	1	126	271	218	280
7*	4	126	271	218	290
8	1	126	271	222	290
9	1	126	271	224	282
10*	3	126	271	224	294
11	1	126	271	226	294
12	1	126	271	230	286
13	1	126	273	218	?
14*	2	126	279	218	312
15	1	126	283	222	292
16	1	126	283	230	312
17	1	126	285	204	294
18	1	126	287	204	288
19	1	126	287	204	298
20	1	126	287	228	294
21	1	126	289	224	298
22	1	128	271	204	276
23	1	128	271	204	298
24	1	128	271	234	290
25	1	130	271	218	294
26	1	130	271	220	294
27	1	130	279	234	286
28	1	130	281	222	294
29	1	130	283	228	294
30	1	130	?	204	298
31	1	132	271	230	304
32	1	134	271	220	280
33	1	134	271	234	298
34	1	134	279	218	280
35*	2	134	281	204	316
36	1	134	285	204	290
37	1	134	?	228	312

Table 3.7. Haplotypes in 22 cases for D6S1550-D6S253-D6S305-D6S980 Individual 23rd was taken out of the analysis because of missing genotypes at three loci. The * indicates four haplotypes that represent 23.4% of the disease chromosome.
Haplotype association between cases and controls resulted in a P-value=0.0075
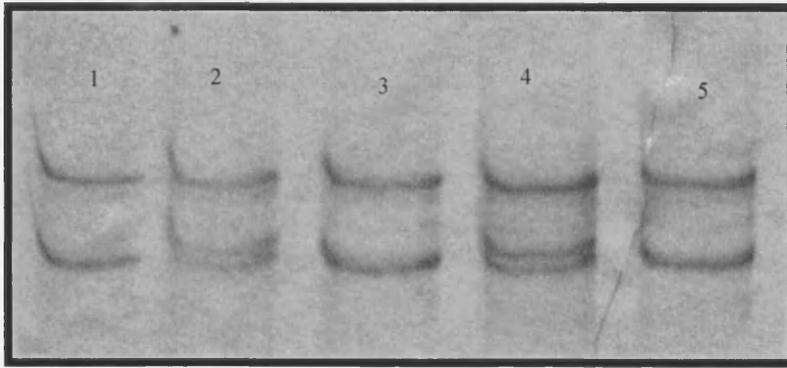
Figure 3.8. SSCP for c.212 G>A polymorphism in exon 2.
Lane 1= V:2, Lane 2= VI:2, Lane 3= Control-1, Lane 4= VI:4, Lane 5= Control-2. VI:2 and
VI:4 are carriers.

D6S1550	1	2	3	4	5	6	7	8
cases	5.56	1.85	44.44	12.96	14.81	3.70	16.67	0.00
controls	4.88	2.44	31.71	26.83	19.51	0.00	6.10	8.54
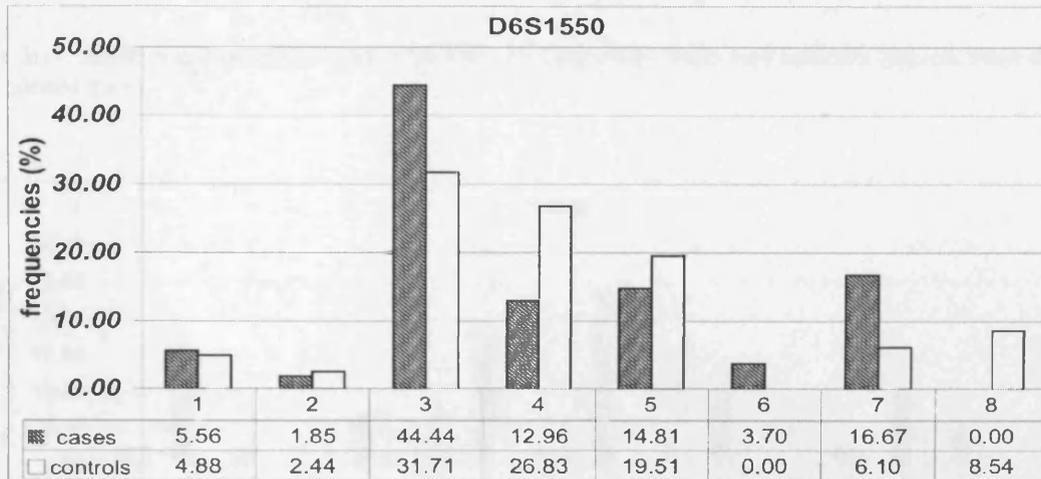
Figure 3.9. Allele frequencies for marker D6S1550 in 23 sporadic cases and controls (parents from the
control trios).

D6S253	1	2	3	4	5	6	7	8	9	10	11
cases	52.63	2.63	0.00	10.53	7.89	10.53	5.26	7.89	2.63	0.00	0.00
controls	73.08	0.00	1.28	5.13	2.56	2.56	2.56	10.26	0.00	1.28	1.28
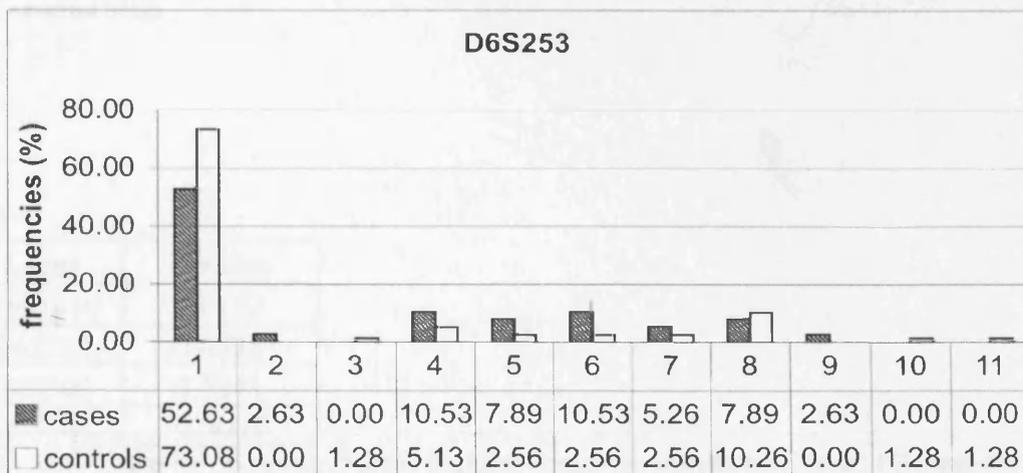
Figure 3.10. Allele frequencies for marker D6S253 in 23 sporadic cases and controls (parents from the
control trios).

D6S305

frequencies (%)

	1	2	3	4	5	6	7	8	9	10	11	12
cases	22.92	0.00	2.08	20.83	4.17	10.42	10.42	4.00	8.33	8.33	6.25	2.08
controls	20.27	1.35	0.00	32.43	1.35	10.81	6.76	4.05	10.81	9.46	2.70	0.00
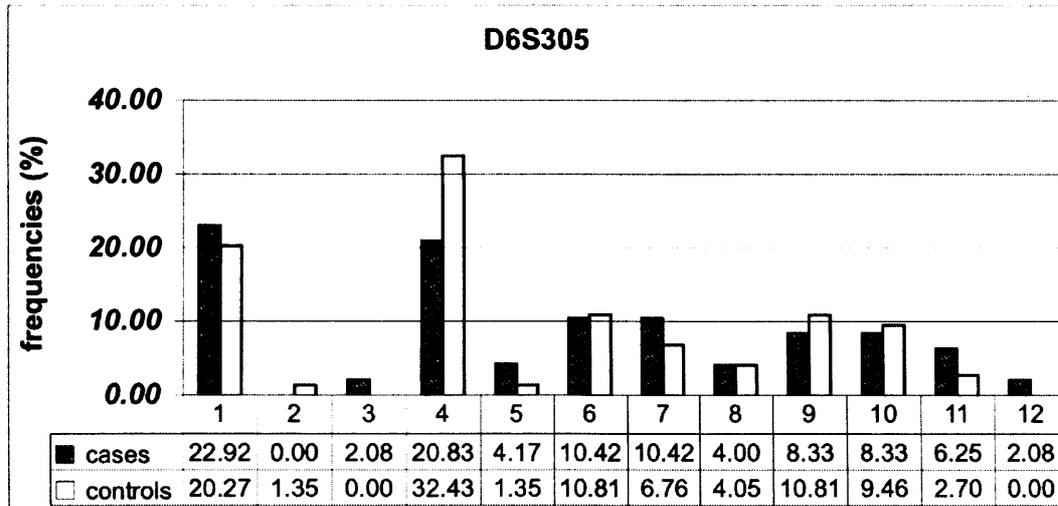
Figure 3.11. Allele frequencies for marker D6S305 in 23 sporadic cases and controls (parents from the control trios).

D6S980

frequencies (%)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
cases	2.38	9.52	2.38	0.00	7.14	4.76	19.05	2.38	23.81	0.00	11.90	0.00	2.38	9.52	4.76	0.00	0.00
controls	2.44	0.00	0.00	4.88	7.32	7.32	28.05	4.88	6.10	15.85	8.54	4.88	1.22	3.66	1.22	1.22	2.44
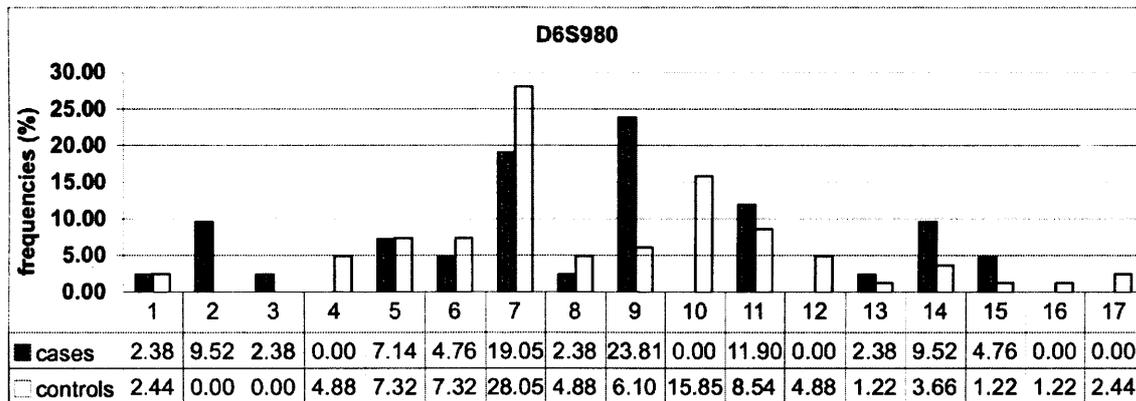
Figure 3.12. Allele frequencies for marker D6S980 in 23 sporadic cases and controls (parents from the control trios).

Locus	P-value
D6S1550	0.0153
D6S253	0.0638
D6S305	0.7841
D6S980	0.0003

Table 3.8. P values for chi-square obtained in global association analysis between 23 PD cases and controls (parents from control trios).

LOCUS	P-value
D6S1550	0.0118
D6S253	1
D6S305	0.6089
D6S980	0.113

Table 3.9. P values for the goodness of fit for Hardy-Weinberg Equilibrium for markers D6S1550, D6S253, D6S305 and D6S980 in parents from control trios.

	Controls (parents)	Cases
Locus pair	P-Value	P-Value
D6S1550&D6S253	0.28	0.473
D6S1550&D6S305	0.932	0.18
D6S253&D6S305	0.229	0.815
D6S1550&D6S980	0.238	0.061
D6S253&D6S980	0.953	0.688
D6S305&D6S980	0.828	0.918

Table 3.10. P values for Linkage Disequilibrium in both cases and controls (Parents).

4. Discussion

The purpose of this chapter was to examine whether those affected individuals, who were carriers of the c.736G>A mutation in *PARK2* were compound heterozygous or were phenotypically affected while being heterozygous for this mutation alone. In addition, polymorphic markers were typed in the vicinity of *PARK2* to evaluate whether the Colombian families carrying the c.736G>A mutation were carrying the same chromosome (haplotype background) as a Spanish family where the same mutation was reported. Moreover, a set of 23 sporadic patients were evaluated for four of these markers as part of a search for association of *PARK2* with the disease in this group of patients.

Haplotype analysis in the families studied revealed a large shared haplotype in family PJF1, which suggests that both parents might share a close common ancestor.

The observation in PJF2 of heterozygosity in marker D6S980 in individual II:3 and given that she is affected and homozygous for the mutation 321-322insGT, suggests that this heterozygosity might be due to a mutation in the marker rather than to a recombination event, since a recombination would have taken away the mutation, leaving the individual as heterozygous. The observation in family PJF3 of one of the mutant chromosomes showing a difference only in marker D6S980 supports the idea of a high mutation rate at this locus.

Admixture analysis of Peque population revealed that 72% of its genetic component is Amerindian (Table 3.4). The Spanish family presents a chromosome carrying the mutation c.736G>A, which is completely different from those in the two Colombian families. This observation, in addition to the admixture results, points towards that a recurrent mutation has occurred independently in the two populations. However, this

59

interpretation needs to be taken with caution since there it is not known how old the Spanish chromosome is and because recombination may have happened in the vicinity of the mutation taking away those alleles that are in linkage disequilibrium in the Colombian mutant chromosomes. Therefore, more closely linked markers (i.e. SNPs) should be typed in c.736G>A vicinity and if haplotype sharing is still not observed, then the hypothesis of a recurrent mutation would be supported.

The finding of a second mutation in the extended PJF3 is remarkable. This is the third *PARK2* mutation in Antioquia. Nonetheless, its frequency among the sporadic cases does not seem to be significant as judged by the haplotypic data.

The association analysis points to other *PARK2* mutations that went undetected, since strong association was obtained for two of the four STRs tested. Moreover, when the haplotypes where considered in the association analysis a highly significant P-value was obtained (Tables 3.6 & 3.7).

Two of the three mutations described are in one of the functional domains of parkin. These are 321-322 insGT and the exon 3-duplication, which are associated with an earlier age at onset compared with the age at onset for individuals homozygous for c. 736 G>A, which lies in the linking region of the protein, suggesting that these two are more severe than c. 736 G>A. The first two mutations affect the ubiquitin-like domain in parkin, which acts as a ligase E3 in the ubiquitination pathway before proteosomal degradation [63].

In conclusion, a third *PARK2* mutation has been found in Antioquia, which is a duplication of exon 3. This mutation seems to account for none of the sporadic cases as judged by the haplotype analysis. In contrast, in the soporadic cases four out of 37

haplotypes account for 23.4% of the patient chromosomes (Table 3.7), which will eventually be associated to *PARK2* mutations. If this were the case, it would constitute another founder effect in this isolate population, where evidence of linkage disequilibrium (LD) due to admixture and founder effects have already been documented [32, 35, 36]. Antioquia appears optimal for gene mapping by LD methods [33].

In addition, the finding of exon 3-duplication in apparently heterozygous affected individuals adds evidence to refute previous reports claiming that carrier condition would be enough for developing symptoms [65, 83, 110, 113, 118, 128, 129].

Finally, in order to establish whether a recurrent mutation has been observed in the Colombian and Spanish chromosomes carrying the c.736 G>A mutation, more analyses need to be done in its vicinity, e.g. typing SNPs more tightly linked than markers D6S253 and D6S305.

CHAPTER FOUR

CHARACTERIZATION OF A

FAMILY WITH A MULTIFORM

MOVEMENT DISORDER

1. Introduction

Distinct forms of Parkinsonism have been reported as caused by mutations in genes

and loci described in the previous chapter. In addition, other genes responsible for

different well characterized disorders have also been associated with Parkinsonism as

part of their phenotype spectrum. In this introduction I will focus on three of those

genes that can lead to a similar phenotype to the one presented by a family presenting

with Multiform Movement Disorder (MMD), where Parkinsonism is the clinical

landmark. These are Huntington's disease, Pantothenate Kinase 2 and Ferritin Light

chain genes. Also, preliminary clinical, pathological and genetic analyses are

discussed.

1.1 Pantothenate Kinase 2 (*PANK2*) gene.

Hallervorden-Spatz syndrome (HSS) is a spectrum of disorders characterized by

neurodegeneration and iron accumulation in the brain with an autosomal recessive

inheritance mode [132]. The classical form of the disease originally described, starts in

childhood with dystonia, dysartria, rigidity and has a relentlessly progressive course,

ending in early death. Pathological characteristics include pigmentary degeneration of

the retina and iron deposition in the basal ganglia is evident in MRI and postmortem

examination; this is described as the "eye of the tiger" sign [132]. The atypical form of

the disease has late onset (adolescence) and its progression is slower than in the

classical form.

The term HSS has now been changed to Neurodegeneration with Brain Iron

Accumulation (NBIA) [133]. To emphasize the fact that mutations in the *PANK2* gene

are the major cause of NBIA, more recently it has been proposed that individuals with mutations in this gene be classified as Panthothenate Kinase-associated neurodegeneration (PKAN) [134] instead of presenting HSS or NBIA that might not be due to *PANK2* mutations.

The process to find *PANK2* mutations began with a homozygosity map strategy that made it possible localize the putative gene to chromosomal region 20p13 [135]. Then, more polymorphic microsatellite markers were added in the candidate region. Since no common haplotypes were identified among families and not knowing what the candidate gene could be, all 21 known or predicted genes were to be analyzed. After studying 15 of those genes, a 7-bp deletion in the coding region sequence of a gene homologous to the murine pantothenate kinase 1, was identified in the index family. Hence, this gene was named pantothenase kinase 2, *PANK2* [134]. Other missense or null mutations were found in 32 of 38 patients with classical HSS. In addition, individuals with atypical HSS also presented missense mutations.

In another study nonsense mutations were found in atypical HSS patients [133], but in contrast to the first report of mutations associated with the disorder, this second study found *PANK2* mutations in all classical HSS patients. A common (~25%) mutation has been identified in both classical and atypical HSS patients [133, 134].

PANK2 spans 35 kb and is a member of a family of eukaryotic genes consisting of a group of six exons that encode homologous core proteins, preceded by a series of alternative initiating exons, some of which encode unique amino-terminal peptides [134]. Human *PANK1*, *PANK2* and *PANK3* are strikingly similar to their murine ortologs and the difference between paralogs is not over 20%. These genes are located on human chromosomes 10q23, 20p13 and 5q35, respectively. Human *PANK4* is located on 1p36 [134].

Pantothenate kinase is an essential regulatory enzyme in the Coenzyme A (CoA) biosynthesis, catalyzing the cytosolic phosphorylation of pantothenate (vitamin B_5), N-pantothenoyl-cysteine and pantotheine [136]. The disease could result from product deficit (Phosphopantothenate) and a secondary excess of metabolites downstream of the block (Cysteine, N-pantothenoyl-cysteine and pantetheine). Pank2 mutations are predicted to result in CoA depletion and defective membrane biosynthesis in those tissues in which this is the major pantothenate kinase or in those tissues with the greatest CoA demand [134].

Intrafamilial variability has been reported including psychiatric symptoms that have resulted in misclassification [137]. Senile chorea has also been associated with NBIA, since the eye of tiger sign has been evident in these patients [138] and cortical dementia has also been added to the spectrum [139]. Although age of onset in most of NBIA patients is in the second or third decade of life [134, 140], there are also some reports of elder patients (68 years old) [141].

1.2 Huntington disease (*HD*) gene.

Huntington disease is characterized by movement disorders, psychiatric symptoms (mainly depression) and dementia with autosomal dominant inheritance. Death will occur approximately 15-20 years after the first symptoms appear. The pathological target in this disorder is the medium spiny neurons of the striatum.

The localization of the *HD* (or IT15) gene was performed in 1983 [142] but it was only ten years later that the actual gene was identified [22]. This gene has 67 exons and spans 169,41 Kb. A mutation in exon 1 was identified, consisting of an expansion of a triplet CAG, that encodes glutamine [22]. Normal alleles range from 6-35 copies of this

trinucleotide. It has been observed that alleles with ≤ 26 copies do not present meiotic instability (trend to expansion) [143, 144]. In contrast, ≥ 27 CAGs are commonly reported as instable in sperm; the range of alleles 25-35 is called intermediate because of being very prone to expansion over 36 CAGs [145, 146]. The bigger the CAG repeat, the higher the instability for the CAG repeat [145, 147, 148]. Anticipation has been observed, with onset tending to become earlier in successive generations [149]

Affected individuals have been observed with ≥36 CAG repeats and penetrance is higher when more repeats are carried by the patient. It has been reported that ≥42 repeats will always be penetrant (100% penetrant) [149] but ≤41 might or might not lead to symptoms since there have been reports of individuals older than 90, with 41 repeats who have died disease-free [150, 151].

The expansion of the polyglutamine tract in the huntingtin (htt) protein confers cytotoxic properties to htt and compromises some of its endogenous functions [152, 153]. It is likely that post-translational modifications such as phosphorylation, ubiquitination, sumoylation, palmitoylation and proteolysis, play an important role in the pathogenesis of HD [154-157].

1.3 Ferritin Light Peptide (*FTL*) gene.

Ferritin plays an important role in iron storage and in the control of intracellular iron distribution. It is composed of 24 heavy (H) and light (L) subunits. H and L are regulated by a common cytoplasmic protein, the iron regulatory protein (IRP), which binds to the iron responsive-element (IRE), located on the 5' end of both H and L chains gnes (reviewed in [158]).

Curtis and colleagues (2001) reported an unknown, dominantly inherited disorder that variably presented with symptoms of Huntington's disease (HD) or Parkinsonism and late onset in a large kindred from Cumbrian-England. Brain histopathology showed widespread reddish discoloration of the basal ganglia. The globus pallidus contained abundant iron-positive, most of which were also positive for ferritin. The imaging and histology data are similar to those of NBIA [159]. Their approach in the genetic study consisted of first ruling out known neurological disorders by routine diagnostic tests and then, performing a linkage analysis study. They found linkage to 19q13.3, where a maximum lod score (Z) of 6.38 at theta (θ) zero was obtained. Refining the region was possible by adding available family members and polymorphic microsatellite markers in the initial candidate region. Haplotype analysis showed a 3.5 cM region of maximum disequilibrium between markers D19S596 and D19S866 corresponding to a 2 Mb physical distance. From this interval they chose *FTL* as a candidate on the basis of the suspected involvement of iron in neurodegenerative disorders [160].

FTL has 4 exons and spans 1.56 Kb. They identified a variant in exon 4 by using SSCP. Sequence analysis revealed an insertion of an Adenine after position 460, in all affected individuals in a large pedigree. This was not found in 300 Cumbrian controls. Five additional patients carrying this mutation presented the disease at 36-58 years. All these individuals had dominant family histories of various neurological disorders. This mutation is predicted to alter 22 C-terminal residues of the protein and extend the chain for four additional residues. Of the five domains of the protein, this mutation affects the two last (D-E) [159].

Another mutation identified in this gene was a 2-bp insertion in exon 4, which was identified in a family with tremor, cerebellar ataxia, parkinsonism and pyramidal signs. The pathology revealed accumulation of both light and heavy ferritin

Polypeptide [161]. The third reported *FTL* mutation associated with ferritinopathy is an insertion of C at nt 646-647 in exon 4 [162]. Mutations in *FTL* have also been associated with hyperferritinemia-cataract syndrome [163-168].

A mutation in the IRE of the heavy-ferritin chain has been reported in a Japanese family affected with a dominant inherited iron overload in the liver and spleen. The mutation c.49A>T significantly increases the binding affinity of IRP [158]. The heavy ferritin chain gene lies on chromosome 11q13.

1.4 Preliminary work.

In this chapter the characterization of a family presenting with various clinically related movement disorders is presented. This family originates in Antioquia-Colombia and was ascertained in NEUROCIENCIAS, Universidad de Antioquia. This family presents with a neurodegenerative disorder that at times seems typical of Parkinson disease, at times more like Huntington and at times more like Hallervorden-Spatz syndrome. Affected individuals have been defined as having any of these diseases. Clinical details for five out of six affected individuals are summarized in Table 4.1.

It is remarkable that there is not clear diagnosis and that there is a lack of a unique and clear clinical entity segregating in the family shown in Figure 4.1. Therefore, we have named this disorder as Multiform Movement Disorder (MMD).

ID	First symptom Age at onset	Evolution	Response to medication	Diagnostic impression
II:4	Psychotic episode, Involuntary movements of upper limbs, mouth and tongue. Onset at 77 years old	Resting distal tremor, marked tremor of both hands. Parkinsonism.	No	Late Parkinsonism, Lewy Body disease
III:1	Progressive recent memory loss. Onset at 53 years old	Parkinsonism, Seizures	YES	Huntington, Lewy Body disease, Hallervorden-spatz syndrome, Parkinsonism
III:6	Recent memory loss and pain in one foot. Onset at 43 years old	Behavioural changes, disphagia, facial tics, involuntary movements of mouth, eyelids.	?	Huntington, Dentato-rubro-palydolysial, Parkinsonism, Hallervorden-spatz
III:7	Diskynesia, recent memory loss. Onset at 53 years old	Irritability, insomnia. Family described he had tremor when drawing or writing	?	Huntington, Halllervorden-spatz syndrome, Lewy Body disease
III:8	Tremor in right hand three days after occipital trauma. Onset at 52 years old	Migraine, generalised stiffnes, marked resting tremor of upper right limb	YES	Parkinson disease of focal origin, secondary diskinesias

Table 4.1. Summary of clinical features of affected individuals in family with Multiform Movement Disorder (MMD). Individuals are identified as in Figure 4.1.

Interestingly, brain pathology in one of the diceased patients (III:6) showed massive iron accumulation in various part of her brain, adding ferritinopathy as a likely diagnosis.

As part of a candidate genes approach, I have tested for linkage to several genes related with the likely diagnoses shown in Table 4.1 in addition to ferritnopathy. Thus, markers tightly linked to *PARK1, PARK2, PARK8, PANK2, HD* and FTL genes/loci were tested in this family (Figure 4.1). This analysis was performed typing polymorphic genetic markers in the parents and, initially, in the affected children. When necessary unaffected children were also typed.

Analyses were performed taking into consideration the reported inheritance pattern for each gene. Thus, *PARK1, PARK8, HD* and FTL were analyzed with a dominant inheritance mode and high penetrance (0.985). The two remaining genes (*PARK2* and *PANK2*) were analyzed using an autosomal recessive inheritance mode and the same penetrance as above.

Marker tested in *PARK1* (D4S1572) was not informative for linkage since the affected father II:4 was homozygous (Figure 4.1). The remaining markers provided information for exclusion. In marker D6S305, affected individuals presented different genotypes, suggesting that linkage to *PARK2* locus was not present. For example individuals III:1 and III:7 did not share any allele at this marker locus.

Pedigree genotype data (parents II:4 and II:5):

	Marker	II:4		II:5	
PARK1	D4S1572	2	2	1	3
PARK2	D6S305	1	2	2	3
PARK8	D12S85	1	3	1	2
	D12S345	1	3	2	3
PANK2	D20S889	2	3	1	1
HD	D4S412	1	3	2	4
FTL	D19S902	1	3	2	2
	D19S571	1	2	1	2

Pedigree genotype data (generation III):

	Marker	III:1	III:2	III:3	III:4	III:5	III:6	III:7	III:8
PARK1	D4S1572	2 3	2 3		0 0	0 0	2 1	2 1	2 1
PARK2	D6S305	1 3	0 0		0 0	0 0	1 2	2 2	1 2
PARK8	D12S85	1 1	3 1		2 3	1 2	1 3	1 2	1 2
	D12S345	1 3	1 3		1 2	1 2	2 3	2 3	2 3
PANK2	D20S889	0 0	0 0		0 0	0 0	1 2	1 3	1 2
HD	D4S412	1 4	3 2		0 0	0 0	0 0	1 2	3 2
FTL	D19S902	2 3	1 2		1 2	1 2	1 2	1 2	1 2
	D19S571	1 1	1 2		0 0	1 2	1 1	1 1	2 2

Figure 4.1. Candidate genes/loci tested in MMD family.
All markers were picked from Linkage mapping set v.2 (Perkin-Elmer), with exception of D6S305.
We missed samples of I:1, I:2, and II:3 individuals. Very in the end of my thesis a sample of individual III:3 became available.

PARK8 was evaluated using markers D12S85 and D12S345. In D12S85 all affected children received an allele 1 from II:4, but unaffected individual III:5 also received allele 1. The second marker in *PARK8* showed that all affected children, with exception of III:1, received allele 3 from II:4. Alternatively, III:1 received allele 1 in D12S345, reinforcing the suggested exclusion in D12S85.

For *PANK2* marker D20S889 was evaluated. Available data suggested exclusion since affected individuals received different alleles from their affected father (II:4). Finally, for the Huntington disease gene, marker D4S412 was tested. Alleles at this marker locus were randomly transmitted from II:4 to affected individuals. Thus, individual III:1 received allele 1 from II:4, whereas III:8 received allele 3.

For FTL marker D19S902 presented random segregation of alleles, since individual II:4 transmitted allele 3 to III:1 and allele 2 to the rest of his children (affected and unaffected). Marker D19S571 presented a particular situation since both the parents were heterozygous for the same alleles (i.e. 1/2), leading to inability to identify the origin of the parental alleles in a heterozygous child. Thus, unaffected individuals III:2 and III:5 are heterozygous 1/2 same as their parents. Even more, affected individual III:1 received allele 1 from II:4, and affected individual III:8 received allele 2. These results indicate that neither linkage nor segregating haplotype was observed (Figure 4.1).

1.5 Aims.

A family presenting with an unclear phenotype, at times resembling Parkinson, at times Huntington and at times NBIA is studied here. Based on this information, candidate genes/loci evaluated, were *PARK1, PARK2, PARK8, PANK2* and *HD*. Having obtained brain pathology results indicating massive iron deposits consistent with ferritinopathy, the *FTL* gene was evaluated by means of two polymorphic markers. No evidence of linkage was obtained.

The focus of this chapter was to perform a genome wide analysis in this family as a strategy for looking for the actual mutated gene causing the disorder.

2. Materials and Methods

2.1 Power simulation

Power simulations were performed assuming automosomal dominant inheritance and high penetrance (0.985). A marker locus with four equally frequent alleles was simulated in 100 replicas, using SLINK [55]. The disease frequency was assumed to be rare (0.0001). The actual simulated pedigree is shown in Figure 4.1. Individuals simulated as available were those with a sample code.

2.2 Genome wide scan

deCODE markers were typed across the genome in all available individuals at an average density of 6 cM . See laboratory methods in chapter two.

2.3 Microsatellite marker typing

Based on the multipoint lod scores identified, more markers were typed in two of the regions identified (Table 4.2). For chromosome 12, markers D12S1601, D12S1703, D12S2195, D12S2206 and D122S337 were added into the candidate region. Marker D12S1294 was typed again since it had presented a high rate of failure in the data provided by deCODE.

For chromosome 14, markers D14S1433, D14S53, D14S983, D14S1020, D14S1008 and D14S1000 were added into the candidate region. In this region, an additional unaffected individual (III:3) was available for analysis.

Marker typing was performed as indicated in the section on microsatellite marker typing in chapter two.

PCR conditions included a denaturing step of 15 min, followed by 94° 30 sec, 55° 15 sec, 72° 30 sec for 30 cycles. A final extension step of 10 min was added.

CHROMOSOME	MARKER	LABEL	SEQUENCE (5' to 3')	ASR
12	D12S1601	FAM	AGCTACTTCAGAAAACTATTTGACA GCAGGAACTTAACTCTTGCTT	193-203
	D12S1703	TET	CTAGCCTGGGAAACATAGTAAGACC CTTTCAGTGAGCGGAGACG	205-243
	D12S2195	HEX	GGGTTGGGAACCCTCAAAC CCTTTTTATTAAATGAATGAGTTCC	201
	D12S2206	HEX	TAAAGGGGTCAATTCAGCAG TCTGGGTACTCCAGTGTTGG	107-127
	D12S337	FAM	CTGCAAAAACCCAGTGCT TTATGGAGGCTCTCCGAGT	133-139
	D12S1294	TET	CCAGTTTAGACACCCAGGAT ATATTTCAGAACCGCAGCAG	197-198
14	D14S1433	HEX	ACAGTTCCCAAGACTAAACAGC TGGAATTCAGTTATTTGTTAATGC	162-176
	D14S53	TET	CAACAAGAGCGAAACTCGC GAAGACTCAAGATATAGCAG	135-155
	D14S983	TET	TGGACTGGTTAGCCTCAGTG GCATCAACTGGCTTCCAATC	222-270
	D14S1020	FAM	GCCTTTACAGAGGGACTCATC TCTACTGGGAGCTAGGGCAC	228-268
	D14S1008	FAM	TGGCAAGGATGTAGAACACT CTTAGCCTGGGTGACAGAG	112-126
	D14S1000	HEX	TGGATTGTATTGCCAACTG TGTGTACCCAAGCATAAGTAGG	107-133

Table 4.2. STR markers used to fine mapping regions on chromosomes 12 and 14 in MMD family. Oligo F was labelled as indicated. Oligo R had no label. ASR= average size range. Sequences and ASRs were obtained from UCSC genome browser (http://genome.ucsc.edu/).

2.4 Two-point, multipoint and haplotype analysis.

Two-point lod scores were calculated using MLINK from the package LINKAGE [44] and assuming autosomal dominant inheritance. Allele frequencies were those provided by deCODE. They, in turn, estimated them from a pool of samples of

individuals with different ethnic backgrounds. Penetrance was set to 0.985 (same as in the power simulation).

Multipoint lod scores and haplotypes were obtained with simwalk2 [49] using same conditions as for two-point analysis.

3. Results.

3.1 Power simulation

Table 4.3 shows power simulation results. This simulation indicated that although no $Z \geq 3$ is likely to be reached in this family, a maximum lod score of 2.37, was expected at zero recombination fraction.

As the power of the pedigree structure was not high enough as to reach a lod score of 3, the pathological findings were greatly valued to make the decision of running a genome wide scan.

θ	Expected Lod Score (Z)		Probability of reaching $Z \geq$ to:		
	Average	Z max	1	2	3
0	1.34	2.37	66	34	0
0.05	1	2.14	52	21	0

Table 4.3. Power simulation of family with MMD.

Power simulation also indicated an approximately 50% chance of reaching a $Z \geq 1$ at $\theta=$ 0.05 was provided by the pedigree structure.

3.2 Genotyping results

Table 4.4 shows statistical summary of deCODE marker typing. Although there was no data at all for marker locus D1S219, and a high failure rate for marker D12S1294, deCODE marker typing efficiency was as high as 97%. The individual with the least average genotypes was III:2 (87%). In contrast the individuals with the most genotypes were II:1, II:2, III:5 and III:6 (99%). Per chromosome, the least genotypes were obtained for chromosome 9 (92%). The highest typing efficiency was obtained for chromosome 22 (100%, Table 4.4).

Marker typing efficiency for markers on candidate regions on chromosomes 12 and 14 was 100%.

3.3 Two-point analysis

Tables 4.6-4.27 show two-point lod scores (Z) for chromosomes 1-22 (see appendix 2), respectively and a summary of lod scores above 0.9 is shown in Table 4.5.

Two-point lod scores above 0.9, at any recombination fraction, were found on chromosomes 1 (Z=1.7), 5 (Z=1.03), 7 (Z=0.97), 11(Z=0.97), 14 (Z=1.18) and 16 (Z=1.23) (See summary in Table 4.5).

On chromosome 1, the highest lod score (Z=1.7 at θ= 0) was obtained for marker locus D1S452 (Table 4.5), where allele 4 was segregating with the disorder. Individual II:2, who is unaffected, received allele 4 and the remaining unaffected individuals alternatively received allele 7 (Figure 4.2).

On chromosome 11 there were three markers where the same score was obtained (Z= 0.97; Table 4.5). Marker locus D11S937 segregated allele 10 to the affected

INDIVIDUAL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Av
												CHROMOSOME												
II:1	0.98	1.00	0.97	1.00	0.96	1.00	0.93	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.93	0.93	1.00	1.00	1.00	1.00	C
II:2	1.00	1.00	1.00	0.97	1.00	0.96	0.97	1.00	0.96	1.00	1.00	0.96	0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	C
II:4	0.98	1.00	0.97	0.97	1.00	0.96	0.97	1.00	0.84	1.00	0.95	0.96	0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.96	C
II:5	0.98	0.94	1.00	1.00	1.00	0.96	0.97	0.91	0.92	1.00	0.95	1.00	1.00	1.00	0.95	0.94	0.95	0.93	0.93	1.00	1.00	1.00	1.00	C
III:1	0.96	0.97	0.94	1.00	1.00	0.96	0.97	1.00	0.96	0.89	1.00	0.93	1.00	1.00	0.95	1.00	1.00	1.00	0.93	0.93	1.00	1.00	1.00	C
III:2	0.80	0.83	0.91	0.93	0.96	0.75	0.87	0.87	0.68	0.89	0.85	0.86	0.82	0.96	0.90	0.89	0.86	0.93	0.93	0.80	0.93	1.00	0.88	C
III:4	0.98	1.00	0.97	1.00	0.96	1.00	0.93	0.91	0.96	1.00	1.00	1.00	1.00	1.00	0.95	0.94	1.00	1.00	0.93	0.93	0.93	1.00	1.00	C
III:5	1.00	1.00	0.97	1.00	1.00	1.00	1.00	1.00	0.96	1.00	1.00	0.96	1.00	1.00	1.00	1.00	1.00	1.00	0.93	1.00	1.00	1.00	1.00	C
III:6	1.00	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.96	0.95	0.96	0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	C
III:7	0.98	0.94	0.97	0.97	1.00	1.00	1.00	1.00	0.88	1.00	0.95	0.89	0.82	0.96	0.95	1.00	0.91	1.00	0.93	0.93	1.00	1.00	1.00	C
III:8	1.00	0.97	0.97	1.00	1.00	1.00	0.97	0.96	0.96	1.00	0.90	0.89	0.94	1.00	0.90	1.00	1.00	0.93	1.00	0.93	1.00	1.00	0.96	C
Average	0.97	0.97	0.97	0.98	0.99	0.96	0.96	0.97	0.92	0.98	0.96	0.95	0.95	0.99	0.96	0.98	0.98	0.98	0.96	0.96	0.99	1.00	0.98	C

Table 4.4. deCODE marker typing efficiency.

individuals with the only exception of individual III:7, who is affected and received

allele 12. Unaffected individuals alternatively received allele 12 (Figure 4.3). For

marker D11S1886 allele 3 was transmitted to affected individuals, with exception of

Locus/ Chromosome	Genetic map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
Chromosome 1								
D1S2628	168.68	-0.35	-0.19	-0.1	-0.04	-0.01	0.5	0
D1S452	176.39	1.69	1.34	0.96	0.55	0.18	0	**1.7**
D1S2818	187.03	-3.1	-0.73	-0.22	-0.04	0	0.5	0
Chromosome 5								
D5S1725	106.88	0.29	0.21	0.13	0.06	0.02	0	0.29
D5S409	114.89	0.26	0.78	0.67	0.42	0.14	0.1	0.78
D5S2027	120.95	0.55	1.02	0.84	0.52	0.17	0.088	**1.03**
D5S2065	122.49	0.55	1.02	0.84	0.52	0.17	0.088	**1.03**
D5S490	134.52	-0.33	-0.18	-0.09	-0.04	-0.01	0.5	0
Chromosome 7								
D7S2459	120.05	-4.46	0.71	0.62	0.37	0.11	0.113	0.72
D7S2418	122.15	-5.27	0.97	0.82	0.51	0.17	0.101	**0.97**
D7S530	133.2	-6.08	0.39	0.44	0.32	0.12	0.173	0.45
Chromosome 11								
D11S4087	73.81	-4.12	0.06	0.23	0.16	0.05	0.204	0.23
D11S937	85.42	-5.27	0.97	0.82	0.51	0.17	0.101	**0.97**
D11S1780	94.12	-5.56	-0.3	-0.1	-0.03	-0.01	0.5	0
D11S1886	107.71	-7.01	0.97	0.82	0.51	0.17	0.101	**0.97**
D11S4206	110.81	0.34	0.25	0.15	0.08	0.02	0	0.34
D11S908	118.66	-0.13	-0.08	-0.04	-0.02	0	0.5	0
D11S4089	126.83	-6	0.43	0.47	0.33	0.12	0.164	0.48
D11S4151	135.73	-4.33	0.97	0.82	0.51	0.17	0.101	**0.97**
D11S2367	148.23	-4.17	0.43	0.47	0.33	0.12	0.163	0.48
Chromosome 14								
D14S81	99.25	-7.09	0.07	0.24	0.17	0.05	0.203	0.24
D14S1054	102.52	1.18	0.91	0.62	0.33	0.1	0.001	**1.18**
D14S987	106.22	-7.49	-0.78	-0.23	-0.03	0.01	0.5	0.24
Chromosome 16								
D16S505	108.07	0.55	1.07	0.93	0.64	0.26	0.098	1.07
D16S763	120.63	1.23	1.07	0.84	0.54	0.2	0	**1.23**
D16S2621	131.46	-6.16	0.11	0.33	0.28	0.11	0.222	0.34

Table 4.5: Summary of Two-point LOD scores above 0.9. '
For a complete information see appendix 2.
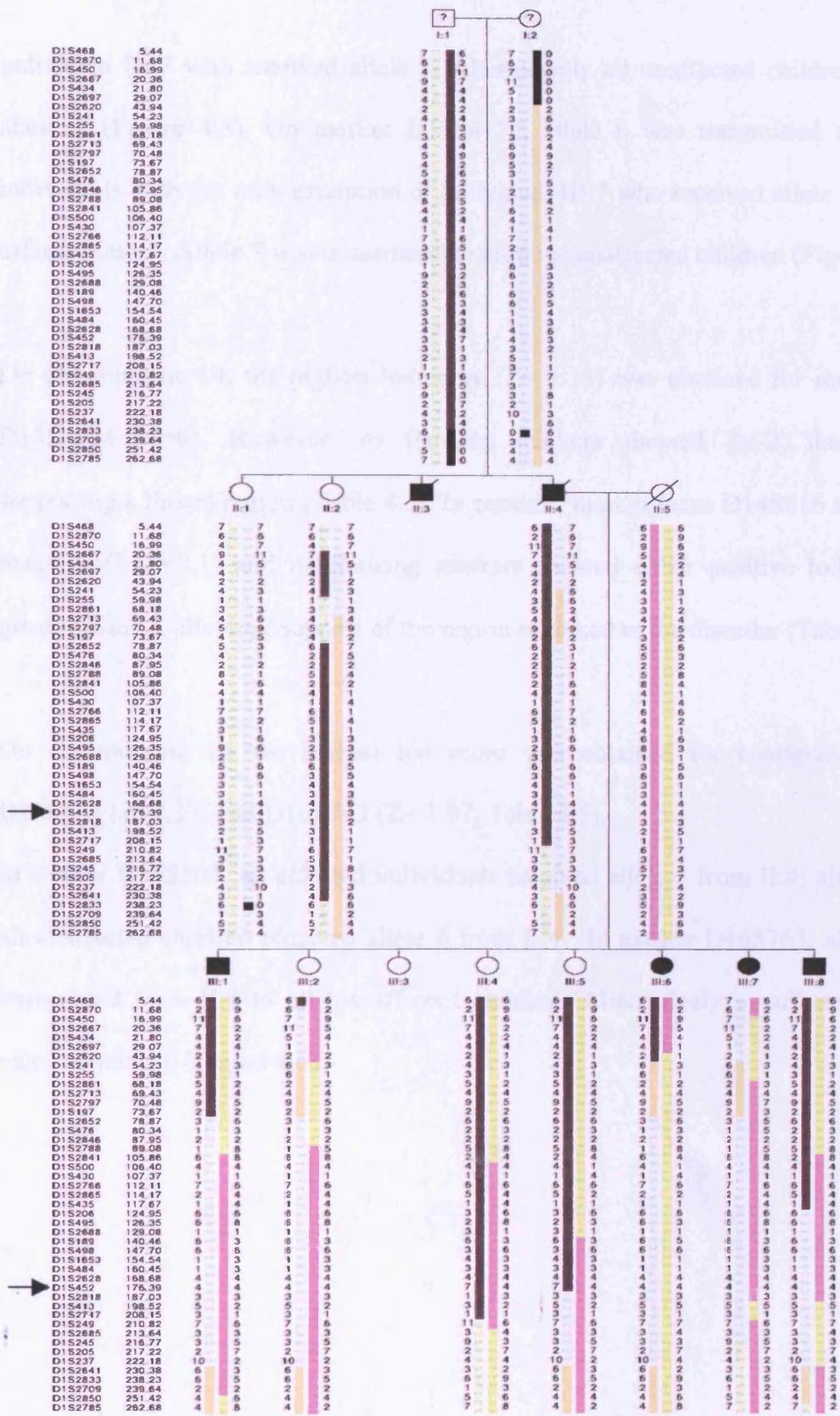
Figure 4.2. Markers tested on chromosome 1. Marker D1S452 showed the highest two-point lod score.

individual III:7 who received allele 1. Alternatively all unaffected children received allele 1 (Figure 4.3). On marker D11S4151, allele 8 was transmitted to affected individuals with the only exception of individual III:7 who received allele 7 from the affected father. Allele 7 was transmitted to all three unaffected children (Figure 4.3).

On chromosome 14, the highest lod score (Z= 1.18) was obtained for marker locus D14S1054 (θ=0). However, its flanking markers showed Z<-2, therefore not supporting a linked region (Table 4.5). In contrast, marker locus D14S616 showed a Z max= 0.97 (θ=0.1) and its flanking markers showed either positive lod scores or greater than -2, allowing support of the region as linked to the disorder (Table 4.5).

On chromosome 16 the highest lod score was obtained for contiguous markers D16S505 (Z=1.23) and D16S763 (Z= 1.07; Table 4.5).
In marker D16S505, all affected individuals received allele 5 from II:4; alternatively, all unaffected children received allele 6 from II:4. In marker D16S763, allele 4 was transmitted from II:4 to all his affected children. Alternatively, unaffected children received allele 6 (Figure 4.5).

80
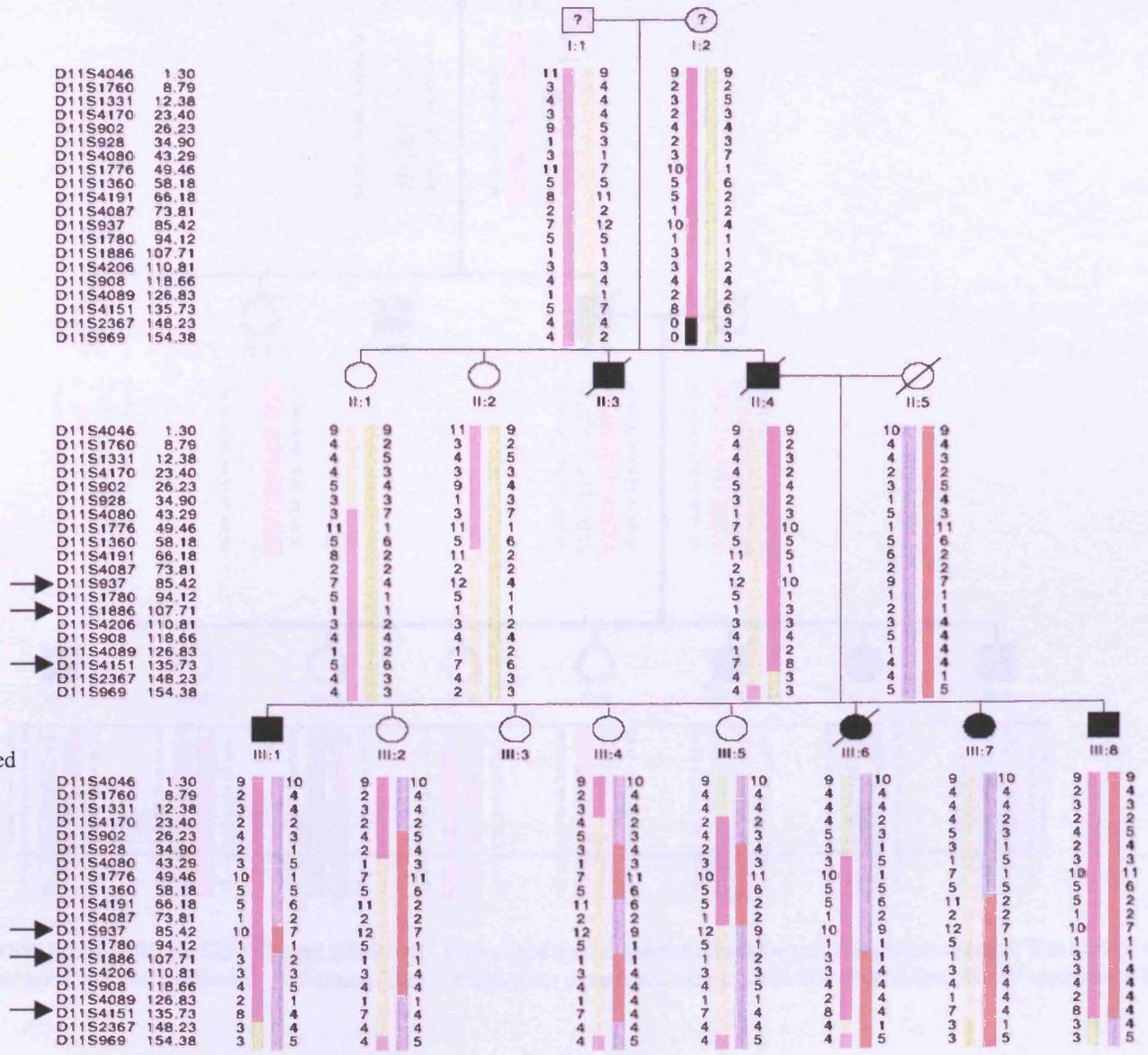
Figure 4.3. Markers tested on chromosome 11. Markers D11S937, D11S1886 and D114151 showed the highest two-point lod scores.

D14S258 0.00
D14S1433 4.94
D14S53 5.31
D14S983 6.32
D14S1020 7.68
D14S1008 10.02
D14S1000 11.18
D14S616 15.17
D14S67 18.40
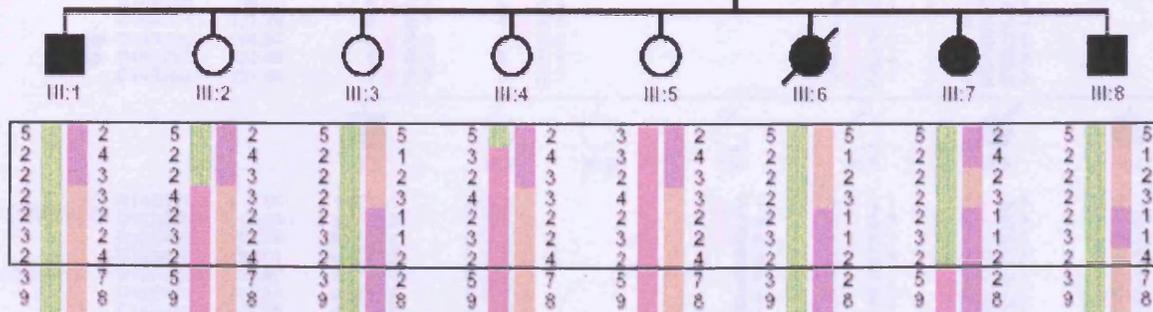
Figure 4.4. Fine mapping on chromosome 14.
Extra markers were D14S53, D14S983, D14S1020, D12S1008 and D14S1000. Underlined appear markers from the original genome search. The mutant chromosome (green) is carried by affected individuals and one unaffected. individual (III:3). Incomplete penetrance may explain this observation. Boxed appears the haplotype segregating with the disease

82

Figure 4.5. Markers tested on chromosome 16.
Markers D16S505 and D16S763 showed the highest two-point lod score.

3.4 Multpoint analyses

Figures 4.6, 4.7, 4.8, 4.9 and 4.10 show positive multipoint lod scores among all 23 chromosomes. Positive multipoint lod scores were obtained on chromosomes 1, 11, 12, 14 and 16. The highest multipoint lod scores were: 1.04 for chromosome 1 at map position 176.39 in the interval D1S2628-D1S2818 (Figure 4.6); 1.27 for chromosome 11 at map position 140.73 in the interval D11S4151-D11S2367 (Figure 4.7); 1.70 for chromosome12 at map position 89.31 in the interval D12S283-D12S351 (Figure 4.8); 2.1 for chromosome 14 at map position 74.94 in the interval D14S258-D14S616 (Figure 4.9); and Z=0.55 for chromosome 16 at map position 118.12 in the interval D16S516-D16S2621 (Figure 4.10). Despite having presented relatively low two-point lod scores (Table 4.17), chromosome 12 presented the second highest multipoint lod score after chromosome 14.

3.5 Haplotype analysis

Additional markers were added in the candidate regions identified by multipoint analysis on chromosomes 12 and 14.

On chromosome 12, addition of five markers in the candidate region led to the identification of a disease-associated haplotype, which included alleles 9-2-4-1-1-1-1-1 at markers D12S83-D12S1601-D12S1294-D12S1703-D12S2195-D12S2206-D12S337-D12S326 (Figure 4.11). However, this haplotype was present in one unaffected individual (III:2) and individual III:5 also presented this haplotype, with the only exception of allele 1 at marker D12S326, where this individual carries allele 4 (Figure 4.13); the presence of this haplotype in four affected siblings in addition to two unaffected individuals indicates no linkage of the disease to chromosome 12.

84
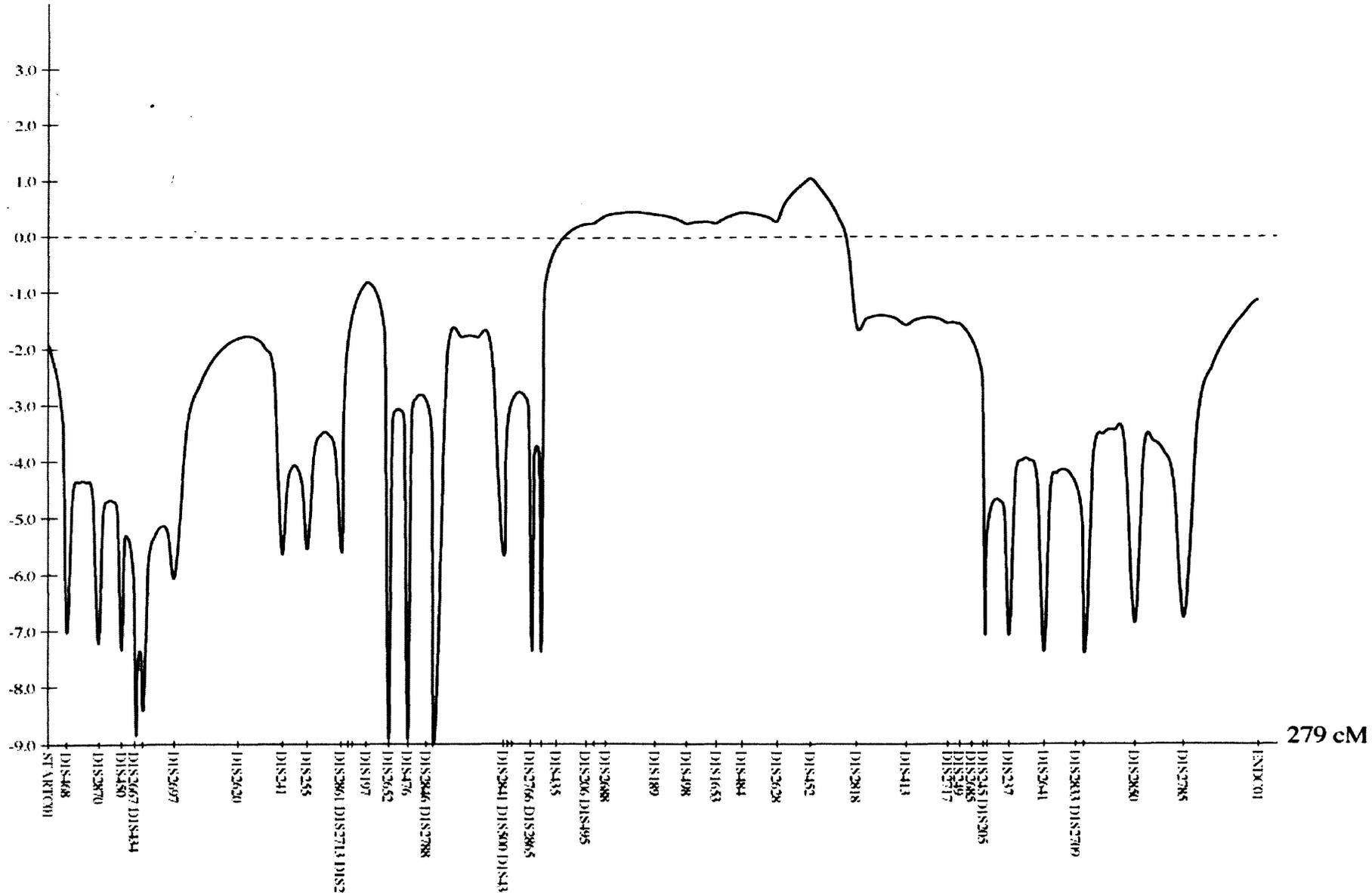
Figure 4.6. Plotting multipoint lod scores of chromosome 1

85
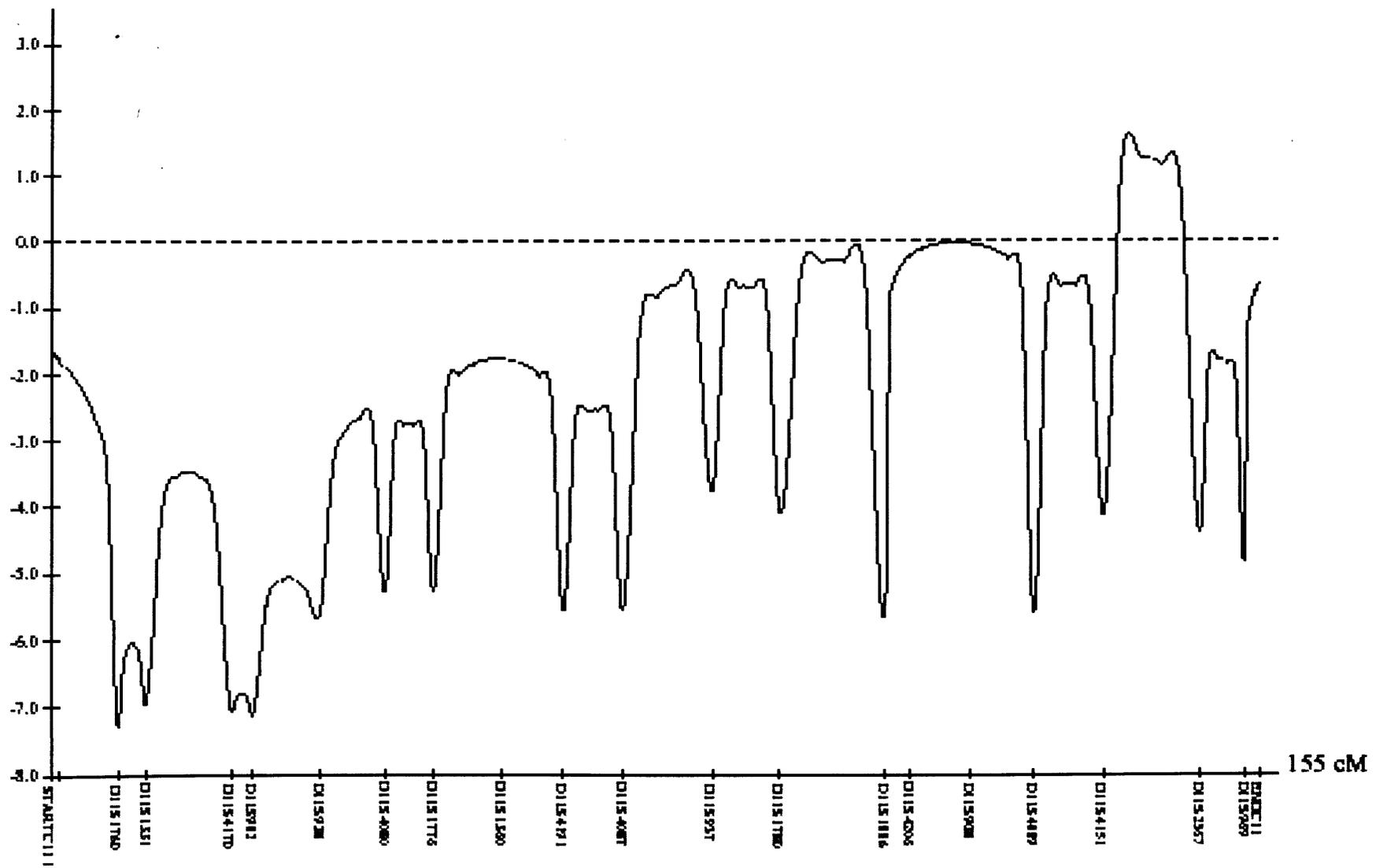
Figure 4.7. Plotting multipoint lod scores of chromosome 11

Tue Nov 16 15:25:56 2004

86
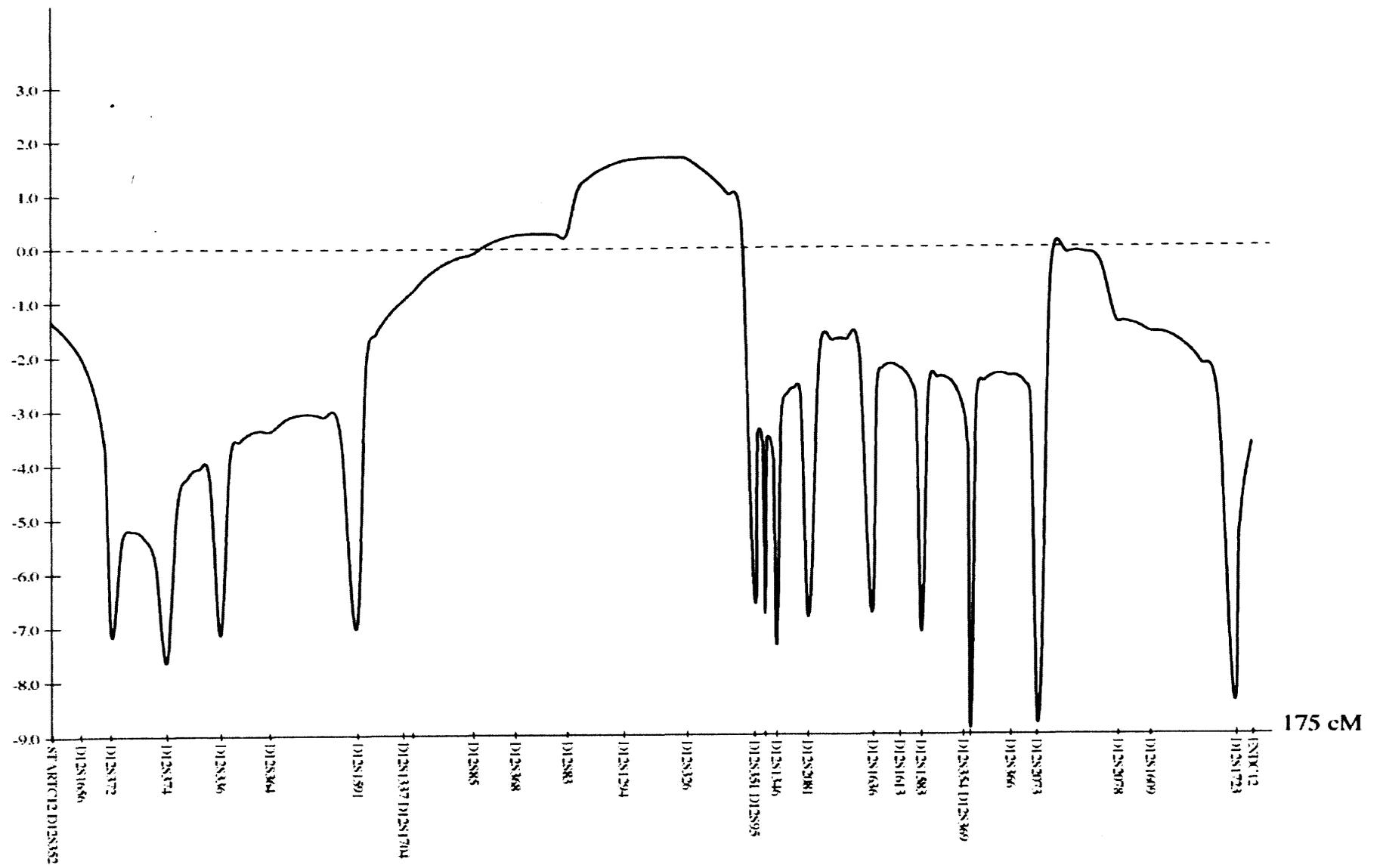
Figure 4.8. Plotting multipoint lod scores of chromosome 12

87
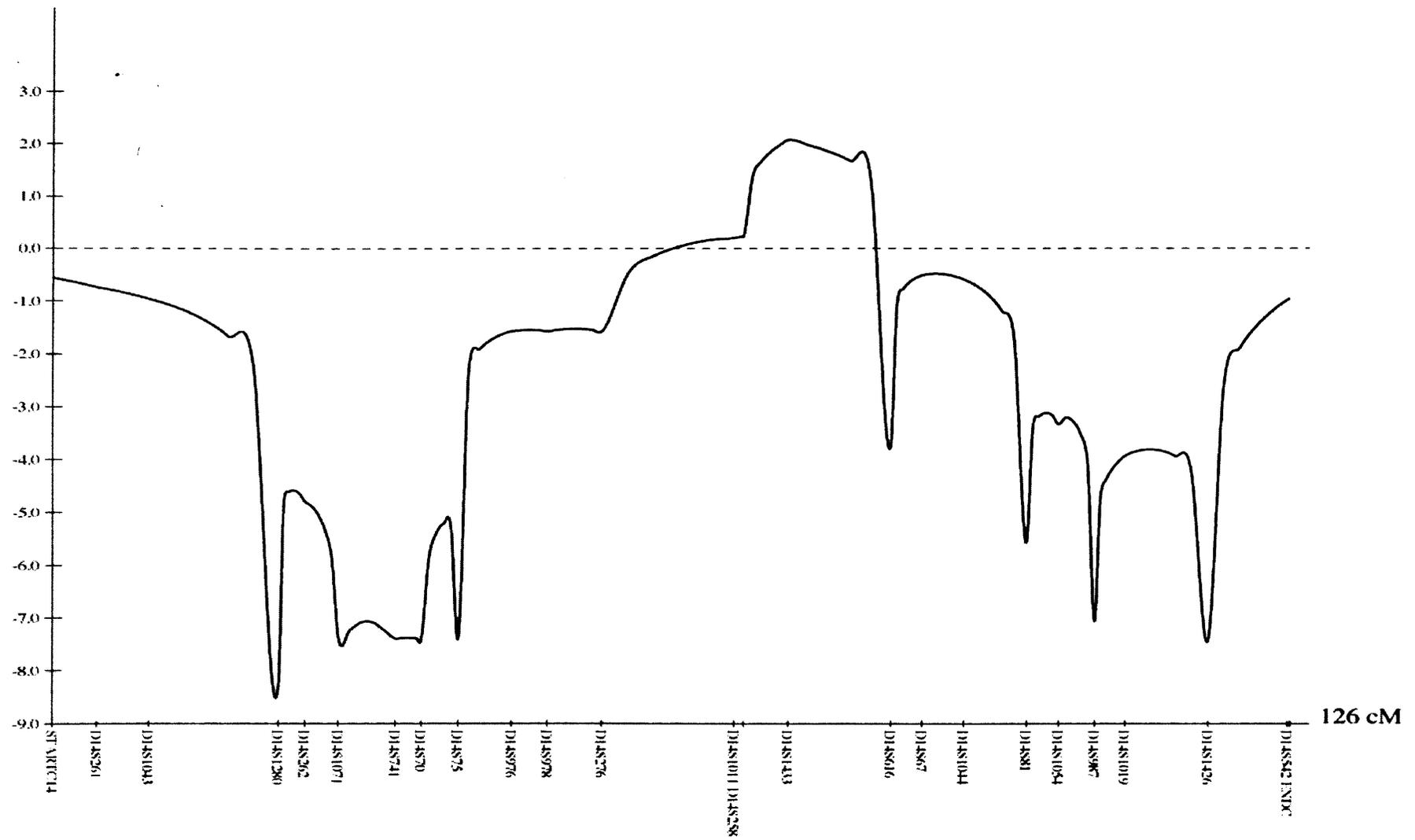
Figure 4.9. Plotting multipoint lod scores of chromosome 14

88
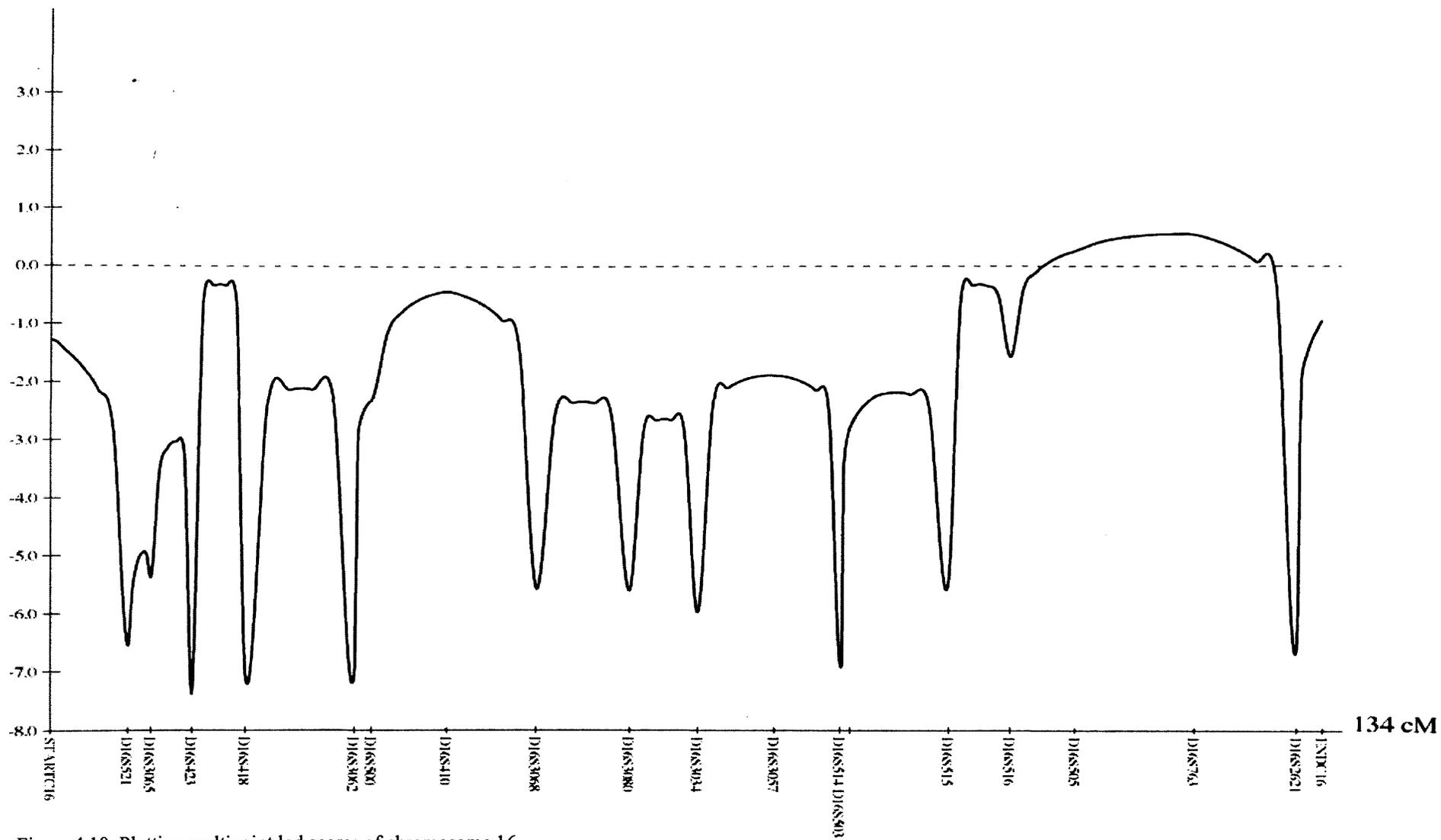
Figure 4.10. Plotting multipoint lod scores of chromosome 16
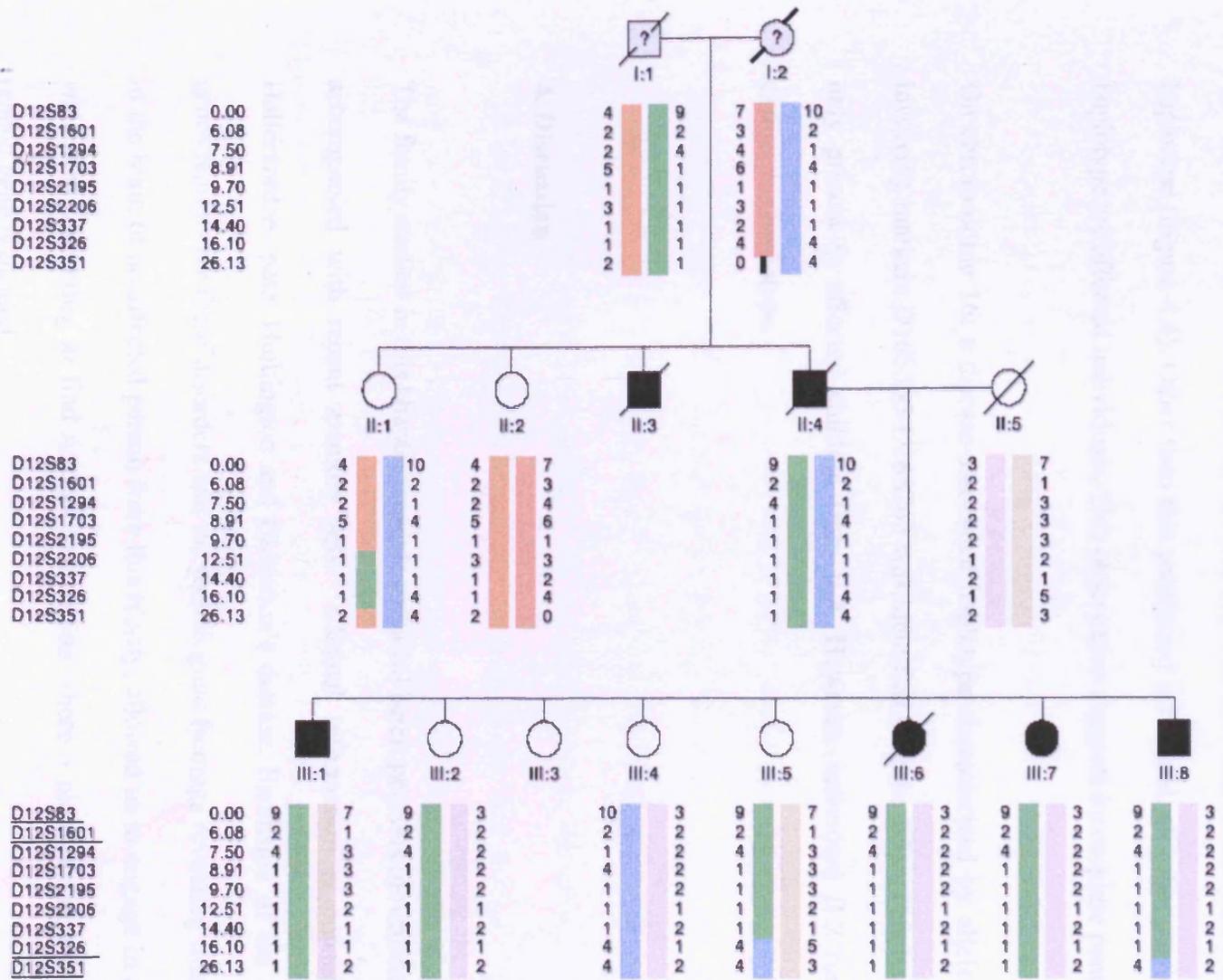
Figure 4.11. Fine mapping on chromosome 12.
Extra markers were D12S1601, D12S1703, D12S2195, D12S2206 and D12S337. Underlined appear markers from the original genome search. The mutant chromosome (green) is carried by affected individuals and two unaffected individuals, suggesting exclusion of this locus.

On chromosome 14, five additional markers were typed, including individual III:3, leading to the identification of a disease-associated haplotype characterized by alleles 5-2-2-2-2-3-2 at markers D14S258-D14S1433-D14S53-D14S983-D14S1020 D14S1008-D14S1000. However, the recently available individual (III:3) carries this haplotype (Figure 4.4). Other than this unaffected individual, all other carriers of this haplotype are affected individuals. This observation suggests incomplete penetrance.

On chromosome 16, a disease-associated haplotype characterized by alleles 5-4 and involving markers D16S505-D16S763 was identified (Figure 4.5). This haplotype is only present in affected children from II:4. However, individual II:2 (unaffected) carries this haplotype.

4. Discussion

The family studied in this chapter presents a clinical heterogeneous movement disorder accompanied with recent memory loss. Clinical information is consistent with Hallervorden-spatz, Huntington and Parkinson's disease. Exclusion of the candidate genes/regions for these disorders and the pathological findings revealing iron deposits in the brain of an affected person from this family, allowed us to engage in a genome wide scan, expecting to find linkage to a region where a plausible iron metabolism related gene is located.

Considering the low power observed in the power simulations (Z= 1.34 at θ= 0) and the fact that not many regions were identified in the two-point analysis, a cut off lower

than the expected average lod score of 1 at $\theta=0.05$ was used to define candidate regions. Thus, considering a two-point lod score ≥ 0.9, chromosomes 1, 5, 7, 11, 14 and 16 were initially identified as candidates. After multipoint analysis for all chromosomes, positive lod scores were obtained only for chromosomes 1, 11, 14 and 16. In addition, Chromosome 12 presented a lod score of 1.70 in the interval D12S1294-D12S326. All lod scores obtained in these analyses were lower than 3 ($Z<3$), suggesting that they all could be false positive results [42].

Fine mapping was carried out for regions on chromosomes 12 and 14. For chromosome 12, it was found that two unaffected individuals were carrying the disease-associated haplotype (Figure 4.11), which suggests exclusion of this region as harbouring a mutant gene causing the disorder in this family. Fine mapping of the region on chromosome 14 included a new recently available individual (III:3). Haplotype segregation was perfect with the disorder, with exception of the new individual, who carries the disease-associated haplotype (Figure 4.4). Incomplete penetrance may explain why an unaffected person carries the mutant chromosome.

The Neuroglobin gene (*NGB*), which is an important candidate gene related to the iron metabolism, lies in this interval. NGB is a brain oxygen transporter [169]. Levels of this transporter have been differentially found in the brain of people with disorders such as Alzheimer's disease, where a reduced capability of oxygen transport has been associated with neuronal death, triggering the disease [170]. It is suspected that other related disorders, such as Parkinson, also have links with the level or efficiency of oxygen transport by this molecule [171], leaving open the possibility that this gene is implicated in the disorder presented here.

If chromosome 14 was definitely excluded, candidate regions on chromosomes 1, 11 and 16 should undergo fine mapping. Based on the multipoint lod scores, the order of further evaluation could be as follows: Chromosome 11, then chromosome 1 and finally chromosome 16. Interestingly, on chromosome 11 there were three hints of linkage involving markers D11S937, D11S1886 and D11S4151. The two first were shown in the two-point analysis but not in the multipoint analysis. Markers D11S4087 and D11S4206 flank this region, which extends approximately 37 cM. The heavy-ferritin chain gene (*FTH1*) lies in this region. Even though the haplotype and multipoint analysis do not support this locus, a double recombinant in individual III:7, in this interval, would fit with *FTH1* as responsible for the disease in this family. IRE mutations of *FTH1* have been reported in a large Japanese family presenting with iron deposits in liver and spleen [158]. *FTH1* mutations could clearly explain the pathology findings and the disease itself in our family. This is assumed based on reports by Curtis and colleagues (2001) of a disorder clinically similar to the one presented here, where brain pathology is also similar to the one presented by our patients. In their study, they found mutations in the ferritin light chain (*FTL*) and iron deposits[159] in a similar manner to our family. We excluded *FTL* and therefore *FTH1* is a very reasonable candidate gene.

However, if no mutation was identified in *FTH1*, despite no evident candidate gene in the second region on chromosome 11, this should be evaluated in more detail by typing extra markers in the region of D11S4151.

Candidate genes in the chromosomal region on chromosome 1 around marker D1S452 are not evident. However, initial testing of extra markers would help us focus in particular chromosomal regions where plausible candidate genes locate.

Candidate genes in the interval D16S505-D16S763 (on chromosome 16) are not clear. However, EFCBP2 presents relatively high expression in brain and is known to bind calcium in neurons.

In addition, chromosomes 5 and 7 are to be taken into consideration for further examination in case of excluding all candidate regions where positive multipoint lod scores were obtained.

In conclusion, no perfect haplotype segregation with the disorder has yet been found. However, up to date data suggests that *NGB* and *FTH1* genes are both positional and functional candidate genes in the family studied here.

Furthermore, if chromosomes 14 and 11 turn out to be excluded, other regions on four different chromosomes (1, 5, 7 and 16) can still be further analyzed in searching for the mutant gene responsible for the disorder in this family.

CHAPTER FIVE

LINKAGE ANALYSIS OF FAMILIES

WITH

GENERALIZED EPILEPSY WITH

FEBRILE SEIZURES PLUS, GEFS+

1. Introduction

1.1 General aspects on epilepsies.

Epilepsy is a group of neurological disorders characterized by a marked increment of the neuronal excitability inducing recurrent auto limited convulsions in the absence of acute triggering factors [172, 173]. Attacks may remain confined to an elemental or complex alteration of the behaviour or can progress to focal or generalized convulsions. Convulsions can be partial or generalized, depending on whether the epileptic discharges involve, in the beginning, one hemisphere or both, respectively [174]. The term epilepsy should not be used to describe single epileptic seizures, seizures that occur during the process of an acute illness or occasional provoked seizures (i.e. febrile seizures) (reviewed in [175]).

Reported prevalence rates in epilepsy vary between 0.2 and 4.2 % [176]. In Colombia, particularly in Medellin-Antioquia, it has been estimated that the prevalence is 2.14 % [177].

There are over 40 clinical forms of epilepsy, which are classified according to their aetiology and/or their clinical presentation [174]. Epilepsies are classified as partial and generalized, which also can be further classified as symptomatic or idiopathic. In the first a discernible cause (e.g. brain or metabolism abnormalities) can be established.

The term idiopathic epilepsy is reserved for those cases affected by convulsions where neither brain structure or metabolism abnormalities are detected, and are presumed to have genetic basis. Idiopathic epilepsy is age related, usually has good response to

treatment and may present either Mendelian or complex inheritance [178]. According to the type of convulsion, idiopathic epilepsy is subdivided into Idiopathic generalized epilepsy (IGE) and Idiopathic partial epilepsy (IPE). Recent findings suggest that some symptomatic epilepsies have an important genetic component [179].

Twin studies have found a high concordance rate for the type of epilepsy. Thus, 94% of monozygotic twins presented the same subtype of epilepsy, whereas 71% of dizygotic twins shared their epilepsy subtype [180], which indicates a strong genetic component. Despite the fact that several studies have confirmed the importance of genetic factors in the expression of idiopathic epilepsy, its genetic basis is complex and very likely involves locus and allelic heterogeneity, in addition to environmental factors [181].

There are four main types of IGE. These are Juvenile myoclonic epilepsy (JME), Childhood and Juvenile absenses epilepsy (CAE and JAE) and epilepsy with grand mal seizures on awakening (EGMA). The commonest of these is JME.

In recent years, several genes related to epilepsy have been mapped by linkage and association studies; so far, over 15 autosomal genes and one mitochondrial one have been identified (Table 5.1).

Moreover, several genetic linkage studies have provided evidence, sometimes inconsistent, of the influence of other loci conferring susceptibility to develop epilepsy. Sometimes the same locus has been linked to different forms of epilepsy and, also, sometimes a single form of epilepsy has shown linkage to different loci (Table 5.1).

Clinical form	Chromosomal location	Gene/Locus	References
Partial			
Nocturnal Autosomal dominant Epilepsy of frontal lobe	20q13.2	*CHRNA4*	182
	1q	*CHRNB2*	183
Autosomal dominant Partial Epilepsy with auditive characteristics	10q	*LGI1*	184, 185
Rolandic Epilepsy	15q14	*CHRNA7?*	186
Generalized			
Progressive Myoclonic Epilepsy type 1	21q22.3	*CSTB*	187-189
Progressive Myoclonic Epilepsy with red ragged fibbers, MERRF	mtDNA	tRNA(lys)	190
Batten disease	16p12.1-11.2	*CLN3*	191-193
Progressive Myoclonic Epilepsy of Lafora's type	6q23-25	*EPM2A*	194, 195
Progressive epilepsy with mental retardation	8p	*CLN8*	196, 197
Idiopathic generalized			
Juvenile Myoclonic Epilepsy	6p21.2	?(HLA) (EMJ1)	198-201
	15q14	? (EMJ2)	202, 203
	5q34	*GABRA1* (EMJ4)	204
Juvenile Myoclonic Epilepsy with absences	1p	? (EMJ3)	205
IGE (JME, CAE, JAE, EGMA)	3q26	*CLCN2*	206
Myoclonic Epilepsy autosomal recessive of the infancy	16p13	?	207
Generalized idiopathic epilepsy	8q24	?	208
	3p	?	209
Benign Neonatal Convulsions	20q13.2	*KCNQ2* (EBN1)	210-212
	8q24	*KCNQ3* (EBN2)	213, 214
Special Syndromes			
· Febrile Convulsions, FS	8q13-21	? (FEB1)	215
	19p13.3	? (FEB2)	216
	5q14-15	MASS1(FEB4)	217, 218
· Febrile Convulsions Plus, GEFS+	19q13	*SCN1B* (GEFSP1)	219
	2q24	*SCN1A* (GEFSP2)	220-222
		SCN2A	223
	5q32-35	*GABRG2* (GEFSP3)	224

· Table 5.1. Molecular genetics of epilepsies. (Modified from [225]).

Even though, IGE is classified as a genetic disorder, the majority of cases lack evidence of a clear genetic basis. Part of the difficulty of identifying susceptibility genes originates from incomplete penetrance, variable expressivity and genetic heterogeneity, in addition to environmental factors.

It is expected that identification of genes responsible for Mendelian forms of common types of IGE will help identify susceptibility genes for the common forms. In addition, large-scale genome screening methods applied to large populations of patients with common types of epilepsy will help to identify variations in other genes that are associated with the disease.

1.2 Genes in GEFS+

In contrast to most cases of IGE, Generalized epilepsy with febrile seizures plus (GEFS+) presents monogenic inheritance. Generalized seizures during fever episodes (febrile seizures, FS) occur in 2-5% of children under six [226]. Of these, about 5% develop non-febrile seizures or epilepsy later in life (FS+). GEFS+ is characterized by FS and FS+, or because they present non-febrile tonic-clonic seizures [227].

Historically, FS and epilepsies have been considered different conditions; however, GEFS+ connects these two entities showing that there is a genetic link between them. GEFS+ has been described in families presenting with febrile convulsions and following an autosomal dominant inheritance mode with incomplete penetrance [228].

Four genes causing GEFS+ and one causing FS have been identified so far.

SCN1B (GEFSP1)

First, the voltage gated B1 sodium neuronal channel encoded by the gene *SCN1B*

presented mutations in an Australian family, in which 26 of its members had GEFS+.

The approach used to identify this gene involved linkage analysis, using a penetrance

of 64% and a phenocopy rate of 3% [85, 215]. They found that a region on chromosome

19q13.1 was co-segregating with the disorder. *SCN1B* lies in this region, which has

five coding exons and extends 9.73 Kb. Single strand conformational analysis

followed by sequencing revealed a transversion of C to G in nucleotide 387 (exon 3),

leading to the substitution of Cys121Trp [219].

This mutation changes a conserved cysteine residue, altering a disulfide bridge that

normally maintains an extracellular fold. A second GEFSP1 family was reported,

where a C121W mutation in SCN1B was found [229]. This mutation was not present in

96 control individuals.

SCN1A (GEFSP2)

The second GEFS+ locus was also identified by linkage analysis. Baulac and

colleagues (1999) studied a large family where *SCN1B* was excluded and identified

linkage to a 20 cM region on chromosome 2q24-q33 [230]. Moulard and colleagues

subsequently studied five additional families with GEFS+ (and one with FS), being

able to refine the linked region on chromosome 2 [231]. The following year, Escayg and

colleagues (2000) [220] found that the family described by Baulac and colleagues (1999)

carried a mutation in the *SCN1A* gene, which leads to the substitution Arg1648His.

They also found that one of the families described by Moulard and colleagues (1999)

carried a mutation in codon 875, causing the substitution Thr to Met. These nucleotide

changes were not found in more than 100 controls.

SCN1A has 26 exons and extends 84.48 Kb. It encodes the sodium channel α-subunit, which is comprised of four domains (D1-D4), each of which is composed of six transmembrane segments (S1-S6).

A wide range of point mutations have now been reported in *SCN1A*, which have been associated with GEFS+ [220-222, 232-235] and with other types of epilepsy including Severe myoclonic epilepsy of infancy (SMEI) [236-239]. Most of SMEI mutations have been reported in the pore forming loop of the protein and the majority of them are nonsense mutations [236], in contrast to those associated with GEFS+, which are all missense mutations.

Functional studies of some of the reported *SCN1A* mutations have shown disruption of channel inactivation, resulting in increased Na+ influx and leading to hyperexcitability and suggesting a gain of function [240]. In contrast, *SCN1B* mutations in GEFSP1 have been identified as loss-of-function mutations, which also indirectly decrease the rate of inactivation of the sodium channel [219].

SCN2A

SCN2A lies in the same chromosomal region as *SCN1A*, 2q24-q33. Together with *SCN3A* and *SCN9A*, they constitute a neuronal sodium channel (SCN) cluster on this chromosomal region [221] (Figure 5.1). Sugawara and colleagues (2001) evaluated nineteen Japanese families with GEFS+. They looked for mutations in *SCN1B*, *SCN1A* and *SCN2A* genes in the index cases from these families. They included *SCN2A* because it is also expressed in high levels in the central nervous system with a tissue specific profile. They found two additional mutations in SCN1A and, for the first time, a mutation in SCN2A, which consisted of a c.1571 G>A transition, leading to the

substitution R187W. This mutation was not found in 112 controls and is highly conserved among the voltage gated sodium channel alpha subunit family members and through evolution [223].

GABRG2 (GEFSP3)

More recently, a fourth gene associated with GEFS+ was identified in 5q34 (GEFSP3). It was known beforehand that a cluster of gamma amino-butyric acid (GABA$_A$) receptor subunit genes is located in this region, including α1 (GABRA1), α6 (GABRA6), β2 (GABRB2) and γ2 (GABRG2) subunits [241, 242]. These genes have nine (GABRG2 and GABRA6) and eleven exons (GABRA1 and GABRB2). After having found linkage to this region in a French family with GEFS+, sequencing of GABRG2, which extends 85.71 Kb, revealed an A to T transversion in exon 8, leading to the substitution K289M [243]. In this family, 13 affected, 2 obligated carriers and one unaffected individuals were heterozygous for this mutation. This mutation lies in a highly conserved segment localized in the extracellular loop between transmembrane segments M2 and M3, and was not found in 800 control chromosomes. Subsequently, more GABRG2 mutations have been reported in GEFS+ [244-246].

GABA$_A$ receptors are heteropentameric proteins, with an integral chloride channel mediating rapid synaptic inhibition, and possess specific binding sites for GABA, benzodiazepines, barbiturics and steroids [247]. The GABA$_A$ receptor gene family includes six major classes of subunits, which are α, β, δ, γ, ε and π. The α-subunit has six subtypes; β-subunit, four subtypes; γ-subunit, three subtypes; and the remaining three types γ, ε and π only have one form [247]. The subunit combination most frequently found in the brain is α1β2γ2 [248].

In addition to the four GEFS+ genes described, mutations in one gene have been found in FS. This study was started in a large Japanese FS family and then 39 additional FS families confirmed linkage to the chromosomal region 5q14-15 [217] (*FEB4*). Based on the report of a naturally occurring murine mutation in the gene *mass1* (Monogenic audiogenic seizure-susceptible), which causes audiogenic seizures in the Frings mouse strain [249], and knowing that the human orthologue maps to 5q14, Nakayama and colleagues looked for mutations in *MASS1* and found the disease causing S2652X mutation, as well as other variations. S2652X was not found in any of 200 control chromosomes and is predicted to disrupt the C-terminal end of the protein. A small repetitive motif of MASS1 shares homology with several sodium-calcium exchangers [218].

Furthermore, three other FS chromosomal regions awaiting for genetic identification have also been associated with FS, but not with afebrile seizures. These are *FEB1* at 8q13-q21 [215], *FEB2* at 19p13.3 [216] and *FEB3* at 2q23-24 [250].

Mutations in these genes account for less than 10% of families with GEFS+, indicating that other genes might more frequently play role in GEFS+ aetiology. This is supported by a number of families reported where the current known GEFS+ and FS genes have been excluded [223, 251]

Despite findings of disease causing mutations in the five described genes, that segregate with the disorder through families supporting the idea of a mendelian-like inheritance mode in GEFS+ and FS, other studies have supported a polygenic inheritance model [252], suggesting that the genes identified so far in GEFS+ might

correspond to Mendelian forms of the complex disorder, similar to what happens in IGE. Interestingly, those suggested genes in the polygenic inheritance model have not yet been identified

1.3 Aims

This study aimed to characterize a set of six families with GEFS+ from Antioquia-Colombia. First, by testing for linkage to the previously reported FEB loci and GEFS+ genes. Families showing exclusion will be studied by whole genome scan.

The characterization of these families is important because it will, eventually, lead to the identification of causing disease mutations that can then be used for diagnostic purposes, predictive tests in as yet unaffected children and will help to evaluate the degree of genetic heterogeneity in GEFS+ in this population. In addition, it might identify families, from this set, where exclusion for the GEFS+ genes and FEB loci would suggest the participation of unidentified genes.

2. Patients and Methods

2.1 Families

Six families are included in this study (Figures 5.3-5.8). All of them were ascertained in the Paediatric Neurology Service of Hospital Universitario San Vicente de Paul and Medicine Faculty of Universidad de Antioquia (Medellin-Colombia). Pedigrees were extended focusing on a history of seizures (febrile or non-febrile) in relatives of each index case. All available individuals with a positive history underwent a full neurological examination and the characteristics of the seizures evaluated by interviewing relatives and reviewing available clinical records. Seizures were classified according to international criteria [174]. Blood samples were collected from available individuals who signed a written informed consent, designed for this particular study. DNA was isolated using the QIAGEN DNA extraction kit following the manufacturer's instructions.

2.2 Power simulation

Power simulations were performed assuming locus homogeneity. It was assumed penetrance of 0.96 and a mutated gene frequency of 1/10000. A microsatellite marker locus with four equally frequent alleles was simulated in one hundred replicas with the software SLINK [55].

2.3 Genotyping and linkage analysis

Polymorphic microsatellite markers tightly linked to GEFS+ genes and FEB loci were evaluated (Table 5.2). Each marker had a fluorescent label to allow for pooling and to run them in one single lane onto an ABI-377 (Applied Biosystems).

105

MARKER	DYE	LOCUS/GENE	CHROMOSOME	ASR	REFERENCE
D8S530	6-FAM	FEB1	8q13	213-231	[215]
D19S177	HEX	FEB2	19p	159-177	[216]
D2S382	6-FAM	FEB3/SCN1A/ SCN2A	2q24	160-180	[230, 231]
D5S644	6-FAM	FEB4	5q14-15	90-116	[217]
D19S425	HEX	SCN1B	19q13	260-290	[219]
D5S1403	TET	GABRG2	5q34	210-240	[224]

Table 5.2. Polymorphic microsatellite markers evaluated in Colombian GEFS+ families. ASR= Average size range.

In some families additional markers were examined, in FEB3: D2S2330 and D2S2157 (Figure 5.1 and Figure 5.3).

In FEB2: D19S894, D19S216 and D19S1034 (Figure 5.4 and Figure 5.5).

In GABR: D5S436, D5S410, D5S1403, D5S1955, D5S2118, SHGC-132760 and D5S422 (Figure 5.2 and Figure 5.6).

Marker typing was performed as indicated in the section on Microsatellite marker typing.

Linkage analysis was performed using same conditions as for the simulations and assuming equal allele frequencies for alleles at every marker locus.

Additionally, analyses were performed with varying penetrances values and phenocopy rates. Thus Family 1 was also analyzed using pentrance= 0.96 and phenocopy rate of 3%. All families were also analyzed assuming penetrance= 0.65 and phenocopy rate of 2%.
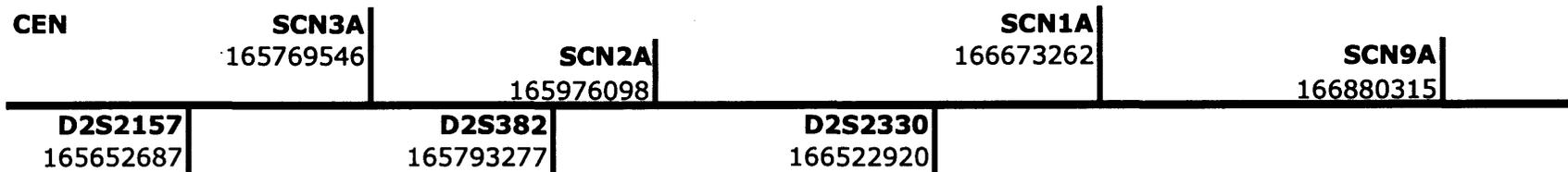
Figure 5.1. SCNA cluster at chromosomal region 2q24. Above are indicated the SCNA genes of this cluster. Underneath are indicated the microsatellite markers evaluated in this study. Physical positions are shown for both markers and genes. CEN= centromere. This diagram is not to scale.
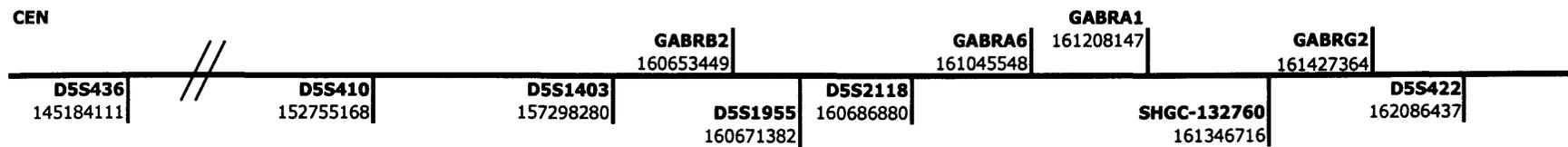
Figure 5.2. GABR cluster at chromosomal region 5q32. Above are indicated the GABR genes of this cluster. Underneath are indicated the microsatellite markers evaluated in this study. Physical positions are shown for both markers and genes. CEN= centromere. This diagram is not to scale.

107

2.4 Mutation screening

SCN1A exons 4, 15,18, 21 and 26 were amplified using primers and PCR conditions reported by Malacarne and colleagues [253]. Mutation screening was carried out by digesting the third fragment of exon 26 with *Tsp451* (New England Biolabs), at 37° C overnight. Digestion products were then examined on 3% agarose gels stained with ethidium bromide.

2.5 Aminoacid sequence alignment

Aminoacid alignment was done using CLUSTAL W in http://www.ebi.ac.uk/culstalw [254].

3. Results

3.1 Pedigrees

Clinical details in all these families were consistent with GEFS+. Families 1 and 7 included twelve affected individuals each (Figures 5.3 and 5.5, espectively). Family 5.3 included seven affected individuals (Figure 5.7); Families 4 and 6 each had six affected individuals (Figures 5.4 and 5.8). Family 5 included eight affected individuals (Figure 5.6). All cases in these families had a history of febrile seizures with the first episode occurring between 3 months and 3 years. Several of the patients who are over six years of age have also had afebrile seizures. All families are consistent with an autosomal dominant inheritance mode with incomplete penetrance. The highest rate of non-penetrance was present in Family 5.

3.2 Power analysis

Table 5.3 shows the simulation results. Assuming locus homogeneity, the set of families would provide a power of approximately 100% to identify significant linkage. Individually, only families 1 and 5 were powerful enough to reach a $Z \geq 3$. Assuming that no phenocopies are included in the families, all pedigrees, with exception of Family 3, provide an expected maximum lod score over 2.

Pedigree	Expected Lod Score (Z)		Probability of reaching a LOD score $\geq$ to:		
	Average $\theta=0$	Max	1	2	3
Family1	2.95	5	93	74	44
Family3	1.48	2.38	70	41	0
Family4	1.1	1.9	55	0	0
Family5	2.1	3.13	94	58	7
Family6	1.79	2.66	79	59	0
Family7	1.74	2.53	83	41	0
Total	11	15.78	100	100	100

Table 5.3. Power simulation for Colombian GEFS+ families.

3.3 Linkage and haplotype analysis

Table 5.4 shows two-point lod scores for markers at the *FEB* loci and *GEFS+* genes in our set of families. Even though markers were at zero cM from the respective genes/loci tested, only Family 3 showed its highest lod score at $\theta= 0$ (Marker D19S177).

FEB3/SCN1A/SCN2A locus

Linkage to the *FEB3* locus was suggested for families 1, 6 and 7 at theta= 0.1 (Table 5.4, marker D2S382). When the analysis assumed a phenocopy rate of 3%, lod scores increased dramatically for family 1 (Table 5.5). In this family, significantly negative lod scores (<-2) were observed at theta 0 for all markers tested with exception of D2S382 at FEB3/SCN1A/SCN2A, which showed a maximum lod score of 2.24. Genotypes at two additional markers in the *FEB3* region resulted in maximum lod

109

scores (at 0 recombination fraction) of 2.83 and 1.19 for D2S2330 (closer to the gene, Figure 5.1) and D2S2157, respectively (Figure 5.3).

An increase in lod scores when analyzing at low penetrance (0.65) and phenocopy rate = 0.02 was also observed in Families 6 and 7 (Table 5.6) for locus FEB3. Family 6 now showed its highest lod score for marker D2S382 at θ=0 (Zmax= 1.21), in contrast to Z max= 0.78 at θ= 0.1 assuming penetrance of 0.96 and phenocopy rate of zero (Table 5.4). Family 7 now presented Z= 0.62 at θ= 0 for marker D2S382 in contrast to Z= -3.12 at θ= 0 in the first analysis (Table 5.4). Similarly, Z at θ= 0.1 increased from 0.7 (Table 5.4) to 0.84 (Table 5.6) in this Family. Markers D2S2157 and D2S2330 at *FEB3* also showed suggestion of linkage in these families. Interestingly, Family 7 only presented its maximum lod score at θ= 0 for marker D2S2157 and not for markers D2S382 and D2S2330. This may be explained by one of two possibilities. One possibility is that actually a recombination has happened in the interval D2S2157-D2S382 taking away the disease-associated alleles in individual III:7 for markers D2S382 and D2S2330. This explanation implies that neither *SCN2A* or *SCN1A* is the carrier of the mutation in this family, and alternatively would probably implicate *SCN3A* (see Figure 5.5). The second possible explanation is that individual III:7 did not receive her allele 6 at marker D2S2157 from her mother and thus she has got a completely different haplotype from the rest of affected individuals in the pedigree and therefore she must correspond to a phenocopy.

Families 1, 6 and 7 together provided Z> 3 for all three marker loci at theta= 0, evidencing linkage to *FEB3* (Table 5.6).

MARKER   D19S177 (FEB2)

	Recombination fraction, θ					
	0	0.1	0.2	0.3	0.4	0.5
Total	-26.49	-5.85	-2.73	-1.25	-0.45	0
Fam7	-9.5	-1.33	-0.58	-0.23	-0.05	0
Fam6	0.15	0.09	0.05	0.02	0	0
Fam5	-9.01	-2.42	-1.1	-0.48	-0.15	0
Fam4	0.06	0.09	0.07	0.03	0	0
Fam3	0.32	0.27	0.21	0.14	0.07	0
Fam1	-8.51	-2.56	-1.4	-0.77	-0.35	0

MARKER   D2S382 (FEB3/SCN1A/SCN1B)

	Recombination fraction, θ					
	0	0.1	0.2	0.3	0.4	0.5
Total	-27.75	-2.49	-0.33	0.16	0.11	0
Fam7	-3.12	0.70	0.60	0.36	0.11	0
Fam6	0.65	0.78	0.6	0.34	0.13	0
Fam5	-11.32	-2.33	-0.9	-0.26	0	0
Fam4	-3.53	-1.38	-0.8	-0.45	-0.19	0
Fam3	-6.56	-0.94	-0.4	-0.16	-0.04	0
Fam1	-3.87	0.68	0.57	0.33	0.1	0

MARKER   D5S1403 (GABRG2)

	Recombination fraction, θ					
	0	0.1	0.2	0.3	0.4	0.5
Total	-38.98	-4.51	-1.14	0.12	0.36	0
Fam7	-9.51	-1.57	-0.78	-0.37	-0.13	0
Fam6	-5.65	0.22	0.45	0.44	0.28	0
Fam5	-1.61	1.1	1.07	0.78	0.38	0
Fam4	-7.02	-0.54	-0.17	-0.01	0.04	0
Fam3	-4.66	-0.01	0.18	0.2	0.13	0
Fam1	-10.53	-3.7	-1.89	-0.92	-0.34	0

MARKER   D8S530 (FEB1)

	Recombination fraction, θ					
	0	0.1	0.2	0.3	0.4	0.5
Total	-50.65	-9.8	-4.59	-2.16	-0.86	0
Fam7	-8.1	0.03	0.22	0.14	0.01	0
Fam6	-10.02	-1.07	-0.5	-0.22	-0.06	0
Fam5	-12.01	-1.62	-0.66	-0.25	-0.06	0
Fam4	-2.78	-0.09	0.01	0	0	0
Fam3	-5.3	-3.67	-1.68	-0.8	-0.3	0
Fam1	-12.44	-3.39	-1.68	-0.8	-0.3	0

MARKER   D5S644 (FEB4)

	Recombination fraction, θ					
	0	0.1	0.2	0.3	0.4	0.5
Total	-45.62	-5.63	-1.76	-0.23	0.21	0
Fam7	-9.46	-2.38	-1.14	-0.52	-0.17	0
Fam6	-13.61	-0.78	0.03	0.24	0.18	0
Fam5	-13.61	-0.78	0.03	0.24	0.18	0
Fam4	-5.11	-0.66	-0.34	-0.19	-0.1	0
Fam3	0.08	0.16	0.16	0.12	0.07	0
Fam1	-9.85	-1.71	-0.59	-0.09	0.07	0

MARKER   D19S425 (SCN1B)

	Recombination fraction, θ					
	0	0.1	0.2	0.3	0.4	0.5
Total	-34.07	-7.07	-3.07	-1.22	-0.32	0
Fam7	-6.19	-1	-0.37	-0.12	-0.02	0
Fam6	-5.91	-1	-0.36	-0.08	0.02	0
Fam5	-6.52	-0.94	-0.35	-0.12	-0.02	0
Fam4	-4.22	-2.83	-1.52	-0.79	-0.32	0
Fam3	-4.43	-2.83	-1.52	-0.79	-0.32	0
Fam1	-6.81	-1.15	-0.44	-0.11	0.01	0

Table 5.4. Two-point lod scores at FS (FEB) and GEFS+ genes in families 1-7. Penetrance= 0.96, phenocopy rate= 0

111
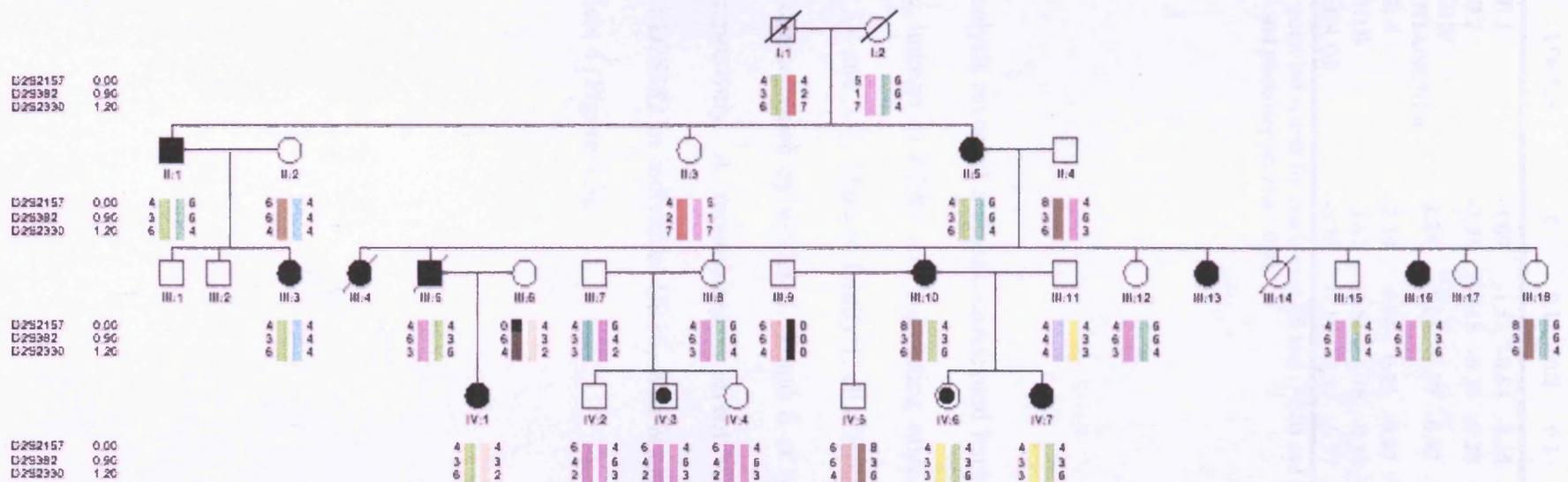
Figure 5.3. Pedigree of family 1. Filled symbols indicate affected individuals with GEFS+, symbols with dots indicate individuals with FS. Clear symbols represent unaffected individuals. Question mark (?) indicates phenotype unknown. Haplotypes were inferred for I:1, I:2, III,5 and III:9 since no sample was available from them. Haplotypes at FEB3 locus are shown.

MARKER	LOCUS	Recombination fraction, θ					Z Max	Theta Max
		0	0.1	0.2	0.3	0.4		
D8S530	FEB 1	-4.09	-1.53	-0.64	-0.25	-0.08	0.01	0.57
D19S177	FEB 2	-2.01	-0.68	-0.34	-0.20	-0.11	0.58	0.9
D2S382	**FEB 3/ SCN1A/SCN2A**	**2.24**	1.83	1.39	0.92	0.44	2.24	0.0
D5S644	FEB 4	-2.19	0.04	0.42	0.47	0.32	0.48	0.27
D19S425	SCN1B	-3.62	-1.51	-0.76	-0.34	-0.11	0.01	0.55
D5S1403	GABR G2	-6.70	-2.59	-1.27	-0.57	-0.18	0.01	0.54

Table 5.5. Two-point lod scores for markers at FS loci (FEB) and GEFS+ genes in Family 1. Penetrance 0.96 and phenocopies rate= 0.03.

Haplotype analysis revealed a disease-associated haplotype for each Family 1, 6, and 7, supporting linkage to *FEB3* and suggesting allelic heterogeneity across families (Figures 5.3, 5.4 and 5.5). Thus, in family 1, all affected individuals, except IV:3 carry a haplotype characterized by alleles 4, 3 and 6 at markers D2S2157, D2S382 and D2S2330, respectively. A recombination event was observed between markers D2S2157 and D2S382 in individual III:15, who is unaffected but carries the disease-associated allele 4 (Figure 5.3).
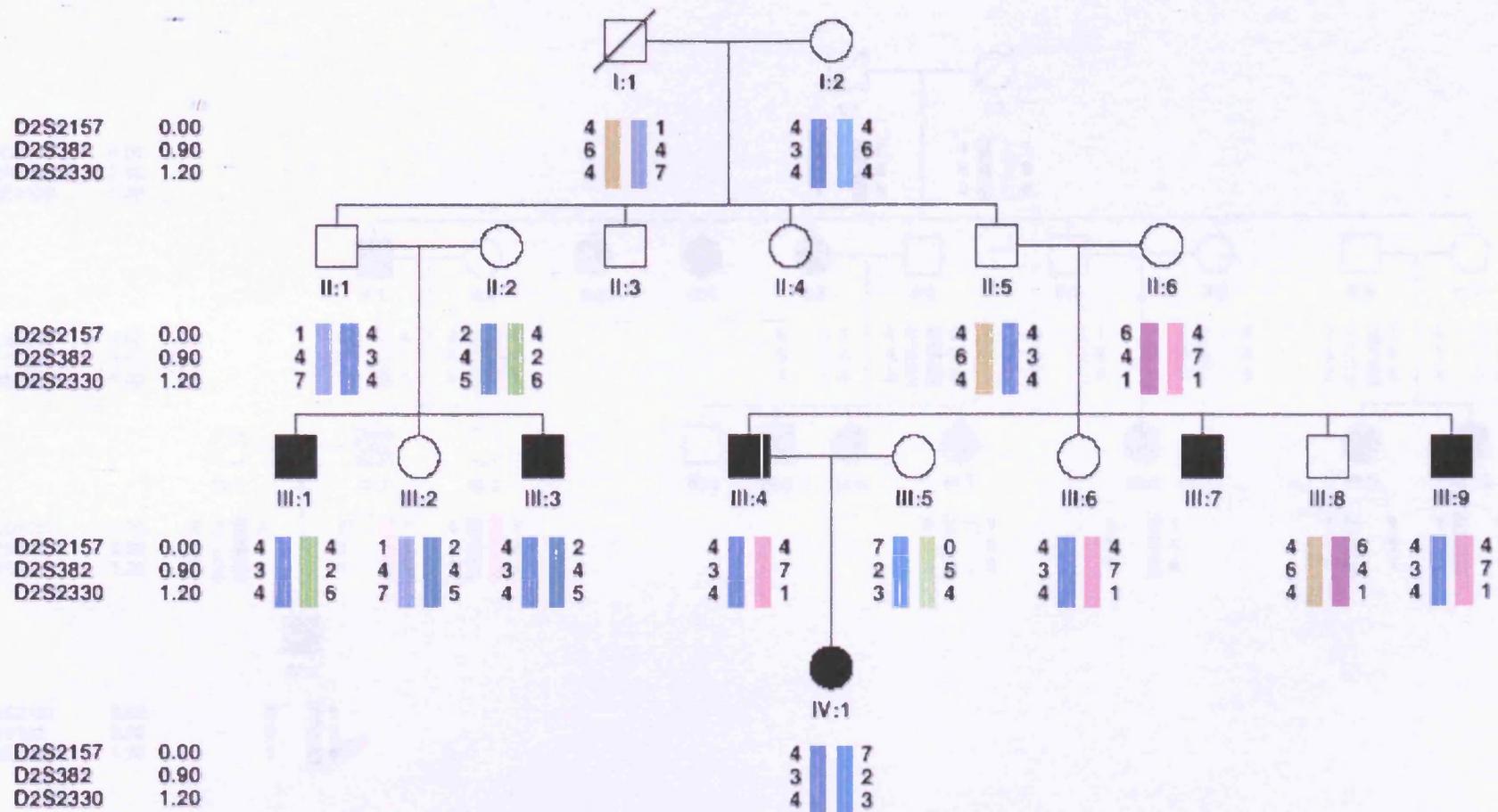
Figure 5.4. Pedigree of family 6. Filled symbols indicate affected individuals with GEFS+. Clear symbols represent unaffected individuals. Haplotypes were inferred for individual I:1 since no sample was available from him.
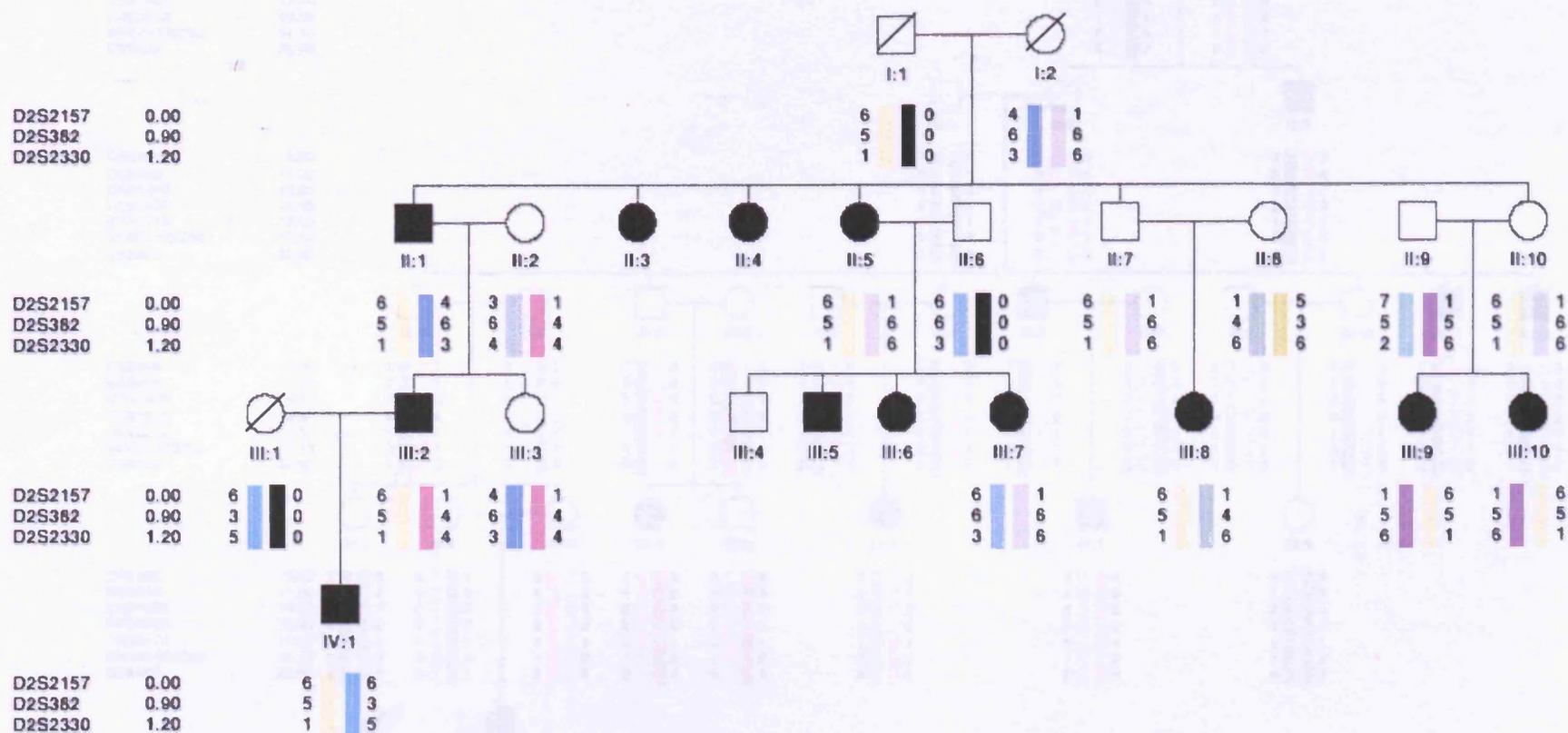
Figure 5.5. Pedigree of family 7. Filled symbols indicate affected individuals with GEFS+. Clear symbols represent unaffected individuals. Haplotypes were inferred for individuals I:1, I:2, II:6, III:1 since no sample was available from them. Individual III:7 corresponds to a phenocopy.

Marker panel (repeated for each individual block):

Marker	Position
D5S436	0.00
D5S410	11.30
D5S1403	14.90
D5S1955	15.90
D5S2118	16.00
SHGC-132760	17.10
D5S422	22.90

Individuals: I:1, I:2, II:1, II:2, II:3, III:1, III:2, III:3, III:4, III:5, III:6, III:7, III:8, III:9, III:10, III:11, III:12, IV:1, IV:2, IV:3, IV:4, IV:5, IV:6, IV:7, IV:8, V:1
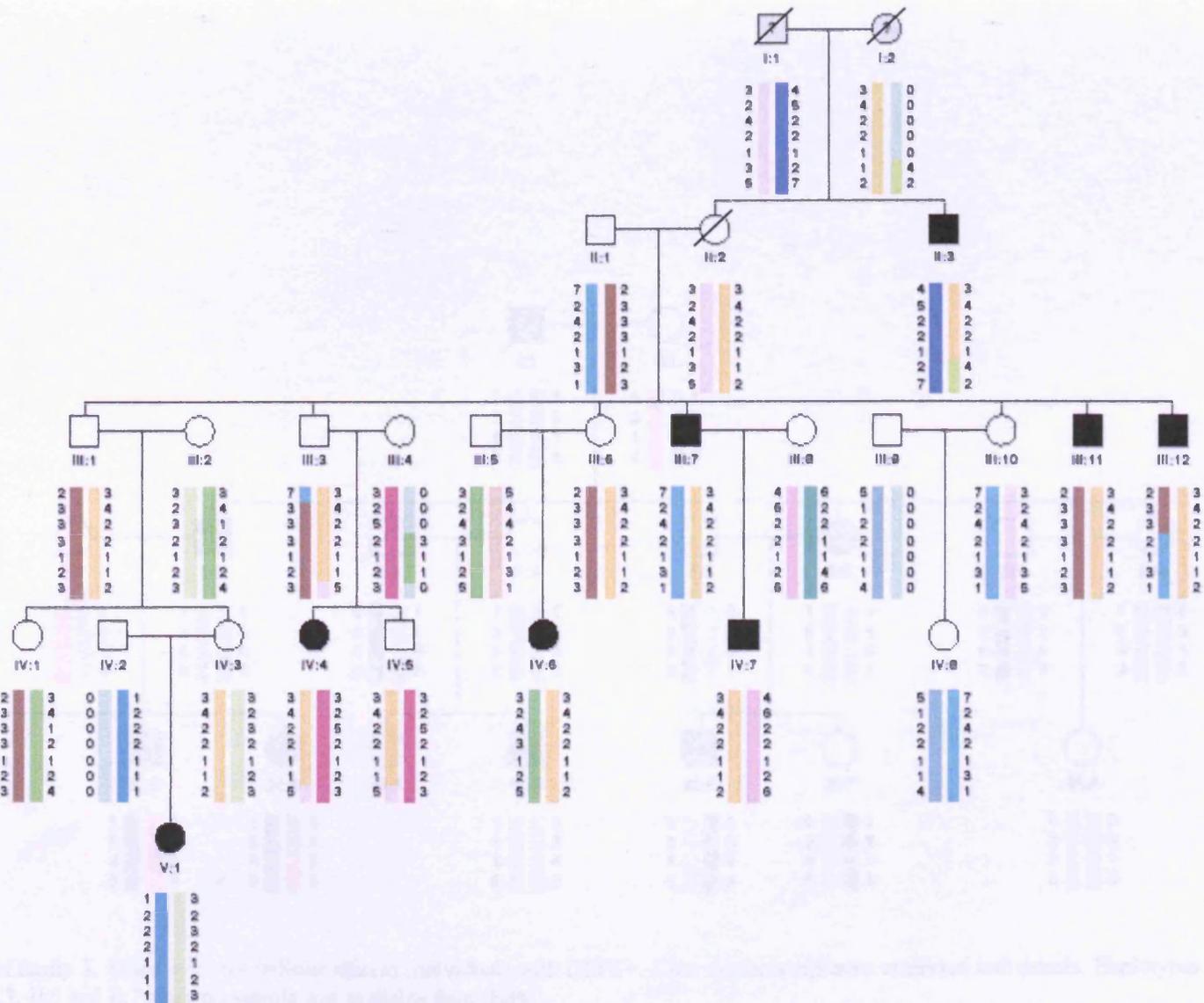
Figure 5.6. Pedigree of family 5. Filled symbols indicate affected individuals with GEFS+. Clear symbols represent unaffected individuals. Haplotypes were inferred for individuals I:1, I:2, II:2, III:4, III:9 and IV:2 since no sample was available from them. Individual V:1 corresponds to a phenotocpy.
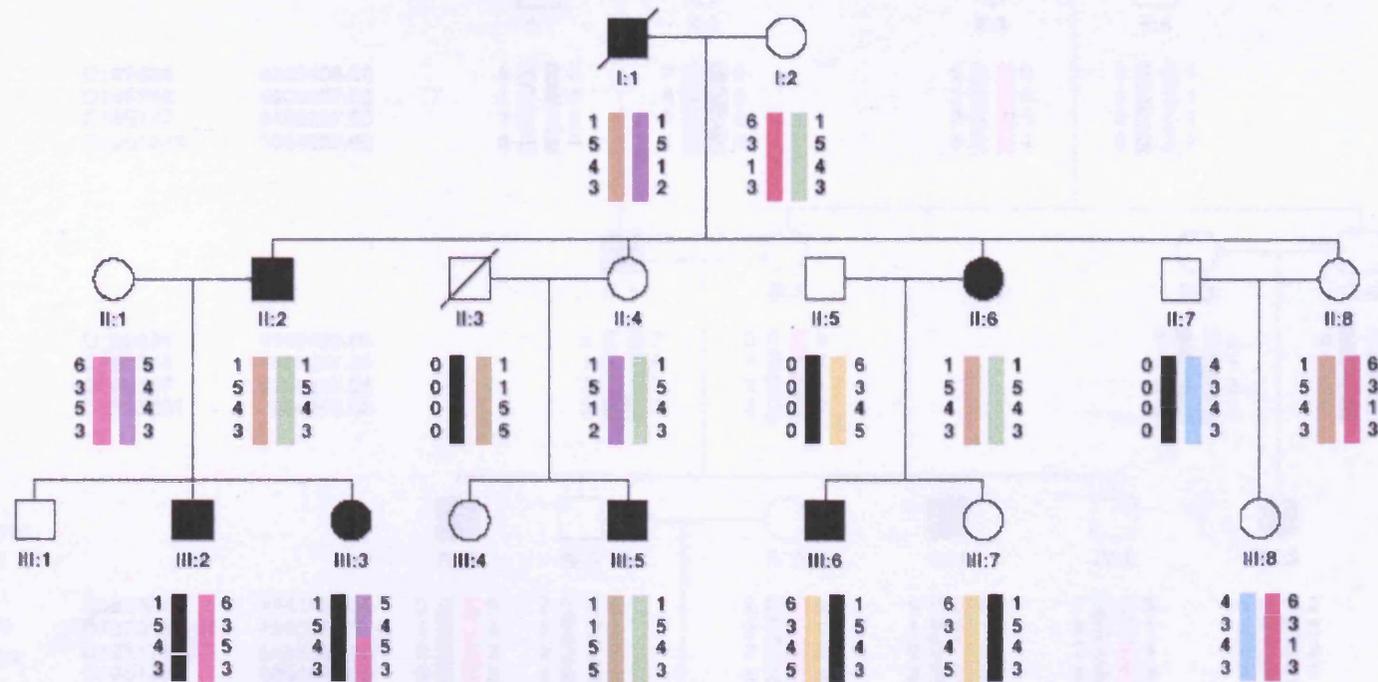
Figure 5.7. Pedigree of family 3. Filled symbols indicate affected individuals with GEFS+. Clear symbols represent unaffected individuals. Haplotypes were inferred in I:1, II:3, II:5 and II:7 since no sample was available from them.
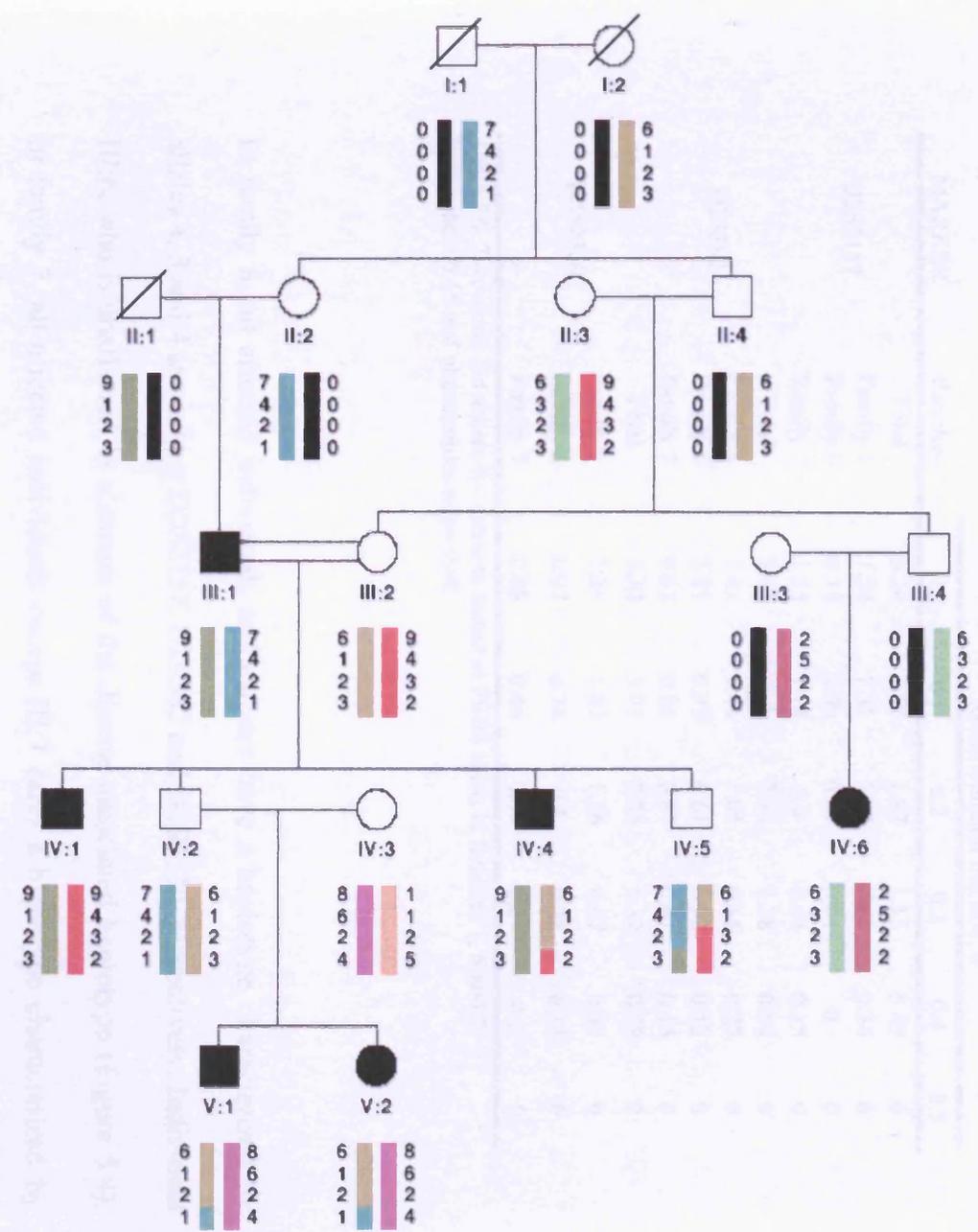
Figure 5.8. Pedigree of family 4. Filled symbols indicate affected individuals with GEFS+. Clear symbols represent unaffected individuals. Haplotypes were inferred for I:1, I:2, II:1, II:2, II:3, II:4, III:3 and III:4 since no sample was available from them

118

MARKER	Families	Recombination fraction, θ					
		0	0.1	0.2	0.3	0.4	0.5
D2S2157	Total	**3.24**	2.58	1.87	1.15	0.49	0
	Family 1	1.54	1.31	1.03	0.7	0.34	0
	Family 6	0.16	0.09	0.04	0.01	0	0
	Family 7	1.55	1.18	0.8	0.44	0.15	0
D2S382	Total	**3.63**	3.2	2.33	1.38	0.54	0
	Family 1	1.81	1.44	1.05	0.65	0.27	0
	Family 6	1.21	0.92	0.61	0.32	0.12	0
	Family 7	0.62	0.84	0.67	0.41	0.15	0
D2S2330	Total	**3.32**	3.03	2.25	1.39	0.59	0
	Family 1	2.26	1.83	1.36	0.87	0.38	0
	Family 6	0.97	0.74	0.5	0.28	0.11	0
	Family 7	0.08	0.46	0.39	0.24	0.1	0

Table 5.6. Two-point lod scores for markers tested at FEB3 locus in families 1, 6 and 7. Penetrance=0.65 and phenocopies rate= 0.02

In family 6, all affected individuals and carriers have a haplotype characterized by alleles 4, 3 and 4 at markers D2S2157, D2S382 and D2S2330, respectively. Individual III:6, who is unaffected, is a carrier of the disease-associated haplotype (Figure 5.4). In family 7, all affected individuals except III:7 carry a haplotype characterized by alleles 6, 5 and 1 at markers D2S2157, D2S382 and D2S2330, respectively (Figure 5.5).

GABR cluster

Linkage to *GABRG2* was suggested for Family 5 (Table 5.4, marker D5S1403). Significantly negative lod scores (<-2) were observed at θ= 0 for all markers tested, with exception of marker D5S1403. This locus showed a maximum Z=1.1 at θ= 0.1 (Z= -1.61 at θ= 0). When analyses were performed using penetrance 0.65 and phenocopies rate= 0.02, the lod score raised to 2.31 at marker D5S1403 at θ= 0 (Table 5.7). Evaluation of additional markers D5S436, D5S410 and D5S422 allowed the

identification of a disease-associated haplotype and also showed a recombinant involving the most telomeric marker, D5S422 (Figure 5.6). Thus there was in individual III:3 a recombination event between markers D5S1403 and D5S422 in the disease-associated chromosome causing the Z max to be obtained at a $\theta > 0$ (Table 5.7) The differences among the lod scores at $\theta = 0$ for the other three markers shown in Table 5.7 may be explained as follows. For marker D5S436 individual II:2 was inferred to be homozygous (3/3) (Figure 5.6) since all her children inherited the same allele and thus was non-informative. For marker D5S410, it is easy to see that allele 4 segregated with the disease. However, individual V:1, who is affected did not have allele 4, which in this type of analysis is interpreted as recombinant and therefore Z max was not at $\theta = 0$. The most likely interpretation is that individual V:1 is a phenocopy (see below). For marker D5S1403 all meioses were informative and no evidence of recombination was observed since every affected individual received at least one allele 2 (including individual V:1) (Figure 5.6).

| MARKER | Recombination fraction, θ | | | | | | θ max | Z max |
	0	0.1	0.2	0.3	0.4	0.5		
D5S436	1.07	0.84	0.62	0.42	0.21	0	0.001	1.07
D5S410	0.47	0.82	0.69	0.45	0.18	0	0.092	0.82
D5S1403	2.31	1.86	1.36	0.85	0.34	0	0.001	2.31
D5S422	0.44	0.8	0.68	0.46	0.2	0	0.096	0.8

Table 5.7. Two-point lod scores for markers tested at GABR locus in family 5. Penetrance=0.65 and phenocopy rate= 0.02

After testing three additional markers in the region (D5S1955, D5S2118 and SHGC-132760), a new recombinant was identified (Figure 5.6). Individual II:3 showed a recombination event involving marker SHGC-132760. Marker D5S2118 was monomorphic. All affected individuals carry a disease-associated haplotype characterized by alleles 3, 4, 2, 2 and 1 at markers D5S436, D5S410, D5S1403,

D5S1955 and D5S2118, respectively. However, the affected individual V:1 does not carry this haplotype, suggesting that she is a phenocopy. Zmax for markers D5S1955, D5S2118 and SHGC-132760 were 0.4 at θ= 0; 0.12 at θ= 0 and 1.29 at θ= 0.077, respectively.

FEB2 locus

Linkage to locus *FEB2* was suggested for families 3 and 4 (Table 5.4, marker D19S177). Evaluation of three additional markers (D19S894, D19S216 and D19S1034) in this region could not either confirm or reject linkage to FEB2 in Families 3 and 4. Lod scores were in the interval -2 < Z < 1 (Table 5.8). However, in Family 4 there was not a common haplotype segregating with the disease, suggesting exclusion of this locus. For example, individual III:1 transmitted his haplotype 9-1-2-3 to his two affected children IV:1 and IV:4 and transmitted his alternative allele to IV:2 who has two affected children. These two affected individuals (V:1 and V:2) have another haplotype characterized by alleles 6-1-2-1. Moreover, individual IV:6 has a different haplotype from the previous two (Figure 5.8).

Even so, in Family 3 all affected individuals carry a haplotype characterized by alleles 1, 5, 4 and 3 at markers D19S894, D19S216, D19S177 and D19S1034. Affected individuals II:2 and II:6 were homozygous for these haplotypes, so that, their meioses were not informative for linkage resulting in weak lod scores (Table 5.8 and Figure 5.7). The fact that I:1 and I:2 share a haplotype suggests that they both have a recent common ancestor. This also assumes that I:2 is a carrier who is non-penetrant and that this is also the case for II:4 and II:8 but is in agreement with linkage to FEB2 on a haplotype basis. Therefore, more markers need to be tested in this region to clearly determine whether this locus is linked to the disease in Family 3 or not.

MARKER		Recombination fraction, θ					
		0	0.1	0.2	0.3	0.4	0.5
D19S894	Total	**-0.14**	-0.26	-0.31	-0.26	-0.13	0
	Family 3	0.24	0.17	0.1	0.05	0.02	0
	Family 4	-0.37	-0.42	-0.42	-0.31	-0.15	0
D19S216	Total	**-1.53**	-0.6	-0.28	-0.09	0.01	0
	Family 3	-0.49	-0.45	-0.35	-0.21	-0.09	0
	Family 4	-1.04	-0.14	0.06	0.12	0.09	0
D19S177	Total	**0.63**	0.47	0.32	0.19	0.08	0
	Family 3	0.37	0.28	0.2	0.12	0.05	0
	Family 4	0.26	0.18	0.12	0.07	0.03	0
D19S1034	Total	**0.86**	0.66	0.46	0.27	0.11	0
	Family 3	0.38	0.27	0.18	0.09	0.03	0
	Family 4	0.48	0.39	0.28	0.17	0.08	0

Table 5.8. Two-point lod scores for markers tested at FEB2 locus in families 3 and 4. Penetrance=0.65 and phenocopies rate= 0.02.

3.4 Sequencing

Families 1, 6 and 7

Sequencing of *SCN1A* exons 4, 15, 18, 19, 21 and 26 was performed in two carriers of the disease-associated haplotype in Family 1 (II:6 and III:10). A heterozygous A to G substitution was identified in exon 26 at position 5213 of the *SCN1A* mRNA (c.5213 A>G, Figure 5.9). This base change destroys a recognition site for enzyme *Tsp45I*. Restriction enzyme digestion of the third fragment of exon 26 PCR products showed that the c.5213A>G base change is present in the individuals that were considered affected, except for IV 3, who presumably represents a FS phenocopy. This DNA mutation was not observed in 60 unaffected controls from Antioquia. The c.5213A>G results in an aspartate for glycine substitution at position 1742 of the protein (D1742G). This substitution lies in the evolutionarily conserved segments S5 and S6 of domain IV in the SCN1A channel (Figure 5.10).

This mutation is not present in either family 6 or 7. SSCP analysis for the same *SCN1A* exons as evaluated by sequencing in Family 1, did not reveal any conformational polymorphism in either Family 6 or 7.

(A)

(B)

III:8   II:5   III:10   III:16   IV:17   III:6   IV:5   III:12

Figure 5.9. Mutation screening of exon 26 of SCN1A. (A) Partial nucleotide sequence of the exon showing the heterozygous A>G substitution indicated with letter R. (B) Restriction digestion of exon 26 with *Tsp45I*. The mutation c.5213A>G destroys the restriction site. Labels of the lanes to the individuals examined as in figure 3.

```
SCN1A_HUMAN    T F G N S M I C L F Q I T T S A G W D G L L A P I L N S K P P D C  D  P N K V N P G S S V K
SCN1A_RAT      . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  .  . . . . . . . . . .
SCN2A_HUMAN    . . . . . . . . . . . . . . . . . . . . . . . . . . . . G . . . . .  .  . D . D H . . . . . .
SCN1A_HUMAN    . . . . . . . . . . . . . . . . . . . . . . . . . . . . A . . . . .  .  . D T I H . . . . . .
SCN9A_HUMAN    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .  .  . K . . H . . . . . E
SCN4A_HUMAN    . . . . I . . . E . . . . . . . . . . . . . . . N . . . . G . . . .  .  . . L E . . . T . . .
SCN5A_HUMAN    . . A . . L . . . . . . . . . . . . . . . . . . S . . . T G . . Y .  .  . T L P . S - N G S R
SCN11A_HUMAN   . . A S . . L . . . . S . . . . . S . . S . M . R . . - - - -  -  - - - - E S C N . S S
SCN1B_HUMAN    I . - - - - - - - I T N - - V T Y N - - - - - - - H . G - - - -  -  - - - - - - - - - - -
```

Figure 5.10. Amino acid sequence alignment of the loop between segments S5 and S6 in domain D-IV of human neuronal sodium channels. The aspartate residue (D) that is substituted by glycine (G) is boxed.

123

Family 5

Sequencing of all nine *GABRG2* exons in Family 5 revealed two nucleotide changes. One of these DNA changes was found on exon1 and consists of an adenine deletion in position +286 of the *GABRG2* mRNA, three bases before the start codon. This variation lies in a tract of twelve adenines and was present in both affected and unaffected individuals. The second variation was found on exon 5 and consisted of a substitution of C to T in position 540 of the cDNA and had previously been reported as a frequent polymorphism [243]. This observation suggested that a mutation in either an intron or the promoter region went undetected. There was also a chance that the mutant gene was other than *GABRG2* (e.g. *GABRA1*, *GABRA6* or *GABRB2*). Evaluation of markers D5S1955, D5S2118 and SHGC-132760, which lie in between markers D5S1403 and D5S422 (see Figure 2), revealed an additional recombinant. Affected individual II:3 presented disruption of the disease-associated haplotype since a recombination event between markers D5S2118 and SHGC-132760 was observed (Figure 5.6). SHGC-132760 lies in between contiguous genes *GABRG2* and *GABRA1* (Figure 5.2)

4. Discussion

The purpose of this chapter was to characterize six Colombian families with GEFS+ initially by evaluating linkage to four *FEB* loci (*FEB*1-3 and *MASS1* gene) and four *GEFS+* genes (*SCN1A*, *SCN2A*, *SCN1B* and *GABRG2*) previously reported. Most of GEFS+ mutations reported frequently are in the *SCN1A* gene. So far, twelve *SCN1A* mutations leading to GEFSP2 have been reported [220-222, 232, 233, 235]. Recently, mutations in SCN1A have also been associated with SMEI [236, 237]. Most of which are

124

located in the functionally critical pore forming of SCN1A (S4-S6) and could therefore have a major impact in the kinetic properties of this sodium channel [237]. Functional studies SCN1A mutations have shown disruption of channel inactivation, resulting in increased Na+ influx and leading to hyperexitability and suggesting a gain of function [240].

Suggestive linkage was found for at least four of the families studied here. Three of these (Families 1, 6 and 7) were linked to markers on 2q23-q24, the region where *FEB3*, *SCN1A* and *SCN2A* map. The fourth family was linked to *GABRG2* (Family 5). Families linked to 2q23-q24 each presented a disease-associated haplotype. Sequencing of *SCN1A* exons in Family 1 revealed a c. 5213 A>G mutation, which lead to the D1742G aminoacid substitution. This substitution lies in the pore forming region of domain IV. The high penetrance of this mutation (100%) and the severity of the clinical manifestations might be explained by the functional importance of its location within the protein. The high penetrance contrasts with the characteristic medium penetrance (65%) reported for GEFS+ [85, 215], where normally a rate of non-penetrants is observed.

Since this mutation was not present in the other 2q23-q24 linked families, allelic heterogeneity of this disorder in Antioquia is indicated.

Mutations in *GABRG2* have also been reported for both GEFS+ and SMEI [243-246]. Three GEFS+ mutations have been reported in GABRG2 and only one for SMEI. This supports a clinical relationship between GEFS+ and SMEI, in agreement with the suggestion that SMEI constitutes the most severe form of the GEFS+ clinical spectrum [246]. Sequencing of all *GABRG2* in two affected individuals from Family 5 did not reveal a mutation. Identification of a recombinant for markers SHGC-132760

confirmed exclusion of *GABRG2* (Figure 5.2 and Figure 5.6) and therefore it is expected that either *GABRA1*, *GABRA6* or *GABRB2* is the mutant gene in this family.

Families 3 and 4 seem not to show linkage to the genes/loci tested. Initially, linkage to the region *FEB2* could not be excluded but once three additional linked markers were tested, linkage was excluded in Family 4. In contrast, all affected individuals in Family 3 carried the same haplotype. However, as individuals I:1 and I:2 are carriers of the same haplotype it is possible that affected individuals carry the same haplotype but from different origin (i.e. paternal or maternal) (Figure 5.6). The fact that two of the affected individuals were homozygous for these haplotypes, meant that these meioses were uninformative and therefore the lod score analysis did not support any decision in favour of or against linkage. If after testing more markers in this region and if exclusion was obtained, Families 3 & 4 and would be candidates for a genome wide scan, aimed at identifying new GEFS+ genes.

In conclusion, four of six families studied here are linked to FEB3 locus and GABR cluster. Family 1 carries D1742G mutation located in the pore forming region of domain IV of *SCN1A*. Families 6 and 7 are currently been further evaluated in SCN1A. Family 5 is linked to the GABR cluster and we have found evidence of exclusion of GABRG2 gene, only gene in the cluster previously associated with GEFS+ [224]. Sequencing of the other GABR genes in this region will reveal which of these is bearing the mutation in this family and will constitute the first genetic evidence of that gene as implicated in GEFS+. Family 4 is not linked to any of the genes/loci evaluated. Family 3 still might be linked to FEB2, where at the moment there is not an evident candidate gene to look at.

The heterogeneity evidenced here allows me to conclude that GEFS+ is a highly clinically and genetically heterogeneous trait, even in families from isolate populations, such as Antioquia.

CHAPTER SIX

LINKAGE ANALYSIS AND

MUTATION SCREENING IN A

FAMILY WITH ATYPICAL BATTEN

DISEASE

1. Introduction

1.1 Generalities

The Neuronal Ceroid Lipofuscinoses (NCLs) are a group of inherited neurodegenerative disorders, characterized by the accumulation of autofluorescent lypopigments (Ceroid and Lipofuscin) in neurons and other cell types within the lysosomes [255, 256] that belong to the group of progressive myoclonus epilepsies (PMEs) [257]. NCLs as a group are the commonest neurodegenerative disorder in childhood. Its incidence has been estimated as over 1 in 25000 live births [258], with an increased prevalence in the populations of north Europe. Common clinical features to all NCLs are progressive cognitive, motor and visual decline, and seizures. Seizures can be generalised, tonic-clonic, myoclonic, astatic or atonic [257].

At least ten NCL subtypes have been identified. Diagnostic criteria are based on age of onset, clinicopathological findings and genetics. Storage material is composed of neutral lipids and phospholipids, and to a lesser extent glycolipids with a major fraction of autofluorescent insoluble proteins.

The major storage components in infantile NCL are two types of Sphingolipid Activator Proteins (SAPs) or saposins (Table 6.1) [259]. SAPs are small, heat stable, lysosomal proteins, originating from a single precursor molecule that, by cleavage, releases four mature SAPs (A-D) [260]. SAPs activate lysosomal hydrolases involved in glycosphingolipid degradation and the glycosphingolipid substrate. SAPs A and D accumulate in infantile NCL [261]. The major storage component in late-infantile, juvenile, and some adult NCLs is subunit c of the mitochondrial ATP synthase complex (Table 6.1) [262, 263]. Subunit c is an essential membrane component of the

129

proton channel of the large oligomeric complex ATP synthase, which generates ATP by the process of oxidative phosphorylation. It is known that indigestible material remains within lysosomes, and as exocytosis in neurons is poor [264]the formation of electron-dense residual bodies is favoured.

The onset of NCLs is usually in childhood. Childhood forms are autosomal recessive, while the rare adult forms can be either autosomal recessive or dominant [265, 266] (Table 6.1). All NCL types were known as Batten disease until 1990s when different genetic loci were identified, but this is still the best-known name for the juvenile type.

1.2 Molecular genetics of NCLs

There are three major clinical forms of NCL: Infantile (INCL), Late infantile (LINCL) and Juvenile or Batten (JNCL). Six genes have been identified that underly these forms: INCL (*CLN1*) [267]}, LINCL (*CLN2*) [268]} and JNCL (*CLN3*) [191]. Genes underlying two variant late infantile and one juvenile forms caused by mutations in *CLN5*, *CLN6* and *CLN8*, respectively, have recently been identified [197, 269-271] (Table 6.1).

Some of the CLN genes have been associated with different clinical forms. Thus, *CLN1* mutations have been associated with INCL, LINCL and JNCL [272, 273]. Similarly, *CLN2* mutations have been associated with both LINCL and JNCL [268, 274, 275].

CLN1

130

CLN1 is located on chromosome 1p32, has 9 exons and extends 24.55Kb [267]. *CLN1* encodes the enzyme palmitoyl protein thiosesterase 1 (PPT1), which removes palmitate residues from proteins. In non-neuronal cells it is transported to the lysosome via the mannose-6-receptor-mediated pathway, and is also secreted [276, 277]. Even though, PPT1 does not seem to localize to lysosome in mature neurons, it is preferentially targeted to axons [278]. Despite having no elevated expression in any particular human tissue type [279], its association with the disease is based on its increased expression during the development of neurons from cortical neurogenesis through cortical development [280].

Forty-three mutations have been reported so far, which are distributed through out the gene. *CLN1* mutations include 20 missense, 9 nonsense, 10 small deletions or insertions and four mutation affecting splice sites (reviewed in [281]).

NCL type	Ultrastructure	gene	Chromosomal region	Major storage material	Age of onset	References
Childhood						
Infantile	GROD	*CLN1*	1p32	SAPs	6-24 months	[267]
Late Infantile						
Classical	CVB	*CLN2*	11p15	Subunit c	2-4 years	[274, 275]
Finnish	FPP/RL	*CLN5*	13p22			[269, 282]
Variant	FPP/RL	*CLN6*	15q21-q23			[270, 271]
Variant	FPP/RL	*CLN7*	ND			
Variant	GROD+/-RL	*CLN1*	1p32			[272, 273]
Juvenile						
Classical	FPP	*CLN3*	16p12	Subunit c	4-10 years	[191]
Progressive	RL	*CLN8*	8p23			[197]
Variant	GROD	*CLN1*	1p32			[283]
Variant	CL	*CLN2*	11p15			[284, 285]
Adulthood						
Kufs disease/ Parry disease	FPP/RL or GROD	*CLN4*	ND	Subunit c	avg. 30 years	[266]
		CLN1	1p32			[265]

Table 6.1. Molecular Genetics of the NCLs.
GROD: Granular osmiophilic deposits; CVB: curvilinear profiles; FPP: fingerprint profiles; ND: Not yet determined. (Modified from [286])

CLN2

CLN2 is located on chromosomal region 11p15, has 13 exons and extends 6.66 Kb [275].

CLN2 encodes the lysosomal enzyme tripeptidyl peptidase (TPP1), which removes tripeptides from the N terminus of small proteins such as subunit c of mitochondrial ATP synthase that accumulates in most of forms of NCL [287, 288]. TPP1 is produced as a glycosylated precursor polypeptide that is processed to the mature enzyme in the lysosome [289]. Enzyme activity occurs in the brain by the age of two years in the

cerebral cortex [290]. Enzyme assay of TPP1 enables accurate diagnosis of NCL caused by CLN2 mutations [291].

Fifty-two mutations have been reported in CLN2, which include 7 small deletions, 2 small insertions, 29 missense mutations, 5 nonsense mutations and 9 mutations affecting splice sites or intronic sequences (reviewed in [281]).

CLN3

CLN3 is located on chromosomal region 16p12, has 15 exons and extends 15 kb [191]. CLN3 encodes a 438 aminoacid glycosilated membrane protein that is believed to localize in the lysosome and also in other additional locations in neurones [292-294]. Although its cellular function is still unknown, it has been implicated in vacuolar homeostasis in yeast [279, 295]. Interestingly, mice models show a preferential loss of GABAergic neurons [296, 297]. Consistently inhibition of GAD by anti-GAD antibody has been shown in cln3-knockout mice serum that associates with brain tissue but is not present in sera or brain of normal mice [298].

Thirty-eight mutations have been reported in CLN3, which include 5 small deletions, 5 small insertions, 4 large deletions, 9 missense mutations, 8 nonsense mutations, 5 mutations affecting splice sites and 1 intronic mutation (reviewed in [281]). A 1.02 kb deletion that is present in about 85% of the disease chromosomes is the most common mutation in CLN3 [191, 299].

Clinical Features of Classical Juvenile (Batten Disease)

For the classical juvenile type (Batten), there are very well characterized clinical and pathological features. Clinical symptoms begin with vision failure between five and ten years of age. It continues with convulsions and progressive deterioration of the cognitive functions until death, in the second or third decade of life. Diagnostic

criteria include the presence of inclusions in many cell types that resemble fingerprints in the electronic microscope [300], macular degeneration, presence of vacuolated lymphocytes in peripheral blood [256].

CLN5

CLN5 is located on chromosomal region 13p22, has 4 exons and extends 13 Kb [269]. *CLN5* encodes a 407 aminoacid glycosylated protein, with an expression increase during cortical neurogenesis [280]. This protein is synthesized as a four precursor forms, which are thought to be lysosomal soluble [301]. Interaction of *CLN5* with *CLN2* and *CLN3* proteins has been reported. All *CLN5* mutants have lost their ability to interact with *CLN2* while keeping their ability to interact with *CLN3*, according to coimmunoprecipitation assays [302].

Four mutations have been reported for *CLN5*. Three of these mutations result in a truncated protein and the fourth is a missense mutation. To date, *CLN5* mutations have been identified in Finnish and Swedish families and in one Dutch patient [269, 303]. All these patients presented a late infantile CLN type.

CLN6

CLN6 is located on chromosomal region 15q21-q23. It has 7 exons and spans 22.69 Kb [270, 271]. It is predicted that CLN6 encodes a 311 aminoacid protein with seven hypothetical transmembrane domains which, similar to CLN3 and CLN5, has no homology with known proteins or recognized functional domains [270, 271].

Eighteen mutations have been reported, including missense, frameshift, splice site and nonsense mutations (reviewed in [281].

CLN8

CLN8 locates at chromosomal region 8p23, has three exons and extends 20.5 Kb [197].

CLN8 is a 286 aminoacids protein that locates in the endoplasmic reticulum (ER) and also shuttles between the ER and the ER-golgi intermediate complex [304]. Five mutations have been reported so far. Four of which lead to aminoacid substitutions, and the fifth mutation (frameshift) is predicted to result in a truncated CLN8 protein (Reviewed in [281]).

Different loci explain different clinical forms of the disease. Also different clinical forms may have their cause in a same gene with its different mutations (Table 6.1). Furthermore, some patients for whom mutation screening of known NCL genes have proved negative results, in addition to animal models have shown that there may be other genes implicated in this disorder. Studying new collections of patients of homogeneous origin or analyzing pedigrees with enough power will result in identification of other until now uncovered genes related with the disease. It is then clear that clinical as well as genetic heterogeneity is common feature in this group of disorders.

1.3 Preliminary work and aims

Little is known about the genetic aspects underlying the group of NCLs in Colombia.

A nuclear family including two siblings (VI:2 and VI:3; Figure 6.1) with a clinical diagnosis of Batten was identified in northwest Colombia. Genealogic inquiries revealed consanguinity between the unaffected parents. Genealogy extension revealed a third affected child in another consanguineous branch of the family who died at age 14 years (Figure 6.1).
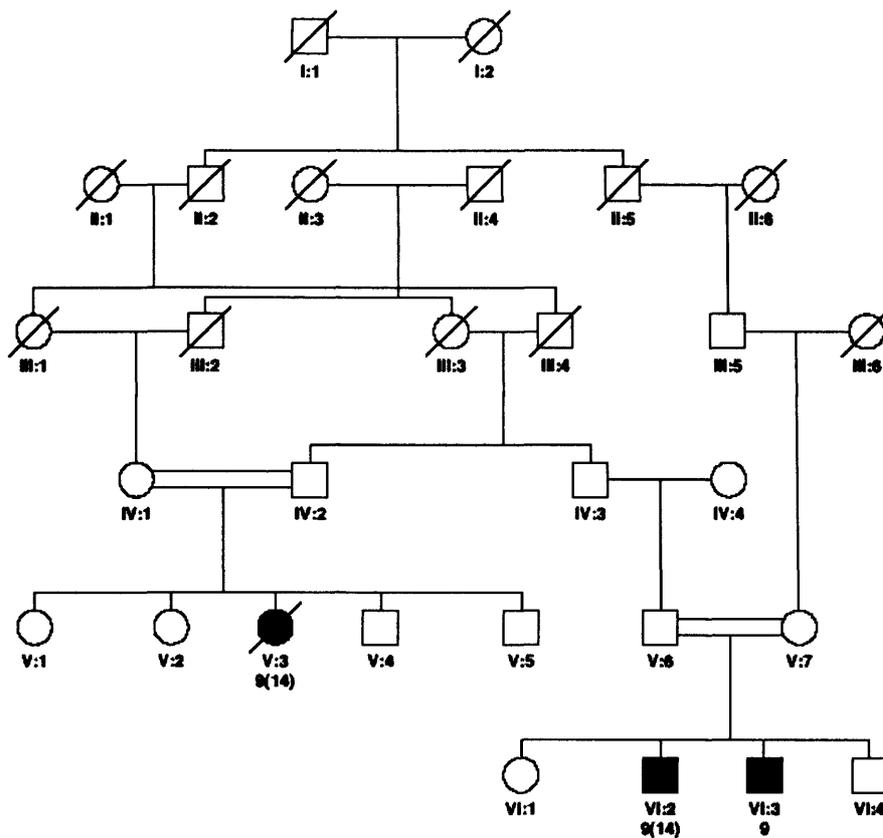
Figure 6.1. Pedigree of the Colombian family with Batten disease studied here.
Two affected siblings in one nuclear family (VI:2 and VI:3) and one affected individual (V:3) in another nuclear family from the same pedigree are shown. Individual VI:2 died during the course of this study. Age at onset (age at death) in affected individuals is shown. Available individuals were all those alive with exception of individuals IV:3, IV:4 and V:6. This was taken into consideration for the power simulation.

Based on the consanguineous nature of the pedigree, a homozygosity search was proposed aiming to identify homozygosity only in the available patients and not in the unaffected relatives.

The population of northwest Colombia is of mixed Spanish, Native American, and African ancestry [32], and no northern European ancestry was documented in this pedigree.

Symptoms started at 9 years of age with visual failure, loss of strength, and tremor of lower limbs. The patients deteriorated rapidly, with blindness and inability to walk occurring within 1 year of the initiation of symptoms. A few months after onset, the

136

patients also had behavioural changes, a gradual loss of language, myoclonus, and seizures.

Evaluation of markers linked to *CLN3* showed that the affected individuals were heterozygous for the same alleles (4/5 and 2/3) at markers D16S298 D16S299, respectively (Figure 6.3), providing lod scores of 0f -1.3 at θ= 0 for both markers (Table 3). However, the affected individuals had the same genotypes for both markers while none of their siblings (Figure 6.3) had these genotypes, making more difficult to exclude *CLN3*.

Given the large number of *Alu* sequences within *CLN3* [305], recombination within the gene was a possible explanation for the heterozygosity observed in the patients. Also, although compound heterozygotes were not expected in this pedigree because of consanguinity, it seemed that two different *CLN3* mutations could be segregating in the family. Consequently, SSCP for all 15 *CLN3* exons and sequencing of exons 11 and 13 was carried out. No genetic variants were found, supporting the idea that any other locus, possibly previously unreported, was involved in this pedigree.

This chapter describes further studies on this family.

The aim of this chapter was to characterize an extended Colombian pedigree presenting with clinically Batten disease (Juvenile type).

2. Methods

2.1 Power simulation

Power simulation was performed assuming autosomal recessive inheritance mode and full penetrance. A mutant allele frequency of 0.0001 was assumed. Also, a marker locus with four equally frequent alleles was simulated in 100 replicas [55]. The actual simulated pedigree corresponds to that shown in Figure 6.1.

2.2 Candidate genes.

Linkage to known NCL genes was evaluated using tightly linked microsatellite markers as shown in Table 6.2. Marker typing was performed as indicated in the section on microsatellite marker typing in chapter one. PCR amplification with reported primers was carried out with 30 cycles of 94° for 30 seconds, 53° to 62° C as annealing temperature, and 72° C for 30 seconds (Table 6.2).

Gene	Marker	Annealing Temperature
CLN1	D1S193	55
CLN2	D11S1338	55
CLN3	D16S298	62
	D16S299	53
	D13S156	55
CLN5	D13S162	50
	D13S1306	55
CLN6	D15S1000	55
	D15S983	56
CLN8	AFM265wb5	55
	44CA	60

Table 6.2. Markers tested for evaluating known NCL genes.

2.3 Linkage and Haplotype analysis

Linkage analysis was performed assuming an autosomal recessive inheritance mode and full penetrance as in power simulation. Allele frequencies were assumed equal to 1/n, n= number of alleles. Haplotypes were reconstructed using simwalk2. Refer to the methods chapter.

2.4 Sequencing

Full sequencing was conducted for the exons of *CLN5* and *CLN8* genes. Refer to the methods chapter.

2.5 Screening for mutation R112H

A sequence change identified in exon 2 of *CLN5* destroys a restriction site for *PshAI*. Screening of this mutation was carried out by digestion of PCR products with 20 units of *PshAI* (New England Biolabs, Beverly, MA) at 37° C for 18 hours. Restriction fragments were resolved in a 3% agarose gel stained with ethidium bromide (0.5 mg/mL) (Figure 6.4B).

3. Results

3.1 Pathology

Electron microscopy of skin biopsy from patient VI:2 revealed inclusions consisting of fingerprint profiles, often condensed, and occasionally associated with lipid droplets, suggestive of a variant NCL (Figure 6.2). No curvilinear or rectilinear profiles indicative of a mixed histopathology were seen (Dr Anderson, Personal communication; Histopathology Department, UCL).
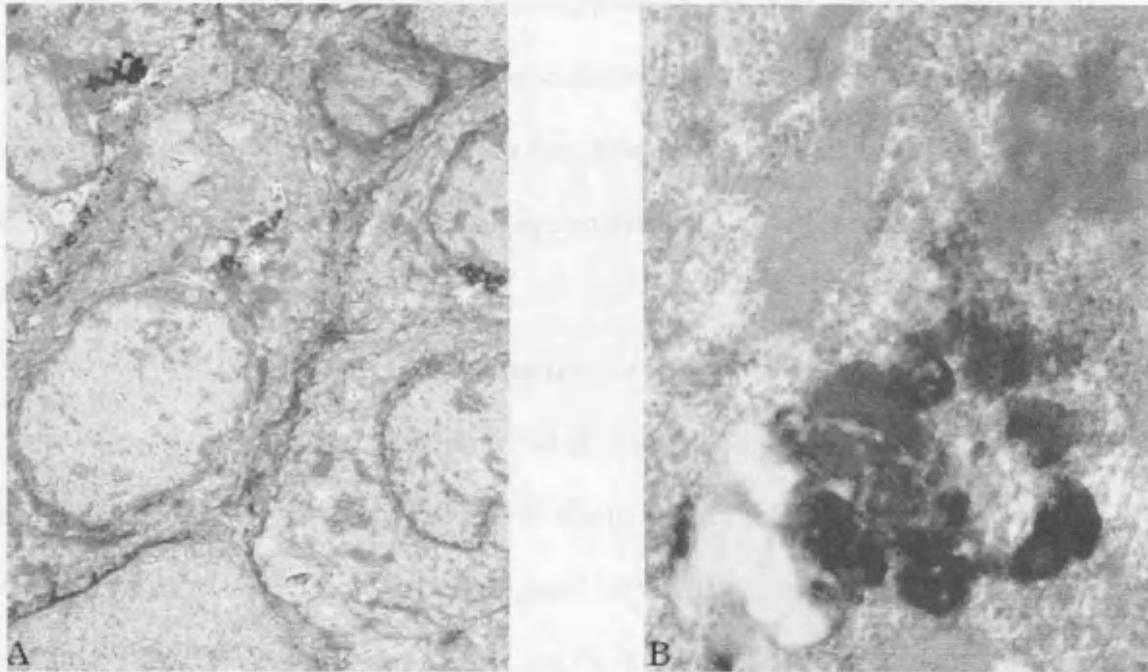
Figure 6.2. Electron microscopy of skin biopsy from patient IV:3. (A) White asterisk indicates storage inclusions in sweat gland epithelial cells. (B) Higher magnification of one inclusion showing a fingerprint pattern, often condensed and occasionally associated with a lipid droplet.

3.2 Power simulation

Table 6.3 shows the power simulation results. These indicate that on average a linked locus will provide a lod score over 2 in this pedigree. The pedigree structure is much more powerful but unavailability of samples from some individuals limited the power.

Expected Lod score (Z) at $\theta=0$		Probability Z$\geq$		
Average	Max	1	2	3
2.22	2.93	96	71	0

Table 6.3. Power simulation for Colombian family with clinical NCL as shown in Figure 2.

3.3 Linkage analysis

Table 6.4 shows the lod scores for markers linked to *CLN1, CLN2, CLN3, CLN6* and *CLN8*. Analysis was performed with the data from the pedigree shown in Figure 6.3.

140

Marker D1S193 (*CLN1*) was found heterozygous in the two affected individuals (alleles 2/3, Figure 6.3). This was a similar situation to that found for markers linked to *CLN3*, where patients were heterozygous for same alleles and none of their siblings had the same genotype as they did. Linkage analysis resulted in a negative lod score for this marker (Table 6.4).

Marker D11S1338 (*CLN2*) was found heterozygous in the two affected individuals and as clear evidence of exclusion, they both had different genotypes (1/2 for VI:2 and 2/3 for VI:3, Figure 3). Z= - infinity at theta= 0 was obtained (Table 6.4).

Two markers linked to *CLN6* were evaluated. D15S1000 was heterozygous same alleles (2/3) in both affected individuals. And marker D15S983 was homozygous (1/1) in one unaffected individual apart from the two affected children (Figure 6.3). Lod score was negative for D15S1000 and Z= 0.5 for D15S983 was obtained at theta= 0 (Table 6.4).

D16S298
D16S299
D15S1000
D15S983
AFM265wb5
44CA

		V:6	V:7
	D16S298		
	D16S299		
CLN1	D1S193D15S1000		0 0
CLN2	D11S13D15S983		1 3
CLN3	D16S29AFM265wb5		2 5
	D16S2944CA		3 3
CLN6	D15S1000		0 0
	D15S983		0 0
CLN8	AFM265wb5		1 2
	44CA		0 0

		VI:1		VI:2 9(14)		VI:3 9		VI:4	
CLN1	D1S193	1	2	2	3	2	3	1	2
CLN2	D11S1338	1	2	1	2	2	3	1	1
CLN3	D16S298	1	5	4	5	4	5	1	2
	D16S299	1	3	2	3	2	3	1	3
CLN6	D15S1000	1	3	2	3	2	3	1	3
	D15S983	1	1	1	1	1	1	1	2
CLN8	AFM265wb5	1	2	2	2	2	2	1	2
	44CA	1	3	2	3	2	3	1	3

Figure 6.3. Nuclear family showing genotypes for markers linked to five NCL genes.
* Identification of individuals is as in Figure 1. Individual VI:2 died during the course of this study. Age at diagnosis (age at dead) is shown.

141

Two markers linked to *CLN8* were evaluated. AFM265wb5 was homozygous only in the affected individuals (2/2) (Figure 6.3). This was consistent with linkage to the disease and the lod score now raised to 1.17 at theta= 0 (Table 6.4). Evaluation of a second linked marker, 44CA, revealed heterozygosity in the affected children for the same alleles (2/3). Lod score was again negative (-1.6 at θ= 0).

NCL gene	MARKER	Recombination fraction, θ				
		0	0.1	0.2	0.3	0.4
CLN1	D1S193	-1.6	0.2	0.2	0.1	0
CLN2	D11S1338	-∞	-0.53	-0.17	-0.05	-0.01
CLN3	D16S298	-1.3	0.4	0.3	0.2	0.1
	DS16S299	-1.3	0.4	0.3	0.2	0.1
CLN6	D15S1000	-1.6	0.2	0.2	0.1	0
	D15S983	0.5	0.3	0.19	0.08	0.02
CLN8	AFM265wb5	1.17	0.78	0.46	0.22	0.06
	44CA	-1.6	0.2	0.2	0.1	0

Table 6.4. Two-point lod scores for markers at *CLN1, CLN2, CLN3, CLN6* and *CLN8* genes.

Three *CLN5* flanking markers were tested (D13S156, D13S162 and D13S1306), which were found to be homozygous in only the two affected individuals and not in unaffected family members (Figure 6.4A). Both affected individuals were homozygous 1/1 for the three markers and typing these markers in the other branch of the family, where the third affected child was identified, revealed that all unaffected individuals where at the most heterozygous for allele 1, with exception of individuals IV:2 and V:1 (Figure 4A). These two individuals were homozygous for allele 1 at marker locus D13S156. The haplotype characterized by alleles 1-1-1 at markers D13S156-D13S162-D13S1306 was only homozygous in both available affected

individuals. Two-point linkage analysis resulted in lod scores higher than 2 for markers D13S162 and D13S1306 (Table 6.5).

Thus it seemed that *CLN5* was linked to the disease in our family. For the marker D13S156, the affected individuals were homozygous 1/1 as was their brother VI:4. Since all the children of V:6 received an allele 1 from him, it was inferred that he was homozygous 1/1 (No sample was available). In addition, individual IV:2 and his daughter V:1 were also found to be homozygous 1/1 for this marker.
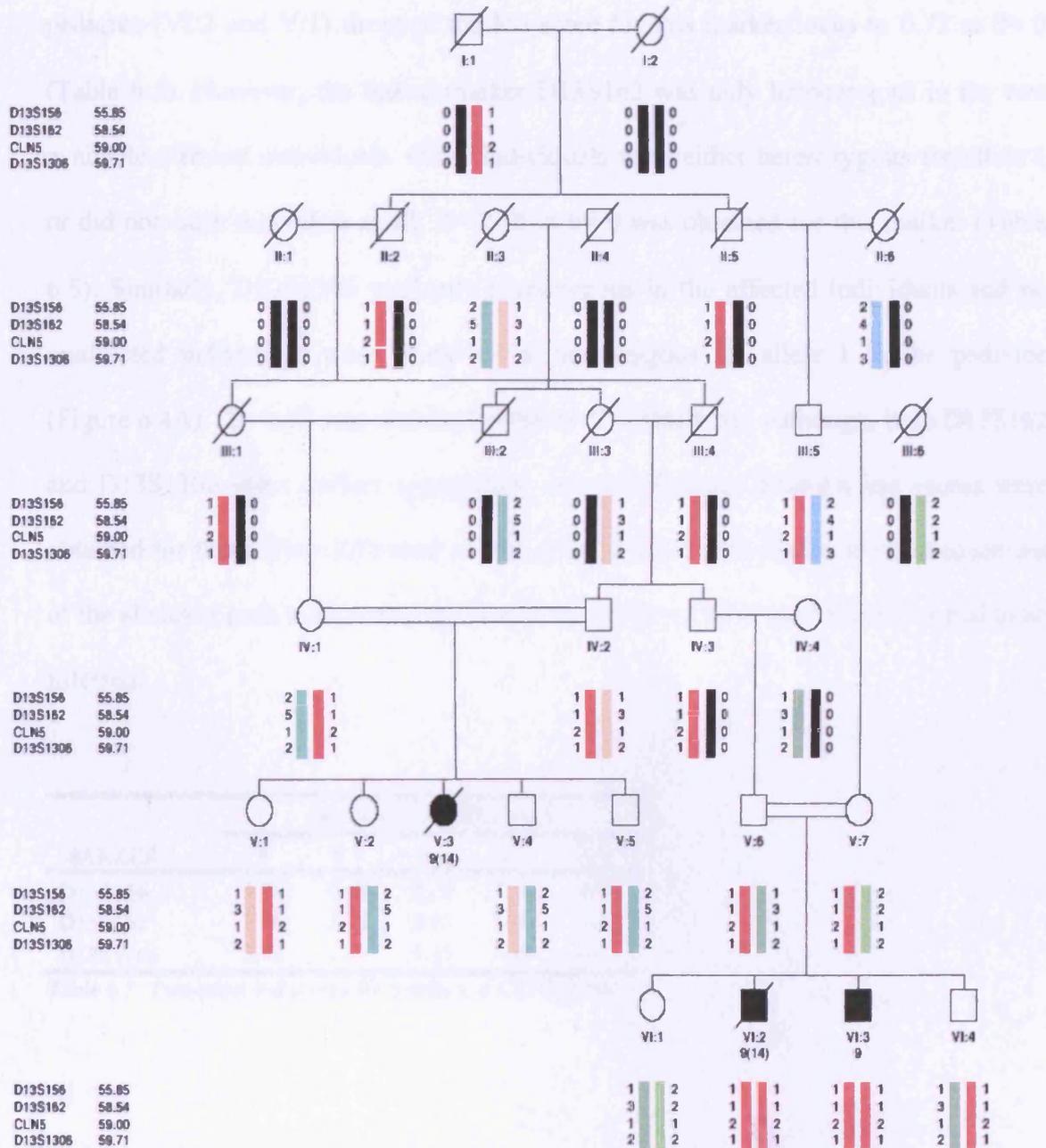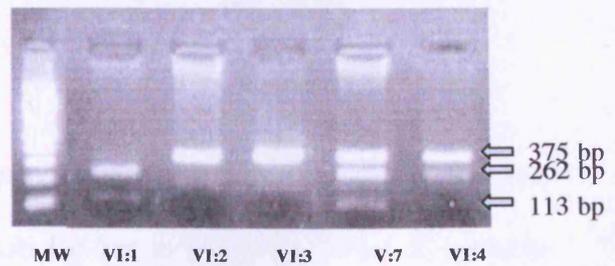
Figure 6.4. (A) Colombian pedigree with clinical NCL. Affected individuals are indicated their age at onset (age at death). Haplotypes for markers at CLN5 locus are shown. In CLN5 a number 2 indicates mutation c.1627G>A. Haplotypes were inferred for individuals IV:3, IV:4 and V:6. (B) Restriction enzyme digestion with *PshAI* of a PCR fragment from CLN5 exon 2 in the nuclear family of the two living affected siblings. Patients present only one undigested band of 375 bp. Unaffected individuals appear as either heterozygous or homozygous for the wild allele (bands 262 and 113 bp).

The meioses of IV:2 and V:6 were therefore not informative for linkage purposes. Moreover, the homozygosity observed in two other unaffected individuals in the pedigree (VI:2 and V:1) dropped the lod score for this marker locus to 0.72 at θ= 0 (Table 6.5). However, the linked marker D13S162 was only homozygous in the two available affected individuals. Other individuals were either heterozygous for allele 1 or did not carry this allele at all. Z= 2.76 at θ= 0 was obtained for this marker (Table 6.5). Similarly, D13S1306 was only homozygous in the affected individuals and no unaffected individuals were found to be homozygous for allele 1 in the pedigree (Figure 6.4A). Z= 2.42 was obtained at theta= 0 (Table 6.5). Although, both D13S162 and D13S1306 show perfect segregation with the disease, different lod scores were obtained for them. This difference in lod scores is due to the difference in frequencies of the alleles at each marker and the fact that the genotypes of individual V:6 had to be inferred.

| MARKER | Recombination fraction, θ | | | | |
	0	0.1	0.2	0.3	0.4
D13S156	0.72	0.49	0.27	0.12	0.03
D13S162	2.76	2.05	1.35	0.74	0.3
D13S1306	2.42	1.77	1.15	0.64	0.27

Table 6.5. Two-point lod scores for markers at CLN5 locus.

3.4 Sequencing of CLN8 and CLN5

Sequencing of the three exons of CLN8 revealed a homozygous c.908 A>G transition in exon 3, which leads to the substitution Asn for Ser in position 225 of the protein (N225S). This change was observed in two unaffected individuals of the family (V:1 and VI:1).

Sequencing of *CLN5* exons identified a G to A transition in exon 2 (c.1627 G>A) that co-segregates with the disease-associated haplotype (Figure 6.5A). This nucleotide change leads to an Arg112His substitution in the protein (R112H). This mutation was not present in any of 29 control individuals from the same population. Alignment of available *CLN5* sequences shows that arginine 112 is conserved across vertebrates, including rodent, amphibian, and bird species (Figure 6.5B).

Figure 6.5. (A) Partial sequence of CLN5 exon 2 from individuals VI:3 (homozygous) upper panel, and V:7 (heterozygous) in lower panel. Arrow indicates the mutant nucleotide position. (B) Alignment of the CLN5 aminoacid sequence encoded by exon 2 in 12 vertebrate species. Arginine 112 is boxed.

4. Discussion

The purpose of this chapter was to characterize an extended Colombian family with juvenile NCL (Batten disease). Patients started their symptoms at age 9 and continued to have a progressive and severe disease that ended in death for two of them at age 14 (V:3 and VI:2). The subsequent onset of seizures was consistent with a diagnosis of Batten disease (Juvenile NCL). However, no vacuolated lymphocytes were found in samples from the patients that were available. Furthermore, pathological examination suggested a variant rather than a classic juvenile form (Dr Anderson, Personal communication; Histopathology Department, UCL), since the classical fingerprints encompassed with curvilinear profiles or lamellar inclusions [306, 307] were not observed in our patients, but instead condensed fingerprint profiles occasionally associated with lipid droplets were found. These two observations suggested that this variant might be due to a mutation in a novel NCL locus.

Candidate NCL genes were then analyzed first. Genetic analysis confirmed that the disease in this family is not caused by mutations in *CLN3*, nor in *CLN1*, *CLN2*, *CLN6* or *CLN8*. However, markers linked to *CLN5* locus were consistent with the transmission of the disease. Thus, available affected individuals were homozygous for alleles 1-1-1 at markers D13S156-D13S162-D13S1306. Consistently, parents of the third affected individual (V:3) were carriers of such a haplotype, leaving the possibility that she was homozygous for the haplotype and therefore for the mutation associated with it.

Sequencing of *CLN5* exons led to the identification of a novel mutation that consisted of a c.1627 G>A transition. This mutation leads to the substitution of the conserved Arg 112 (Figure 3B) and was not found in the control individuals.

147

To date, *CLN5* mutations have been reported only in Finnish and Swedish families and in one Dutch individual [269, 303]. All these patients presented an identical late infantile onset NCL. There are clear clinical differences between northern European patients described with *CLN5* mutations and the patients that we are reporting here. First, the age of onset was on average 3.25 years old [303]. In contrast, for our three patients the age of onset was 9 years old. The first symptoms observed in the European patients were clumsiness and disturbance of concentration, whereas our patients presented first with symptoms of visual failure. However, the age of onset of seizures was approximately the same in the European and the Colombian patients. It was on average 9.3 years old for the Europeans [303] and 9.5 years old for the Colombian patients.

This is the first *CLN5* mutation outside northern Europe. These findings also show that mutations in *CLN5* can result in a juvenile as well as the late infantile NCL. *CLN5* should thus be considered as a candidate for late infantile and juvenile NCLs, even outside northern Europe.

Recently, it was reported that the CLN5 protein is targeted to the lysosome [301], where it interacts with CLN2 and CLN3 proteins [302]. Pathogenic mutations may affect interaction with other CLN proteins or trafficking of CLN5, as well as possibly causing loss of function. Evaluation of the molecular effect of disease-causing mutations should facilitate an understanding of disease mechanism, including the formation of lysosomal deposits

CHAPTER SEVEN

GENETIC CHARACTERIZATION OF
AN EXTENDED PEDIGREE WITH
FAMILIAL ABDOMINAL PAIN

1. Introduction

1.1 An extended family with many member suffering from Abdominal pain.

A four generation family in which many members presented with unexplained abdominal pain was recruited from the Servicio de Neurologia (NEUROCIENCIAS) from Universidad de Antioquia-Colombia. Eight affected females and seven affected males were identified. Both sexes transmitted the disease (Figure 7.1).
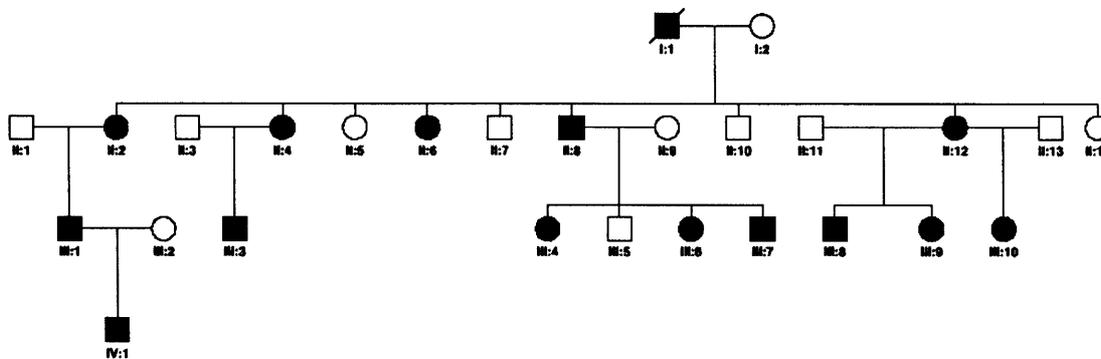
Figure 7.1. Colombian family with abdominal pain. It includes twenty-seven individuals. No samples were available from individuals I:1, II:1, II:3 and II:4.

Age of onset of this condition was 1 day-10 months. Most frequently the first crisis was between 3-4 months. The crises were in general more frequent in infancy than in other periods of life. This may be the case because adults have learnt how to inhibit development of a crisis. Young children have had up to two crises per day.

Once the crisis has begun it can sometimes be stopped by having food, usually sweet, or by applying heat or resting. Crises can be stopped in the first 5-15 minutes; but after this time having any food will cause nausea and vomiting.

Usually, the pain begins in the shoulders, chest and back area, then it descends and generalizes to all the body, with exception of the head, face and visceral structures. Pain is described as a burning sensation and is associated with loss of strength, difficulty of breathing, abundant sweating, coldness and generalized paleness. In addition, perioral cyanosis, hand spasms and generalized muscular contracture can also be present. It resembles colic in that is spasmodic and the pain ascends and descends during the crisis. However, in contrast to the pain in colic, in this disorder there is pain of the abdominal wall rather than visceral pain. This disease mainly affects the striated muscular system (shoulders, arms and abdomen).

During a crisis, the intensity of the pain increases gradually. When pain reaches its maximum, the patient falls into a deep sleep for one to two hours (Postictal sleep). When he/she wakes up the pain is gone but there is a feeling of weariness. Postictal sleep can be inhibited by stress (This was the case of one patient in the army). During a crisis there is no loss of consciousness but it leads to a post-ictal period characterized by somnolence and muscular tiredness. In all, crisis may last half an hour to an hour.

Some of the family members observed that crises are more frequent during other illnesses. In addition, hangover, excessive exercise or fasting are clear triggering factors. The grandmother of the family observed that the crises seemed to be identical in all affected individuals in the family, "It looks like a carbon-copy in the all".

Further, clinical characterization of these patients is underway but it is likely that the most appropriate diagnosis is Abdominal Epilepsy (AE), which is a form of autonomic epilepsy.

1.2 Literature review on Abdominal Epilepsy (AE).

Partial and generalized epilepsies alter autonomic function during ictal (during crisis), post-ictal and inter-ictal states. All aspects of autonomic function can be affected, including the parasympathetic, sympathetic and adrenal medullar systems [308]. Autonomic symptoms can be divided into cardiovascular changes, respiratory manifestations, gastrointestinal symptoms, cutaneous manifestations, pupillary symptoms, genital and sexual manifestations as well as urinary symptoms [309]. Gastrointestinal symptoms can be subdivided into epigastric auras, abdominal epilepsy, ictal spitting and ictal vomiting and ictal retching. Whereas epigastric auras frequently precede or are associated with complex partial seizures, abdominal sensations as the sole manifestation of epileptic seizures are rare [310, 311]. However, gastrointestinal signs and symptoms may be the primary or only manifestation of seizures, and are therefore referred to as abdominal epilepsy (AE) [312].

The main criteria to establish a differential diagnosis of AE are the occurrence of paroxysmal gastrointestinal complaints that are unexplained after complete clinical evaluation, symptoms of a central nervous system disturbance and abnormal electroencephalogram (EEG) with findings specific for a seizure disorder, and a sustained abolition of symptoms on anticonvulsant medication [313]. EEG has rarely been evaluated during an AE crisis and is not necessary for the diagnosis [314, 315]. Moreover, not all of the patients have presented an abnormal EEG [316].

However, some authors stress the importance of both disturbances of consciousness and specific EEG abnormalities as the types of evidence that could truly separate AE from other causes of recurrent abdominal pain [310].

Most reported cases are children and adolescents [317, 318]. In a recent review of the 36 patients reported in the last 34 years in the English literature, Zinkin and Peppercorn [319] showed that age of onset of AE ranged between 1 to 66 years old. They also found that males and females are affected about equally (53% females). Gastrointestinal manifestations of seizures included abdominal pain (86%), nausea and/or vomiting (28%), diarrhoea (5%) and bloating (3%) [319].

The abdominal pain was commonly sharp or colicky, in the periumbilical area and upper abdominal, including epigastric and right upper quadrant.

Neurological manifestations included lethargy, fatigue and/or post-ictal sleep (36%); some loss of consciousness was reported in 64% of the 36 patients, including complete loss of consciousness or generalized tonic-clonic seizures (36%). Less frequently, dizziness (8%), headaches (14%), pallor/sweats (11%), fever (6%) and blindness (6%) were found. Treatment was with anticonvulsants, most commonly phenytoin, Phenobarbital or carbamazepine, improving, most with complete or near complete resolution of symptoms. The duration of episodes was typically no more than a few minutes, but 5 (14%) had symptoms lasting for half an hour or more [319].

Most cases of AE are of temporal lobe origin [313]. However, several cases have been traced to parietal lobe lesions [320].

Rather than considering AE to be a separate entity, it may be more appropriate to simply recognize that some patients will have significant gastrointestinal symptoms as

a manifestation of their seizures. When abdominal pain is followed by a generalized

seizure, it might be referred to as an "abdominal aura" [319].

1.3 Aims

Even though AE was documented several years ago, no indication of a genetic link

was reported previously. It is even the case that the occurrence of two affected

individuals in the same family has never been documented.

However, this large Colombian family has 15 individuals affected with a condition

classifiable as AE. Although we have clear assignment affected and unaffected

individuals, this genetic study was undertaken ahead of the full clinical

characterization since, for example, no EEG has yet been taken for any individual in

the family. The disorder in this family presents as a clear autosomal dominant

condition.

Given that there were no clear candidate genes to test for, the aim of this chapter was

to evaluate genetic markers across the genome as part of a genome wide approach. In

this way it might be possible to find linkage to a chromosomal region that could indeed

be a novel locus associated with epilepsy (i.e. AE) . If this were achieved a further

fine mapping step would be performed with the aim of identifying the mutant gene.

2. Methods

2.1 Power simulation

Power simulations were performed assuming autosomal dominant inheritance and high

penetrance (0.985). A marker locus with four equally frequent alleles was simulated in

one hundred replicas using SLINK [55]. The disease frequency was assumed to be rare

(0.0001). The actual simulated pedigree is shown in Figure 7.1.

2.2 Genome wide scan

A genome wide scan of microsatellites markers were typed by deCODE in all

available individuals at an average density of 6 cM, as described in Chapter 2.

2.3 Microsatellite marker typing

In addition to the deCODE scan, eight additional markers were typed in the interval

D8S512-D8S279. These were D8S533, D8S1775, D8S1767, D8S1792, D8S1117,

D8S543, D8S1795 and D8S1807. Primer sequences and fluorescent labels were as

shown in Table 7.1.

Marker typing was performed as indicated in Chapter 2.

MARKER	LABEL	SEQUENCE 5' to 3'	ASR
D8S1795	HEX	TGAGCCCAATATGACAATGC CCAAAAATNCACAAACCTG	219-227
D8S1807	TET	CCAAAACATTGCTCCC CGCCTGTCCCTAGTTC	226-240
D8S543	FAM	TGGTGTCATTGCTTTCTAGTCT TGCACAGGTGAGTAAATTTGTAA	116-140
D8S1792	FAM	AACTCCAAAACCTATAGANCAT TCCGTTCCATGCTCTTC	238-290
D8S1767	HEX	GCATGAGGTTTACAGTTAATGATGT CAGTCAAATAAATTACATGTCCATC	137-159
D8S533	TET	CTTTGCCAGGGTGTTCAGAG AGAGCCTTGTTTCATGGGAC	159-169
D8S1775	FAM	TGGCAAATACACACTCTGCT CCAAATAGGCTGATGAGAAACT	136-162
D8S1117	HEX	TTATTTTCCTGCAGACTCCG CAGGAGAAAAGCATGCAGAT	183-204

Table 7.1. STR markers used to fine mapping the AE locus. Oligo F was labelled as indicated. Oligo R had no label. ASR= average size range. Sequences and ASRs were obtained from UCSC genome browser (http://genome.ucsc.edu/).

2.4 Two-point, multipoint and haplotype analysis.

Two-point lod scores were calculated using MLINK from the package LINKAGE [44] and assuming autosomal dominant inheritance. Allele frequencies were those provided by deCODE, who had estimated them from a pool of samples of individuals with different ethnic backgrounds. Penetrance was set to 0.985 (same as in the power simulation).

Multipoint lod scores and haplotypes were obtained with simwalk2 [49] using same conditions as for two-point analysis.

3. Results

3.1 Power analysis

Table 7.2 shows power simulation results. This simulation indicated that this family on its own is powerful enough as to obtain evidence of linkage at $\theta=0$. If phenocopies are not present in the pedigree, there is 30% chance of obtaining evidence of linkage (i.e. $Z \geq 3$) to a locus 5 cM away from the actual tested marker locus ($\theta= 0.05$) is still feasible in this pedigree.

	Expected Lod Score (Z)		Probability of reaching $Z \geq$ to:		
	Average	Z max	1	2	3
$\theta= 0$	3.32	5.38	94	87	46
$\theta= 0.05$	2.37	4.96	81	64	30

Table 7.2. Power simulation of family with Abdominal epilepsy.

3.2 Genotyping results

Table 7.3 shows a statistical summary of the deCODE marker typing. Although there were no data at all for marker locus D1S219, on average the deCODE marker typing efficiency was as high as 96%. The individual with the fewest genotypes was II:9 (91%). In contrast, the individual with the most genotypes was II:7 (99%). Per chromosome, the least genotypes were obtained for the X chromosome (91%). The highest typing efficiency was obtained for chromosome 22 (100%; Table 7.3).

INDIVIDUAL	CHROMOSOME																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X
I:2	1.00	1.00	0.94	1.00	1.00	1.00	1.00	1.00	1.00	0.96	0.95	0.96	1.00	0.96	0.95	1.00	1.00	1.00	0.87	1.00	1.00	1.00	1.00
II:2	1.00	1.00	0.97	1.00	1.00	0.96	1.00	0.96	0.96	1.00	1.00	0.93	0.94	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00
II:5	0.98	1.00	0.94	0.97	0.96	0.96	1.00	1.00	0.92	1.00	1.00	0.96	0.94	0.96	1.00	1.00	0.95	1.00	0.93	1.00	1.00	1.00	0.96
II:6	0.98	0.94	0.94	0.97	0.93	0.86	0.83	0.91	0.92	0.89	0.95	0.93	0.76	0.91	0.90	0.89	0.95	0.93	0.87	0.93	0.93	1.00	1.00
II:7	1.00	1.00	0.94	1.00	1.00	0.93	1.00	0.96	0.96	1.00	1.00	0.96	1.00	1.00	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
II:8	1.00	1.00	0.94	0.97	1.00	0.96	0.97	0.96	0.92	1.00	0.95	1.00	1.00	0.96	1.00	1.00	1.00	1.00	0.93	0.93	0.93	1.00	0.88
II:9	0.98	0.92	0.97	1.00	0.96	0.96	1.00	0.96	0.92	1.00	0.90	0.96	0.94	1.00	0.90	0.83	1.00	0.93	0.93	0.93	1.00	1.00	0.00
II:10	0.98	0.97	0.94	0.97	0.96	0.93	1.00	0.91	0.92	1.00	0.85	0.89	0.94	0.91	0.90	0.94	1.00	0.93	1.00	0.87	0.93	1.00	1.00
II:11	0.98	1.00	0.97	0.97	0.93	0.96	1.00	0.96	1.00	1.00	1.00	0.93	0.94	1.00	0.95	1.00	1.00	1.00	0.93	0.93	0.93	1.00	0.00
II:12	0.96	1.00	0.94	1.00	1.00	0.93	1.00	1.00	0.96	0.96	1.00	0.96	1.00	0.96	1.00	0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00
II:13	1.00	1.00	1.00	1.00	0.96	1.00	0.97	1.00	1.00	1.00	0.95	1.00	1.00	1.00	0.95	0.94	1.00	1.00	0.93	0.93	1.00	1.00	0.96
II:14	1.00	0.94	0.97	0.97	0.93	0.93	0.97	1.00	0.96	0.93	1.00	0.96	1.00	0.96	0.85	0.89	0.95	1.00	0.93	0.87	1.00	1.00	0.96
III:1	0.98	1.00	0.91	0.90	0.96	0.96	0.97	0.87	0.88	1.00	0.95	0.89	0.94	0.96	1.00	0.94	0.95	1.00	0.87	1.00	1.00	1.00	0.96
III:2	1.00	0.97	1.00	0.97	0.93	0.89	0.97	1.00	0.88	0.89	0.95	0.93	0.94	0.96	1.00	0.94	1.00	0.93	0.93	0.87	1.00	1.00	1.00
III:3	0.96	0.94	1.00	0.97	0.96	0.96	0.97	0.96	0.88	0.96	0.90	0.89	0.94	1.00	0.95	0.94	0.91	0.93	0.87	0.93	1.00	1.00	0.96
III:4	0.96	0.97	1.00	1.00	0.96	1.00	0.97	0.96	0.92	1.00	0.90	0.89	0.82	1.00	0.90	0.83	1.00	0.93	0.93	1.00	1.00	1.00	1.00
III:5	1.00	0.94	0.89	0.97	0.93	0.93	1.00	0.96	0.92	1.00	1.00	0.96	1.00	0.96	0.90	0.89	1.00	1.00	1.00	1.00	1.00	1.00	1.00
III:6	0.98	1.00	0.97	0.86	0.93	0.96	1.00	0.96	0.92	1.00	0.90	1.00	1.00	0.96	0.85	0.89	1.00	0.93	0.87	0.93	1.00	1.00	1.00
III:7	1.00	1.00	0.97	1.00	0.96	0.96	1.00	0.96	0.96	1.00	1.00	0.96	1.00	0.96	1.00	0.94	1.00	1.00	0.93	1.00	1.00	1.00	1.00
III:8	0.98	0.97	0.94	0.93	0.96	0.93	1.00	1.00	0.92	0.96	1.00	0.93	0.94	1.00	0.95	1.00	1.00	1.00	1.00	0.93	1.00	1.00	0.00
III:9	0.98	1.00	0.94	0.90	1.00	0.96	0.93	1.00	1.00	0.96	1.00	1.00	0.94	1.00	1.00	0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00
III:10	0.96	0.97	0.94	1.00	1.00	0.96	0.93	1.00	1.00	0.89	0.95	1.00	0.94	1.00	1.00	1.00	0.95	1.00	0.93	0.93	0.93	1.00	1.00
IV:1	0.98	0.97	0.94	1.00	1.00	0.93	1.00	1.00	0.96	1.00	1.00	0.89	0.94	0.96	1.00	0.94	0.95	1.00	1.00	0.93	1.00	1.00	0.96
Average	0.98	0.98	0.96	0.96	0.96	0.95	0.97	0.97	0.94	0.97	0.96	0.95	0.95	0.98	0.95	0.93	0.98	0.98	0.94	0.95	0.99	1.00	0.91

Table 7.3. deCODE marker typing efficiency.

158

184 genotypes were required for fine mapping (eight markers in 23 individuals). From these only one failure was obtained in individual II:11 at marker D8S543. Marker typing efficiency in fine mapping corresponded to 99.46%.

3.3 Two-point analysis

Tables 7.5-7.26 show two-point lod scores (Z) for chromosomes 1-22, (see appendix 3), respectively and a summary of lod scores above 2 is shown in Table 7.4.

Locus/ Chromosome	Genetic map	Recombination fraction					Theta Max	Z max
		0	0.1	0.2	0.3	0.4		
Chromosome 2								
D2S337	86.05	-12.22	0.7	1.15	0.99	0.54	0.212	1.15
D2S2152	94.62	-12.19	**2.24**	2.04	1.5	0.75	0.115	2.24
D2S2110	99.58	-8.8	-0.09	0.15	0.14	0.04	0.5	0.12
Chromosome 3								
D3S3551	99.23	-13.34	-0.05	0.54	0.56	0.31	0.251	0.6
D3S3653	106.92	**2.73**	2.28	1.77	1.18	0.52	0	2.73
D3S1271	115.24	-8.26	1.53	1.32	0.93	0.41	0.091	1.53
Chromosome 8								
D8S1763	73.88	-2.23	2.05	1.74	1.22	0.56	0.078	2.06
D8S512	78.06	**4.18**	3.48	2.7	1.83	0.86	0	4.18
D8S279	86.52	-5.91	1.57	1.69	1.35	0.77	0.164	1.72
D8S1707	97.46	**2.11**	1.79	1.43	1.02	0.55	0	2.11
D8S1778	106.89	**2.7**	2.28	1.83	1.31	0.71	0	2.7
D8S1762	108.03	2.76	**2.92**	2.39	1.68	0.82	0.045	3.04
D8S1470	118.7	-3.01	1.36	1.23	0.93	0.52	0.104	1.36
Chromosome 9								
D9S2149	55.3	-13.55	0.09	0.62	0.6	0.29	0.238	0.66
D9S1777	67.34	**2.14**	1.78	1.37	0.9	0.37	0	2.14
D9S1876	70.42	-12.83	0.9	1.14	0.92	0.45	0.189	1.14
Chromosome 13								
D13S1315	119	-13.51	-3.05	-1.36	-0.55	-0.14	0.5	0
D13S293	131.92	**2.4**	2.04	1.63	1.16	0.63	0	2.4

Table 7.4. Summary of Two-point LOD scores above 2.
For a complete information see appendix 3.

Two-point lod scores over 2 were found on chromosomes 2, 3, 8, 9 and 13. For four of

these, the maximum lod score was obtained at theta= 0 and for one it was obtained at

theta= 0.1. Thus, on chromosome 2 a Zmax= 2.24 at θ= 0.1 was obtained (Table 7.4);

for chromosomes 3, 9 and 13, Zmax= 2.73, 2.14 and 2.4 were obtained, respectively

(Tables 7.4).

The only statistically significant two-point lod score was obtained for chromosome 8.

A Zmax= 4.18 at θ= 0 for marker D8S512 was obtained (Table 7.4). Flanking markers

at both sides of D8S512 were negative (<-2) at θ= 0. They presented their Zmax (1.72

and 2.06) at θ= 0.164 and 0.078, respectively (Table 7.4). Figure 7.2 shows two-point

lod scores at θ= 0 for markers on chromosome 8 and illustrates the very negative lod

scores for the flanking markers to D8S512, i.e. D8S1763 and D8S279. Nevertheless,

telomeric markers to D8S279 (D8S1707, D8S1778 and D8S1762) also presented Z>2

(Table 7.4 and Figure 7.2).
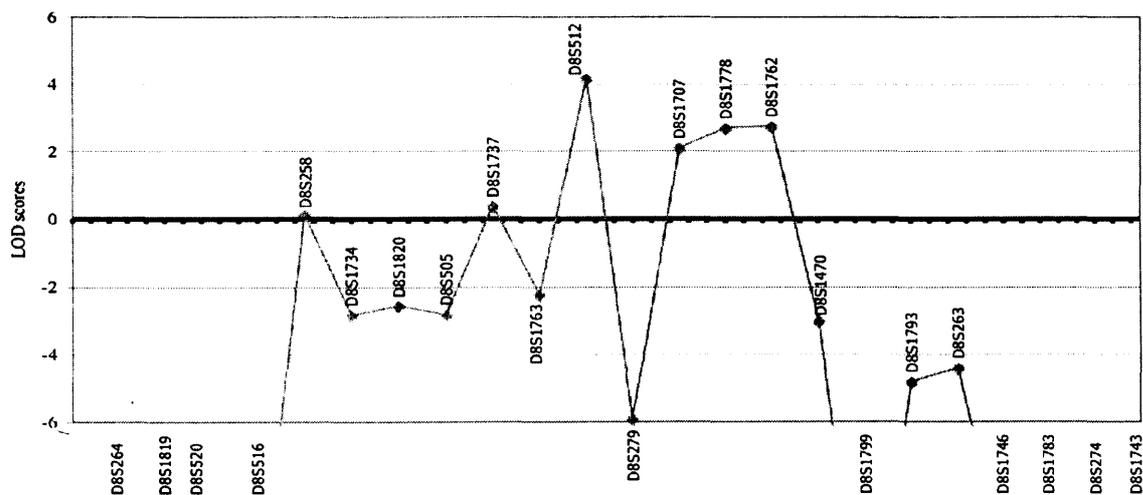
Figure 7.2. Two-point Lod scores for markers on chromosome 8, at θ=0

The second highest lod score was obtained for marker D3S3653 (map position 106.92), on chromosome 3 with a Zmax= 2.73 at θ= 0 (Figure 7.3). Its flanking markers presented Z <-2 at θ=0 and their Zmax were 0.6 and 1.53 for markers D3S3551 and D3S1271, respectively (Table 7.4).
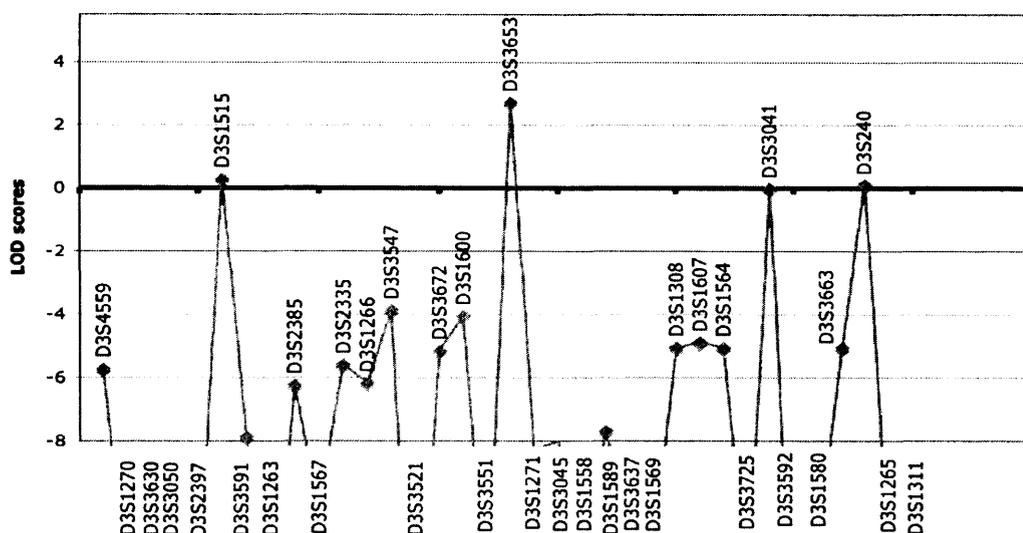
Figure 7.3. Two-point Lod scores for markers on chromosome 3, at θ=0

3.4 Multipoint analysis

Multipoint analysis was performed for both chromosomes 3 and 8. Multipoint scores for chromosome 3 resulted all in exclusion (Z≤-2), with exception of Z= -1.44 at map postion 94.06 and Z= -1.05 at map position 122.78 (Figure 7.4). These two locations are distant from D3S3653, which maps at 106.92 cM.
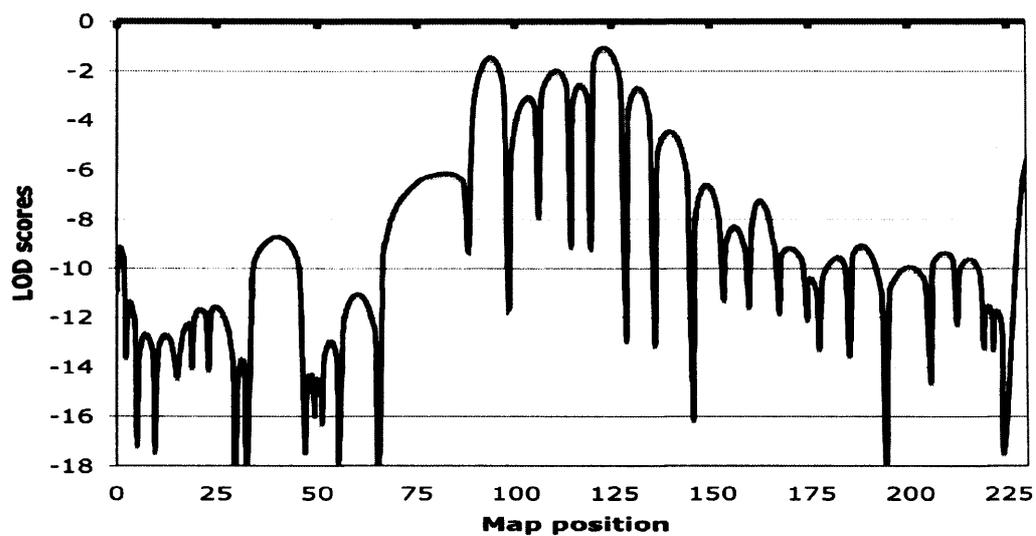
Figure 7.4. Multipoint LOD scores for chromosome 3.

Multipoint analysis for chromosome 8 showed a maximum lod score of 4.42 in the interval D8S512-D8S279, at map position 79.75 cM (Figure 7.5). Statistically significative lod scores ($Z \geq 3$) were also obtained for the telomeric region to marker D8S279 (Figure 7.5), where $Z > 2$ was obtained in the two-point analysis (D8S1707, D8S1778 and D8S1762; Figure 7.2 and Table 7.4).
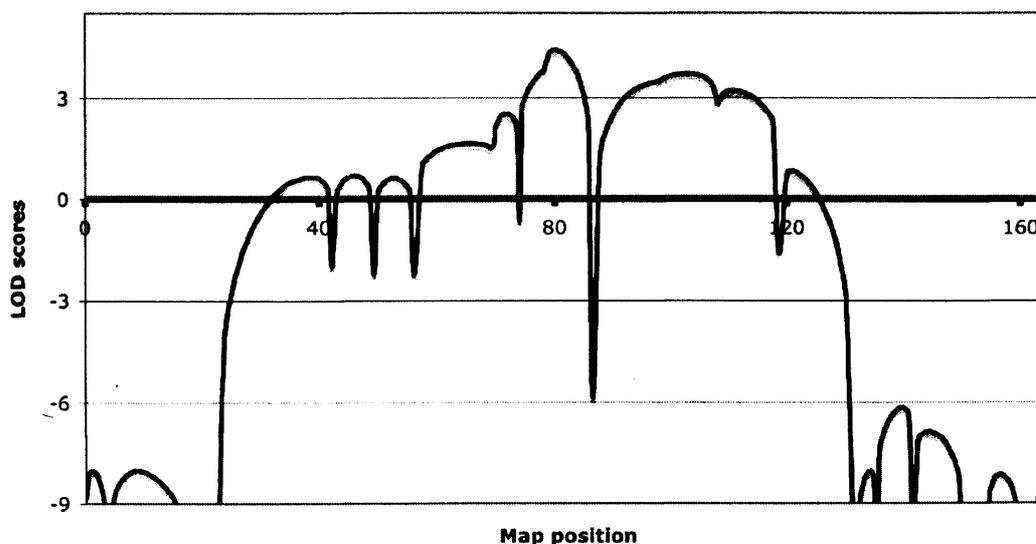
Figure 7.5. Multipoint LOD scores for chromosome 8. The highest peak corresponds to Z= 4.42 at map position 79.75 cM.

Eight additional markers in this interval (D8S512-D8S279) were therefore tested and this resulted in a maximum multipoint lod score of 4.57 at map position 80.5 between markers D8S1767 and D8S1792. Analyzing the data with full penetrance instead of 0.985 (as was used in the simulation) a raise in the maximum multipoint lod score was observed to 4.63, at the same genetic position (Figure 7.6).
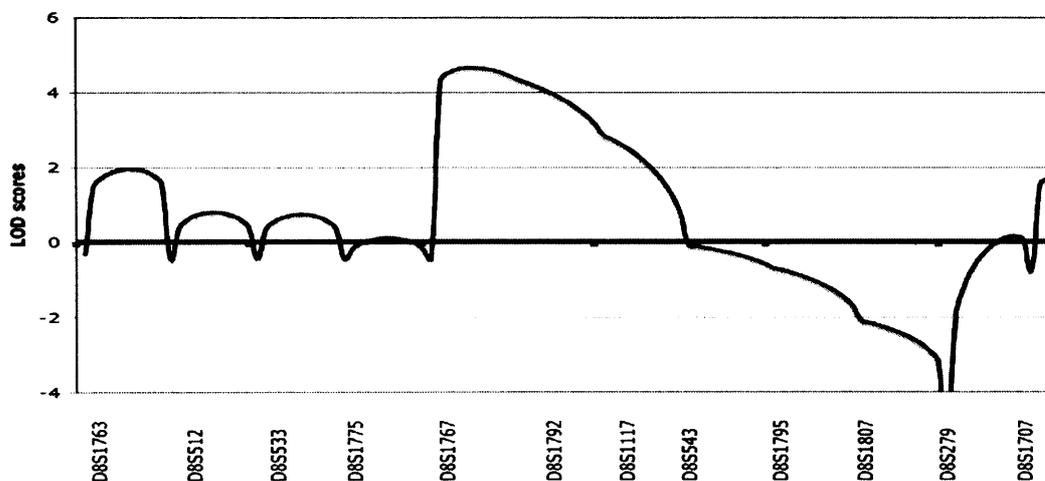
Figure 7.6. Fine mapping of AE locus. Multipoint lod scores for eight additional markers in the interval D8S512-D8S279.

3.5 Haplotype analysis

Figure 7.7 shows haplotype reconstruction for chromosome 8. Evaluation of the eight additional markers at a density of 1 marker every cM in the interval D8S512-D8S279 resulted in the identification of a disease-associated haplotype (Figure 7.8). All affected individuals carry a disease-associated haplotype characterized by alleles 2-3-2-2-2 at marker loci D8S1792-D8S1117-D8S543-D8S1795-D8S1807. A recombination event was evident in the interval D8S1767-D8S1792 in the affected individual III:8. Similarly, three recombinant chromosomes were observed in unaffected individuals II:5, II:7 and II:14 in the interval D8S1117-D8S543. Since

163

these three individuals carry alleles 2-2-2 at marker loci D8S543-D8S1795-D8S1807, the disease-associated haplotype can then be narrowed down to alleles 2-3 at marker loci D8S1792 and D8S1117, respectively.

Looking in more detail at marker locus D8S512, it can be noticed that all affected individuals have inherited an allele 3. However, in individual III:8 there is uncertainty as to which allele 3 he has inherited. This is due to the fact that his parents (II:11 and II:12) are both heterozygous for same alleles as III:8 (Figures 7.7 and 7.8). Nonetheless, considering the haplotypic information, it can be observed that in individual III:8 allele 3 was not received from his affected mother. This is inferred by considering the next marker locus down in the haplotype (i.e. D8S533), where an allele 4 could only be inherited from his mother. This allele 4 is in coupling (same chromosome) with allele 1, which are only present in his mother. This clarifies why the highest multipoint lod score was not obtained at $\theta=0$ in D8S512 but in the interval D8S512-D8S279 (Figure 7.6). The normal maternal chromosome was transmitted to III:8 up to the interval D8S1767-D8S1792, where a recombination was evident.
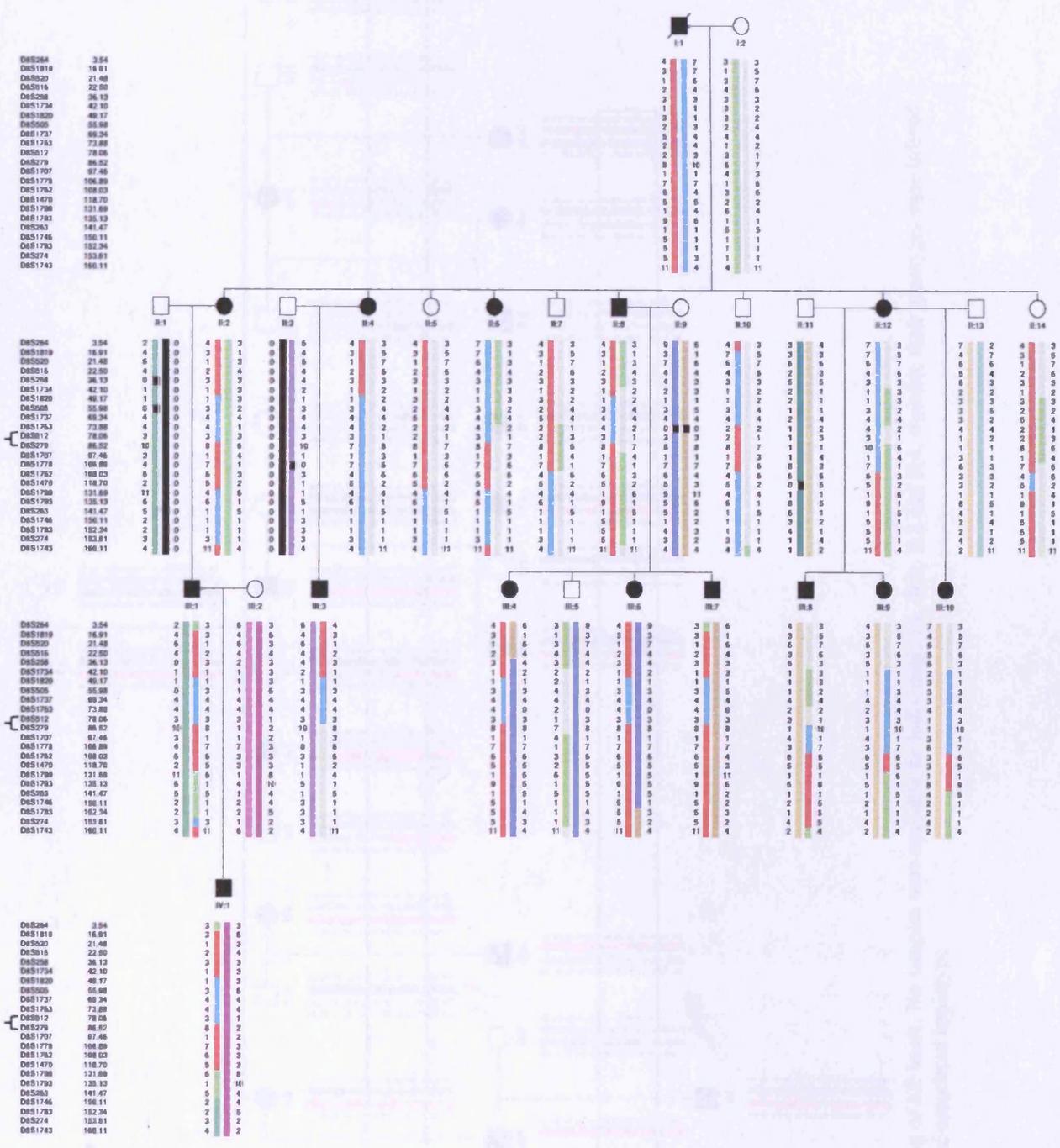
Figure 7.7. Haplotype reconstruction for deCODE markers typed on chromosome 8. Genotypes were inferred for individuals I:1. II:1, II:3 and II:4. The maximum two-point lod score was obtained for marker D8S512 and the maximum multipoint lod score was found in the interval D8S512-D8S279.

Figure 7.8. Fine mapping of AE locus. No samples were available for individuals I:1, II:1, II:3 and II:4, therefore their genotypes were inferred. Boxed appears the disease-associated haplotype

4. Discussion

The inherited condition in the family examined here clearly shows an autosomal dominant inheritance with high penetrance (probably 100%). Evaluation of markers across the genome, with a density of 6 cM, resulted in the identification of a locus on chromosome 8. Initial two-point linkage analysis showed a significant lod score (Z=4.18 at θ=0) for marker D8S512. Then, multipoint analysis showed that the highest lod score was obtained in the interval D8S512-D8S279, which comprises an 8 cM region. Fine mapping of this region resulted in identification of a disease-associated haplotype and multipoint analysis including these eight new markers showed a maximum lod score of 4.63 at map position 80.5 cM, which corresponds to the interval D8S1767-D8S1192. All affected individuals carry the haplotype characterized by alleles 2 and 3 at marker loci D8S1792 and D8S1117. The centromeric flanking marker was defined in individual III:8, who presentes a recombination event in the interval D8S1767-D8S1792. The telomeric boundary was identified in individuals II:5, II:7 and II:14 as being in the interval D8S1117-D8S543. The interval D8S1767-D8S1792-D8S1117-D8S543 covers a region of 1.3 Mb, which corresponds to 1.85 cM.

Interestingly, in this chromosome arm there are at least four other epilepsy-related loci. These are Benign Adult Familial Myoclonic Epilepsy (*BAFME*) at 8q23.2-q24.11 [321], Familial febrile convulsions (*FEB1*) at 8q13-q21 [215], Childhood absence epilepsy (*ECA1*) at 8q24 [322] and Benign Familial Neonatal Epilepsy (*EBN2*) at 8q24 (*KCNQ3*) [214]. Strikingly, our candidate region is within *FEB1* locus. *FEB1* is flanked by markers D8S533 and D8S279 [215]. The candidate genes they suggested are outside our

candidate region. However, it could be the case that both disorders are due to mutations in the same gene, located in our candidate region. There are many examples of clinical heterogeneity of syndromes that are relative to mutations in a single gene. For example, mutations in *SCN1A* have been described in both Generalized epilepsy with febrile seizures plus (GEFS+) and severe myoclonic epilepsy of the infancy (SMEI) [236, 237]; also, *GABRG2* mutations have been associated with GEFS+, SMEI and childhood absence epilepsy (CAE) phenotypes [243-246]. Likewise, *CLN1* mutations have been associated with different clinical forms of Neuronal Ceroid Lipofuscinoses (NCLSs) [281]. So, although the clinical entity reported by Wallace and colleagues (1996) [215] is very different from the one I am reporting here, both are types of epilepsy. Another difference between *FEB1* and *AE* expressivity is the associated penetrance. The mutant allele in *FEB1* is suggested to be 60-65% penetrant, while I am proposing that AE is highly penetrant (probably 100%).

Future work

In this candidate region there are only three identified genes, *DEPDC2* and *VEST1* in addition to a hypothetical protein, *DKFZp547E186*. *DEPDC2* has 45 exons and extends 279.58 kb. There is no evidence of its function but it has been suggested that it might play a role in an intracellular signalling cascade, and have Guanyl-nucleotide exchange factor activity and protein binding [323, 324]. *VEST1* has 20 exons, extends 488.30 Kb and, to date, its function is completely unknown [324]. A possible approach to uncover the mutant gene, responsible for the disease in this family, could involve the evaluation of functional domains of other implicated proteins in epilepsy, to see whether they have some degree of homology with DEPDC2, VEST1 or DKFZp547E186 proteins. Once identified a level of homology, sequence of that particular gene would reveal the actual genetic change underlying the functional

168

alteration. If no homology was identified, sequencing of all the three genes would be the next step.

In conclusion, a highly significant lod score was identified for chromosome 8, which after fine mapping showed an increase of the lod score. Moreover, a small haplotype at 8q13.2 that clearly segregates with the disease was identified. The candidate region is flanked by markers D8S1767 and D8S543. This is the first locus reported for abdominal epilepsy so far.

CHAPTER EIGHT

LINKAGE ANALYSIS OF TYPE 1

DIABETES MELLITUS IN AN

EXTENDED PEDIGREE

1. Introduction

1.1 Generalities

Diabetes mellitus (DM) is a group of metabolic illnesses characterised by hyperglycaemia resulting from defects in the secretion and/or action of insulin. Symptoms of a clear hyperglycaemia include polyurea, polydipsia, loss of weight, sometimes polyphagia, and blurred vision [325].

Most cases of diabetes are located in two large etiopathogenetic groups: in one of them, the Type 1 Diabetes Mellitus (T1D), the cause is the absolute deficiency of secretion of the insulin and it can be subdivided in autoimmune (T1AD) and idiopathic (T1BD). The second category, the Type 2 Diabetes Mellitus (T2D), is caused by a combination of both resistance and an inadequate secretion of insulin [325]. A third category is named other types of DM. This category includes MODYs (Maturity onset of Diabetes in the Young), which is characterized by age of onset < 25 years old, autosomal inheritance with high penetrance, and no antibodies have been associated with this disorder [326]. These features together with non dependence on insulin, at diagnosis, led them to be classified in the past as a form of T2D. And the fourth type of Diabetes is referred to as Gestational Diabetes.

DM is one of the most common illnesses around the world, with a frequency around 3% [327]. T1D is a disease mostly of white populations and its prevalence varies among different countries. The highest prevalence is in northern Europe ranging from 3.7 to 20 per 100.000 [328]. At United States a prevalence of 0.4% has been reported [329]. A

concordance rate of 50% or less in T1D monozygotic (MZ) twins contrasts with the familial aggregation and a concordance of almost 100% in monozygotic twins in T2D [330]. Incomplete concordance observed in MZ twins indicates that environmental factors also contribute to the aetiology. Some studies have provided insights that viruses, cow's milk proteins amongst other suggestions, contain the antigens which provoke the antibodies involved in the autoimmune response; for example, congenital rubella infection has been shown to be associated with increased risk of T1D [331].

Nevertheless, T1D shows increased relative risk in siblings and children of affected parents, suggesting an important genetic component. The risk to siblings has been reported as 15-fold higher than is the population prevalence [332]; similarly, children from affected parents have been reported at an increased risk and particular increase has been noticed when the affected parent is the father [333].

1.2 Genetics of T1D.

To date more than twenty different putative loci have been described (Table 8.1) and it has been estimated that approximately another 50 loci remain to be identified. Two different approaches have been used extensively to identify susceptibility genes in T1D: Candidate genes and positional cloning approaches. The first studies were based on the candidate gene approach. Thus, because of the function of the *HLA* genes (Human Leucocyte Antigens) in the immune response and the fact that autoimmunity was strongly implicated in the pathogenesis of this disease, these genes were likely candidates in T1D.

Indeed human susceptibility to T1AD is associated to the *HLA* locus at 6p21.3, named *IDDM1*. Two particular loci in *HLA* have been shown to account for approximately 50% of the heritable component of susceptibility in T1D.

Locus	Localisation	Markers
IDDM1	6p21.3	*HLA-DRB1, DQB1*
IDDM2	11p15	*INS-VNTR*
IDDM3	15q26	D15S507
IDDM4	11q13	*FGF3,* D11S1337
IDDM5	6q25	*ESR*
IDDM6	18q21	*JK,* D18S487
IDDM7	2q31	*HOXD8,* D2S152
IDDM8	6q27	D6S264, D6S446
IDDM9	3q21-q25	D3S1576
IDDM10	10p11-q25	D10S193
IDDM11	14q24.3-q31	D14S67
IDDM12	2q33	*CTLA4*
IDDM13	2q35	D2S164
IDDM14	--	--
IDDM15	6q21	D6S283
IDDM16	14q32.3	D14S542, IGH
IDDM17	10q25	D10S554
IDDM18	5q33-q34	*IL12B*
Unnamed	1q42	D1S1617
Unnamed	16q22-q24	D16S3098
Unnamed	19p13	D19S247
Unnamed	19q13	D19S225
Unnamed	Xp13-p11	DXS1068
Unnamed	7p13	*GCK*
Unnamed	12q14-q15	*IFNG*

Table 8.1. Genetics of T1D. Modified from [334, 335].

HLA-DRB1 and HLA-DQB1 are the major components of IDDM1 [336-343]. Approximately 95% of white patients with T1D have either alleles DR3 or DR4, and 55% -60% have both alleles [344], though different susceptibility *HLA* alleles may be involved in other ethnic groups than white populations [345]. The *DR2* allele has been observed to have a protective effect [344]. In the case of *HLA-DQB1*, the substitution of Asp57 for Ala, Val or Ser in the molecule has been strongly associated with T1D [336, 338]. This change on its own is not enough to fully explain the *HLA* effect in the disease.

However, if an allele *DQB1*0302* is associated with a DR4 allele, thus makes a strongly predisposing haplotype [346].

The second most significant susceptibility locus (*IDDM2*) is the 5′ flanking polymorphism of the insulin gene (*VNTR-INS*), located on chromosomal region 11p15 [347]. *IDDM2* consists of a VNTR (variable number of tandem repeats) located 596 bp upstream of the start site of transcription of the INS gene [348-350]. The *VNTR-INS* presents three alleles (S, M, L) that have been associated with different levels of the insulin mRNA in both thymus [351, 352] and the pancreatic islets [353, 354]. Alleles at a SNP located 23 bases upstream from the *INS* gene have been associated with specific alleles in the VNTR [355].

Other loci have been identified by a positional cloning approach based on scanning the genome and identifying chromosomal regions, rather than particular genes. Uniformly-spaced previously mapped microsatellite marker loci are used to identify regions linked to the disease. This method is applied in both families and in population based studies. However, in polygenic traits such as T1D is difficult to identify large families with multiple affected members. Therefore, the affected sib-pairs (ASPs) method is an alternative and it has been proven to be especially valuable in identifying most of the susceptibility loci shown in Table 8.1. This method requires determination of the genotypes of pairs of siblings affected with the disease to compare the average proportion of alleles shared at each marker locus [356]. By these means, locus *IDDM12* on 2q33 [357 1999], where the *CTLA4* gene lies, was identified. This gene has been implicated in T1D [357-360] as well as in Grave's disease [361] and Hashimoto tryroiditis [362]. This gene is expressed only in activated T-lymphocytes and functions by down regulating T-cell function [363] and so seems to be a likely candidate for involvement in

174

disorders with autoimmune pathogenesis, such is T1D.

Even though the majority of T1D patients present a pathogenic process related to autoimmunity, there is a group (~10%) that does not present autoimmune pathogenesis (T1BD). By studying families and ASPs with T1BD (Non-autoimmune) locus *IDDM11* was identified on chromosome 14 [364]. Most support for linkage in this locus came from the subset of families where there was not increased sharing of HLA alleles.

1.3 Polygenic versus Oligogenic inheritance in T1D.

A completely different approach used to map susceptibility genes in polygenic diseases is animal models. Studies in animal models have provided indications that human T1D may have both polygenic and oligogenic inheritance. Studies on the non-obese diabetic mouse strain (NOD) have suggested approximately 15 loci contributing to disease [365, 366]. In this mouse strain a chromosomal region was identified on chromosome 1 that was associated with the susceptibility for T1D, which has syntheny with human chromosome 2 [367]. Evaluation of this region in humans led to the identification of a 23 cM region on chromosome 2, which contains three susceptibility T1D loci including *IDDM12* (*IDDM7, IDDM12 and IDDM13*) [368-370]. Despite the fact it is not possible to assess the mode of inheritance, several genes conferring susceptibility to T1D have been identified in this chromosomal region, such as (*CTLA4, HOXD8, IGF2, IGFBP2 and IGFBP5*) [357, 368, 371-373].

An important challenge of studying the genetics of T1D, where there is accumulation of weak effects at many different loci, is to identify the loci themselves and determine

175

to which a particular locus influences disease susceptibility. Evidence for oligogenic inheritance of T1D has been provided in BB rats [374-376] and LETR rats [377] with one or two non-MHC (equivalent to HLA) genes interacting with MHC susceptibility. In humans, there are two reported examples of a single gene that increases the risk for diabetes independently of *HLA*. The first was the autoimmune regulator gene (*AIRE*) on chromosome 21. AIRE mutations cause autoimmune polyendocrine syndrome type 1 (APS-1), which includes insulin-dependent diabetes [378]. The second example was found in a large family of Bedouin with multiple affected individuals with T1D. this study led to the identification of locus IDDM17 on chromosomal region 10q25, [379]. Interestingly, no evidence of linkage to *HLA* was found in this family, possibly due to a markedly increase of high-risk alleles.

1.4 Preliminary results.

In Colombia, DM occurs at a frequency of 7% in both sexes, and the prevalence of the intolerance to the oral test glucose has been found to be 5% in men and 7% in women [380]. In 1996, Montoya *et al.* studied *HLA* alleles in T1D patients from Antioquia–Colombia, and identified two susceptibility haplotypes: *DRB1*0301-*DQA1*0501-*DQB1*0201 with a frequency of 50% in cases versus 12.5% in controls; and DRB104-*DQA1*03-*DQB1*0302 with a frequency as high as 53.8% in cases and 28.6 in controls [381].

More recently, Uribe and colleagues (2004) analyzed a large Colombian family presenting with T1D, where 11 affected individuals were identified in two generations (Figure 8.1). An affected father transmitted the disease to 8 daughters. Two sons were

unaffected. Two other affected individuals were also identified in another branch of

the family. There was evidence of dominant inheritance which was consistent with

both an autosomal gene and x-linkage [382] (see Figure 8.1).

In the first study, we evaluated linkage to several candidate loci, including *IDDM1,*

IDDM2, IDDM4, IDDM7, IDDM11, IDDM12 and *IDDM13*. In addition DXS1068

was also analyzed. Furthermore, we also evaluated markers every 10 cM (Linkage

panel set v.2, Applied Biosystems) on chromosomes 2 and 11, as part of the beginning

of a genome scan. AntiGAD (Anti Glutamic Acid Decarboxilase) antibodies were

evaluated in serum from patients where this was available. Linkage analsyis showed

lod scores of either zero or exclusion ($\leq$-2) in all the candidate loci tested and

interestingly also showed a peak of lod score of 2.31 for telomeric marker D2S319 at

θ=0, at chromosomal region 2p25. AntiGAD antibodies were positive in two late onset

patients, confirming an autoimmune process (T1AD) in this family and ruling out

other possible diagnoses such as MODY [382].
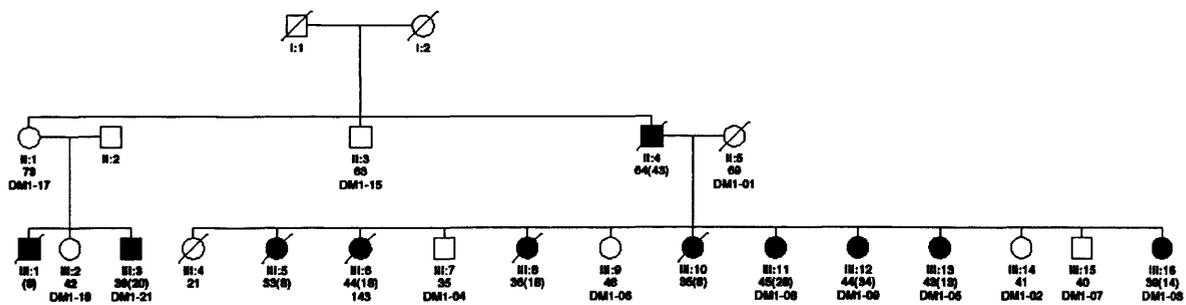
Figure 8.1. Colombian family with T1D. Below each individual is shown current age/age at dead (age at diagnoses) when available. Sample code is shown for available individuals. Individuals III:11 and III:12 were positive for antiGAD antibodies.

1.5 Aims

The aim of this chapter was to further persue to linkage to marker D2S319 on chromosome 2. Also, to evaluate whether a haplotype was segregating with the disease in this Colombian family or if, on the contrary, random allele segregation was observed between the disease and markers at 2p25. Identification of a haplotype segregating with the disease would add evidence supporting the original hypothesis of the presence of a novel susceptibility T1D locus on chromosomal region 2p25 [382]. Identification of the actual mutant gene might allow the identification of the molecular processes that lead death of cells β islet in this family, adding knowledge to the complex set of cellular events that are presumed to participate in triggering this disease.

2. Patients and Methods

2.1 Patients

Patients were diagnosed as T1D according to international criteria [325]. The parents and 13 children compose the main branch of this pedigree (8 alive and 5 dead). The father was diagnosed at age 43. He was never obese and got progressively thin until six months after diagnosis. He only responded to treatment with insulin. He died due to an insulinic comma (double insulin dose) at age 64. Of 13 children, 8 were diabetic (all women), 7 required insulin from the stage of the initial diagnosis and one (III:11,

Figure 8.1) started treatment with diet for the first two years. For the following three years she required oral hypoglyceamiants, and finally has required treatment with insulin.

Four affected children (III:5, III:6, III:8 and III:10) died due to complications directly related to Diabetes. Three of them (III:5, III:8 and III:10, Figure 8.1) died because of chronic renal insufficiency, on average 22.6 years after diagnosis. Individual III:6 died due to a coronary disease of early onset. A living patient (III:13) developed chronic renal insufficiency and required transplantation [382].

2.2 Power simulation

In order to estimate the power to detect linkage, a simulation was carried out including the following assumptions: Autosomal dominant inheritance with a penetrance of 98.5% [382] and 92%; an allele frequency of the mutated gene of 0.0001, and a marker locus with four alleles of equal frequency. 100 replicas were run using the SLINK program [55].

2.3 Microsatellite marker typing

Ten polymorphic DNA markers were tested in the vicinity of D2S319. These markers were D2S2980, D2S323, D2S2393, D2S2245, D2S304, D2S205, D2S2166, D2S2891, G62096 and D2S2133 (Figure 2). They all were amplified using 30 cycles of 94° C 30 sec, 52-60° C 15 sec, 72° C 30 sec, and a final elongation step of 72° C during 10 min. Markers were labelled with different fluorescent tags in the F-oligo to allow pooling before loading them onto an ABI-377 (Table 8.2).

2p25
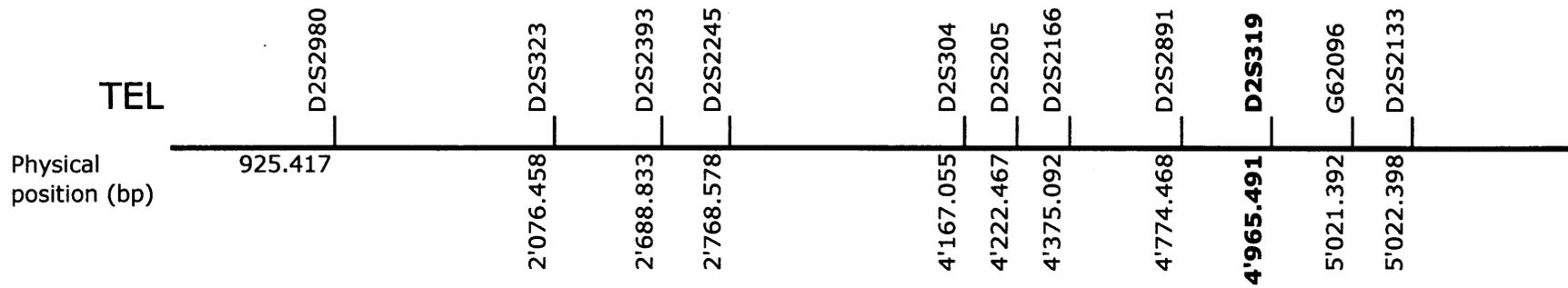
2p25
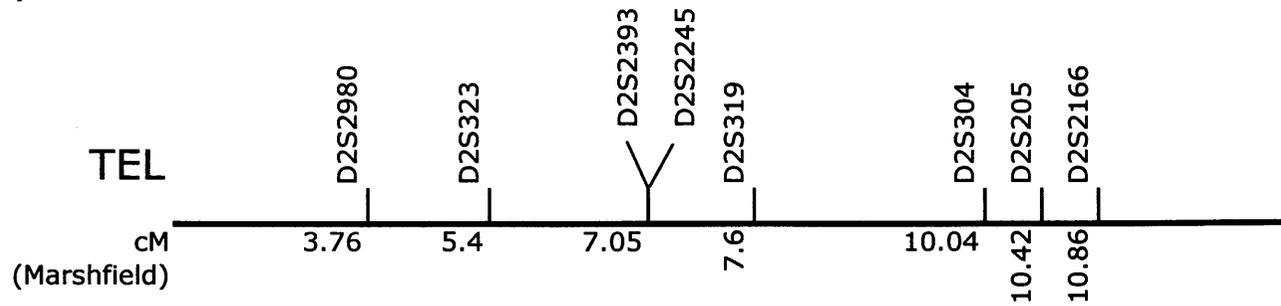
Figure 8.2. Physical and genetic map of markers tested in this study. Above is shown the physical map for 11 markers including D2S319. Below is shown the genetic map for eight markers including D2S319. Genetic positions of markers D2S2891, G62096 and D2S2133 were not available so had to be excluded from the genetic map. TEL= Telomere.

Marker	Annealing To	Label	ASR
D2S2980	55° C	HEX	232-284
D2S323	55° C	FAM	177-193
D2S2393	55° C	FAM	244-270
D2S319	55° C	FAM	116-138
D2S304	52° C	FAM	177-189
D2S205	55° C	HEX	143-171
D2S2166	55° C	HEX	208-250
D2S2891	57° C	FAM	141
G62096	55° C	HEX	145
D2S2133	60° C	FAM	98-102

Table 8.2. Marker loci typed around marker locus D2S319 at 2p25 in a Colombian family with T1D. ASR=Average size range. These markers were evaluated at different times, so that they were pooled at least in pairs.

2.4 Linkage and Haplotype analysis

Linkage analysis was carried out using same assumptions as for the power simulation.

Allele frequencies were assumed as 1/n (n=number of alleles) and a phenocopy rate of

zero. Two-point lod scores were calculated with MLINK [44]. Haplotypes were

reconstructed using simwalk2 [49].

3. Results

3.1 Power simulations

Table 8.3 shows the results of the power simulation. At both penetrances an average

lod score of about 2 is expected. The simulation also indicated that approximately 29%

chance of finding a $Z \geq 3$ in this family was possible, when penetrance 0.985 was used.

Using penetrance of 0.92 a 25% chances of finding a $Z \geq 3$ was obtained.

181

	Expected Lod score (Z) at $\theta= 0$		Probability of reaching $Z \geq$ to:		
Penetrance	Average	Zmax	1	2	3
0.985	2.39	3.57	74	57	29
0.92	1.98	3.14	76	60	25

Table 8.3. Power simulation for family with T1D.

3.2 Two-point analysis

Table 8.4 shows the two-point lod scores for markers according to genetic map shown

in Figure 8.2. It can be noticed that marker D2S319 presented the highest lod score

(Z= 2.53) at $\theta= 0$. The second highest lod score (Z= 2.0) was found for marker

D2S2393 at $\theta= 0.072$. Markers D2S323 and D2S2245 presented maximum lod scores

close to zero (Z=0.4 and Z= 0.37, respectively) because they were uninformative for

linkage. The remaining markers were informative for linkage and their Z max was

over 0.85 at different recombination fractions. These results support the notion of a

locus in T1D at 2p25.

The locus by locus analysis was focused on the main branch of the pedigree (Figure

8.3).

For marker D2S2980 all affected individuals received an allele 1 from II:4, with

exception of the affected individual II:12 who is a recombinant and received allele 2.

Consistently, Z<-2 ($\theta= 0$) was obtained, which indicates exclusion (Table 8.4).

For marker D2S323 lack of informativity of individual II:4 explains the low lod score.

As both affected and unaffected children of II:4 received an allele 3, it was possible to

infer that II:4 was homozygous 3/3 for this marker (Figure 8.3). A recombination

seems to have occurred in individual III:12 in the interval D2S323-D2S2393.

However, since marker locus D2S323 appeared as homozygous, it is not possible to

determine whether this is between D2S323-D2S2393 or D2S323-D2S2980 (see Figure 8.3). This indicates that the telomeric border of the candidate region is marker D2S2980.

Locus	Recombination fraction, θ						θ max	Z max
	0	0.1	0.2	0.3	0.4	0.5		
D2S2980	-5.23	0.8	0.86	0.62	0.25	0	0.16	0.89
D2S323	0.4	0.28	0.18	0.08	0.02	0	0.001	0.4
D2S2393	-1.52	1.98	1.63	1.09	0.44	0	0.072	2
D2S2245	0.37	0.27	0.17	0.09	0.02	0	0.001	0.37
D2S319	2.53	2.17	1.68	1.1	0.44	0	0.001	2.53
D2S304	-7.52	0.78	0.83	0.58	0.21	0	0.159	0.86
D2S205	-6.97	1.03	1.03	0.72	0.28	0	0.144	1.1
D2S2166	-3.19	1.48	1.23	0.8	0.3	0	0.086	1.48

Table 8.4. Two-point lod scores for markers flanking D2S319 locus. Analysis included penetrance= 0.985

Marker locus D2S2393 showed a Zmax= 2.0 at θ= 0.072 (Table 8.4). Every affected child in the large nuclear family received an allele 3 at this marker locus from II:4. In contrast, the unaffected individuals received an allele 1 from him. In contrast to the observation in the large branch where allele 3 was segregating with the disease, affected individual III:3 received an allele 1 from his mother. The lack of allele sharing caused the Z maximum to be at $\theta \neq 0$ (Table 8.4). If we excluded individuals III:1, III:2 and III:3 from the analysis, the highest lod score is then obtained at θ= 0 (Z= 1.76, Table 8.5). For marker D2S2245 there was no linkage information since this marker was monomorphic. Everybody in the pedigree was homozygous 2/2 with exception of II:3, who was 1/2.
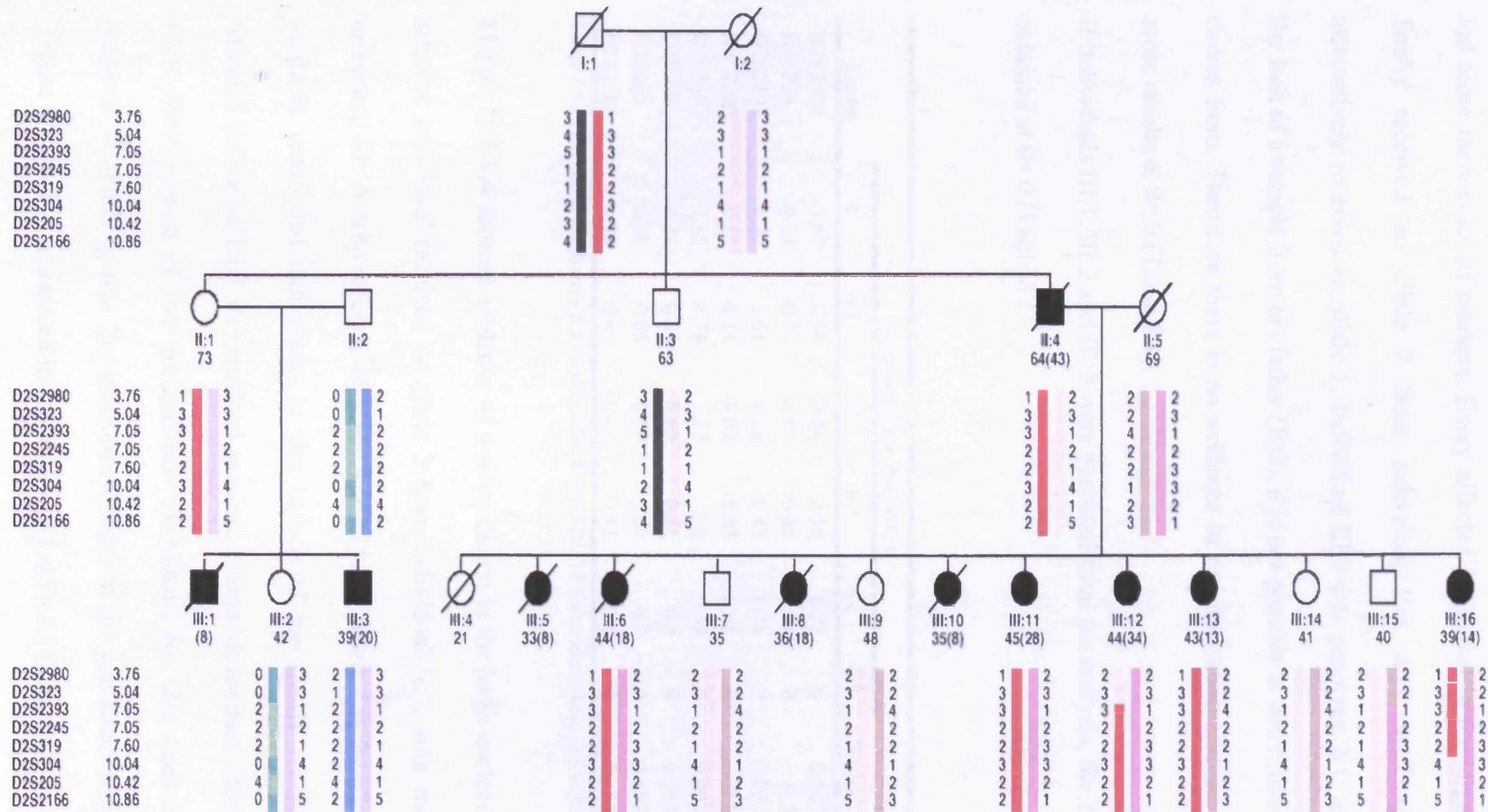
Figure 8.3. Pedigree of Colombian family with T1D. Haplotypes were inferred for individuals I:1, I:2, II:2 and II:4 since no sample was available from them. Below each symbol is shown current age/age at dead (age at diagnosis) when available. Haplotypes for markers according to genetic map in Figure 1 are shown. Two key recombinant events were observed in III:12 and III:16 allowing establish a candidate region between marker loci D2S2980 and D2S304. III:11 and III:12 were positive for antiGAD antibodies.

For marker D2S319, a lod score of 2.53 was obtained, at θ= 0, representing the highest lod score for this set of markers. Every affected individual in the large branch of the family received an allele 2 from individual II:4, and unaffected individuals alternatively received an allele 1. Individual III:3 was genotype 2/1, and considering the lack of a sample from his father (II:2), it is not possible to determine where allele 2 comes from. Therefore there is no evidence of recombination and the maximum lod score results at θ= 0 (Table 8.4).

If individuals III:1, III:2 and III:3 were excluded from the analysis, the Z max still is obtained at θ= 0 (Table 8.5).

Locus	Recombination fraction, θ						θ max	Z max
	0	0.1	0.2	0.3	0.4	0.5		
D2S2980	-3.67	0.98	0.86	0.58	0.22	0	0.107	0.98
D2S323	-0.33	-0.17	-0.09	-0.04	-0.01	0	0.5	0
D2S2393	1.76	1.61	1.29	0.87	0.36	0	0.001	1.76
D2S2245	-0.25	-0.14	-0.07	-0.03	-0.01	0	0.5	0
D2S319	2.12	1.78	1.37	0.9	0.37	0	0.001	2.12
D2S304	-4.24	0.65	0.69	0.51	0.2	0	0.157	0.71
D2S205	-4.28	0.65	0.69	0.51	0.2	0	0.157	0.71
D2S2166	-4.24	0.65	0.69	0.51	0.2	0	0.157	0.71

Table 8.4. Two-point lod scores for markers flanking D2S319 locus excluding individual II:1's family.

Marker D2S304 showed evidence of a recombinant in the large nuclear family. Every affected individual received an allele 3 from individual II:4, with the exception of individual III:16, who received an allele 4 from him. Allele 4 is the allele that all four available unaffected individuals in this branch of the family received from II:1. Affected individual III:3 also received an allele 4 from his mother. These observations allow identification of two recombinant individuals for this marker locus in the pedigree, establishing that the centromeric limit in the candidate region is D2S304 (Figure 8.3), and its associated lod score is -7.52 at θ= 0 (Table 8.4).

When analyses were done only for the main nuclear family and using a penetrance= 0.92, the lod scores were slightly higher than when using penetrance= 0.985. However, more important than the lod score itself is the recombination fraction at which those lod scores were obtained. So, maximum lod scores at q= 0 were obtained for marker D2S2393 and D2S319 (Table 8.5). Thus, exclusion of linkage was obtained for markers D2S2980, D2S304, D2S205 and D2S2166 (Table 8.5). The remaining four markers formed a consistent disease-associated haplotype.

3.3 Haplotype and Multipoint analysis

In the main nuclear family, all affected individuals carry a disease-associated haplotype characterized by alleles 3-3-2-2 at markers D2S323-D2S2393-D2S2245-D2S319. Among the unaffected individuals, II:1 is the only carrier of this haplotype. II:1 has two affected children, one of them (III:3) seems to correspond to a phenocopy since he does not carry the disease-associated haplotype (Figure 8.3).

Figure 8.4 shows the multipoint analysis for the complete pedigree. Lod scores were positive for the interval D2S323-D2S319 (Zmax= 0.148 in the interval D2S323-D2S2393), providing little evidence in favour of linkage.
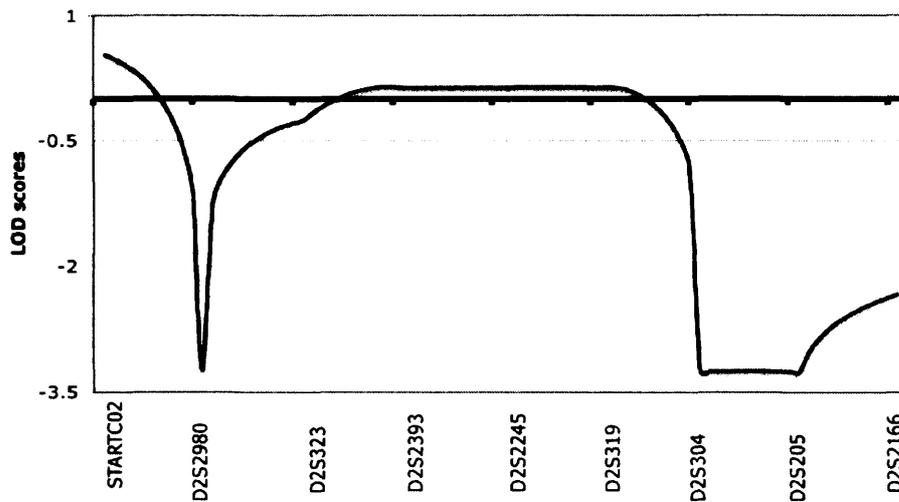
Figure 8.4. Multipoint analysis for markers according to the genetic map shown in Figure 2.

However, further analysis excluding II:1's family resulted in a raise of the multipoint

lod scores in the interval that corresponds to the candidate region (i.e. for the interval

D2S2980-D2S304, Figure 8.5). Its maximum lod score is 1.146 At the interval

D2S2393-D2S319.
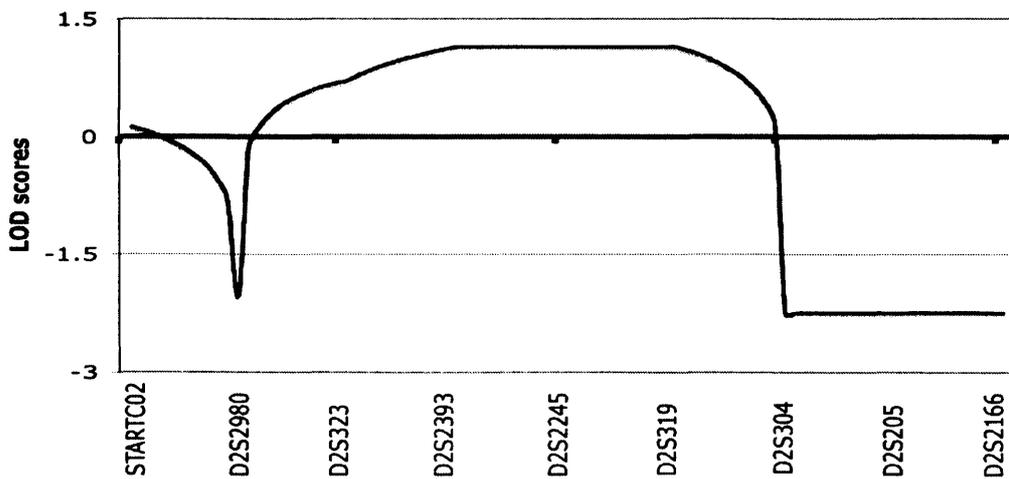
Figure 8.5. Multipoint analysis excluding individual II:1's family. Penetrance= 0.92.

4. Discussion

The purpose of this chapter was to evaluate whether a haplotype was segregating with

T1D in 2p25 in order to confirm the findings reported by Uribe and colleagues [382].

187

Although this family presents with a high penetrance, resembling MODY, autosomal dominant disease with presence of antiGAD antibodies in addition to severity and disease related complications points to autoimmune T1D rather than other diagnoses. Evidence of autommunity was found in individuals III:11 and III:12, who presented high titers of antiGAD antibodies.

Analysis of seven additional polymorphic markers linked to D2S319 led to the demonstration of a maximum two-point lod score of 2 for marker loci D2S2393 (Table 8.4), suggesting that true linkage of 2p25 to T1D is present in this family. A higher lod score than 2.31 for D2S319 was obtained in our analyses because we assigned a lower value for the suggested population allele frequency of the mutation, since all indications are that this is an extremely rare allele. It was decreased to 0.0001, from 0.003 in the model used by Uribe and colleagues [382].

Moreover, a disease-associated haplotype characterized by alleles 3-3-2-2 at markers D2S323-D2S2393-D2S2245-D2S319 (Figure 8.3) was identified. Marker D2S2980 corresponds to the telomeric limit identified in this study. Similarly, centromeric marker to D2S319 (i.e. D2S304) corresponds to the centromeric limit of the candidate region identified here (Figures 8.3). The candidate region is 3.24 Mb long and spans 6.28 cM. Typing extra markers (e.g. STRs or SNPs) in between D2S319-D2S304 will result in a narrower centromeric limit of the candidate region. Narrowing the telomeric limit could also be possible. Considering that marker D2S323 was not informative for linkage, the recombination delimiting the telomeric limit can still be more precisely identified. Thus, by typing extra markers in the intervals D2S2980-D2S323 and D2S323-D2S22393, further refining of the region could be achieved. Three already typed markers that have not got a reported genetic position are shown in

Table 8.6. Even so, two of these are non informative for linkage, since they were monomorphic, are presented here mainly to prevent that somebody else type them again.

Individual	D2S2891		G62096		D2S2133	
II:1	1	1	1	1	2	2
II:3	1	1	1	1	0	0
II:5	1	1	1	1	1	2
III:2	1	1	1	1	1	2
III:3	1	1	1	1	1	2
III:6	1	1	1	1	1	2
III:7	1	1	1	1	2	2
III:9	1	1	1	1	2	2
III:11	1	1	1	1	1	2
III:12	1	1	1	1	1	2
III:13	1	1	1	1	2	2
III:14	1	1	1	1	2	2
III:15	1	1	1	1	1	2
III:16	1	1	1	1	1	2

Table 8.6. Genotypes for three extra markers with genetic positions not available. These markers were not used in multipoint analysis because of the inability to distinguish between centromeric and telomeric flanking markers to D2S319 in the genetic map.

The age at onset shows a large range in this family, 8-43 years. Considering that the oldest person diagnosed was II:4 (onset at 43 years old) and that the other affected individuals presented age at onset much younger than him (8-34), there is the possibility that anticipation is involved. Thus, expansion (of a trinucleotide repeat array) is a possibility, since when the array increases over certain size, the disease will appear earlier. For example, in Huntington disease it has been documented that the larger the expansion, the younger the age at onset [149].

Oligogenic inheritance has previously been reported for two autoimmune disorders that usually are thought of as polygenic [378, 379, 383]. In one of these studies a single locus on chromosome 10 (*IDDM17*) provided most of the genetic component in the

189

family studied. No evidence of linkage to *HLA* was found [379]. The present results are similar to those of Verge and colleagues (1998) [379], since only linkage to 2p25 has been identified and no linkage to any of four STRs in *HLA*-class II was observed [382].

Looking at candidate genes in this region, it appears that at least one is a very good candidate. This is *TPO*, which encodes the enzyme thyroid peroxidase. Antibodies against TPO in T1D patients have been reported [384]. TPO gene has 17 exons and expands 129.16 kb and locates just centromeric to D2S2980. Mutations in *TPO* have been reported in families with total iodine organification defect (TIOD), which leads to congenital hypothyroidism [385]. *CTLA4* (*IDDM12*), which has been associated with T1D [357-360] and has also been associated with thyroid disorders (Graves disease [361] and Hashimoto thyroiditis [362]). So, in a similar way to *CTLA4*, it will not be surprising if *TPO* mutations were associated to T1D in the family presented here.

ACP1 (Erythrocyte acid phosphatase) is located just telomeric of D2S2980 (outside the candidate region). Genotypes for low activity of this enzyme have been associated with higher glyceamic levels in T1D [386].

In conclusion, a haplotype segregating with T1AD has been identified in an extended Colombian family. This haplotype spans 3.4 Mb on chromosomal region 2p25. Further fine mapping needs to be done before engaging in evaluating candidate genes in the region. However, *TPO* gene seems strikingly good candidate. This family is the third evidence of oligogenic inheritance of T1D.

CHAPTER NINE

GENERAL DISCUSSION

This thesis aimed to make progress on the genetic characterization of several Mendelian forms of complex disorders, including Parkinson's disease, Generalized Epilepsy with Febrile Seizures Plus [GEFS+], Batten's disease, Type 1 Diabetes [T1D], Multiform Movement Disorder [MMD] and Abdominal Epilepsy [AE]. All individuals studied here are from Antioquia – Colombia.

In Parkinson disease (PD) –Chapter 3-, progress was made in three aspects. First, three affected individuals (V;1, V:12 and VI:2, Figure 3.4), who originally were identified as sporadic cases and only carriers of mutation c.736 G>A in *PARK2*, were found to be members of the same extended pedigree from the same town (Peque). In addition, family PJF3, in which this mutation was originally reported [36] was also found to be a branch of the same extended pedigree in Peque. Haplotype analysis of the extended pedigree at the *PARK2* locus indicated that the three affected individuals shared a second haplotype (Table 3.5) that eventually could be associated with a second mutation segregating in the family. An exon 3 duplication was found in individuals V:1 and VI:2. Individual V:12 was not tested and based on the haplotype shared I assumed him to be also carrier of such a duplication. This observation provides evidence against those reports claiming that being carrier (heterozygous) of *PARK2* mutations is a genetic risk factor for being affected by PD [65, 83, 110, 113, 118, 128, 129]. An important factor for those studies not to have found second mutations (e.g. deletions of duplications) is that they have mainly used SSCP and sequence, in addition to lacking large pedigrees where haplotype analysis could provide sights of a second mutation in those affected individuals with only one mutation identified.

A second aspect is, when analysing four intragenic STRs in a group of 23 sporadic cases, it was found that association to *PARK2* was detected for alleles at marker loci D6S1550 and D6S980 (Table 3.8). Furthermore, when the analysis included the

haplotype made of the four marker loci, an increase in the statistical significance was observed. None of these 23 cases carries the c.736 G>A mutation; only one of them is a carrier of 321-322insGT mutation (previously reported in Antioquia) and judging from the haplotype analysis none of these carry the exon 3 duplication. However, these results suggest that mutations that went undetected by SSCP and sequencing will eventually be found in this set of sporadic cases in PARK2 gene.

The third aspect was the fact that the Spanish family where c.736 G>A was reported, showed a completely different chromosome from the one in the Colombian families. This observation suggests that recurrent mutation could be a good explanation; however, since nobody knows how old is the Spanish chromosome carrying the c.736 G>A mutation, and considering that recombination could have taken away the shared regions with the Colombian chromosomes, it is necessary to evaluate markers more closely linked to the mutation (SNPs) and if still no haplotype is shared, then the hypothesis of recurrent mutation would be favoured.

In Multiform Movement Disorder (MMD) –Chapter 4- study of one family with the disorder and based on the most appropriate differential diagnoses, several candidate genes were excluded, including PARK1, PARK2, PARK8, PANK2, and PARK8. Brain pathology in one of the deceased patients (individual III:6) revealed massive iron accumulation in various parts of her brain. This observation was similar to the reports by Curtis and colleagues [159] and evaluation of linkage to the gene FTL, reported by them provided no evidence in favour of linkage. Thus, considering the power limitations of the pedigree studied, a genome wide scan was performed in this family hoping to find linkage to a region where a clear candidate gene involved in iron metabolism lies. Results indicated that there are two candidate genes to be evaluated since both of them are associated with iron metabolism. These are the Neuroglogin

gene (*NGB*) on chromosome 14 and the Ferritin Heavy Chain gene (*FTH1*) on chromosome 11. For the case of NGB, one unaffected individual (III:3, Figure 4.4) carrying the disease associated haplotype was identified. This observation could be explained by incomplete penetrance or also by random segregation (i.e. no linkage to the disorder).

FTH1 seems a very plausible candidate, since a previous report linked the light chain (FTL) of the ferritin to a disorder presenting with pathological features very alike to those presented by the family studied by us [159]. However, two-point analysis showed that the maximum lod score (Marker D11S4151, Zmax= 0.97) is obtained at θ= 0.1. Typing markers more tightly linked to *FTH1* would evaluate whether this gene is the one implicated in the disorder in this family.

In GEFS+ -Chapter 5- a set of six families were analyzed for candidate genes/loci involved in GEFS+ and Febrile Seizures (FS). From this set, three families (Families 1, 6 and 7) showed linkage to markers at 2q23-24, where *FEB3/SCN1A/SCN2A* genes lie. In Family 1, a combination of SSCP and base sequence of *SCN1A* gene revealed a c.5213 A>G mutation, which lead to the D1742G substitution. This mutations lies in the pore forming loop of domain IV of the protein and its conservation through related genes might explain the severity of its phenotype, which contrasts with the mild severity and penetrance reported for GEFS+ [215, 387]. Families 6 and 7 do not carry this mutation and are being further evaluated in *SCN1A* gene.

Family 5 showed linkage to GABR cluster. This cluster involves *GABRG2*, *GABRA1*, *GABRA6 and GABRB2* genes. Of these, only *GABRG2* has been associated to GEFS+ [243 2002, 246]. Sequencing of the coding region of this gene failed to identify disease causing mutations. In addition, an affected individual appeared as recombinant for

marker SHGC-132760, confirming exclusion of *GABRG2* in this family (Figures 5.2 & 5.6). Evaluation of *GABRA1*, *GABRA6* and GABRB2 genes is currently underway. Families 3 and 4 seemed not to be linked to any of the genes/loci evaluated. However, Family 3 requires more evaluation in locus *FEB2* since exclusion has not been obtained; but at the same time very low lod scores have been obtained in this family since all affected individuals carry a haplotype (Figure 5.7), which is also carried by unaffected individuals as well as by both parents in the family.

In Batten disease – Chapter 6- an extended family was studied. Three affected individuals in two consanguineous nuclear families from the same pedigree presented with clinical Batten disease (i.e. Juvenile onset). However, pathological analysis of skin biopsy revealed inclusions consisting of fingerprint profiles, often condensed, and occasionally associated with lipid droplets, suggestive of a variant NCL. Genetic analysis excluded *CLN3*, since both affected children appeared heterozygous for STRs linked to this gene (Figure 6.3). SSCP and sequencing confirmed exclusion. Similarly, typing markers linked to *CLN1*, *CLN2*, *CLN6* and *CLN8* genes resulted in exclusion (Figure 3, Table 4). Evaluation of markers linked to *CLN5* showed that only affected individuals were homozygous for a haplotype characterized by alleles 1-1-1 at markers D13S156-D13S162-D13S1306. These markers flank the *CLN5* gene. Sequence analysis of *CLN5* revealed a transition of G to A at position 1627 in exon 2 (c.1627 G>A; Figure 5A) and was not found in any of 29 controls. This mutation leads to the substitution Arg112Hys and was found to be highly conserved among vertebrates (Figure 6.5B). c.1627 G>A is the first *CLN5* mutation reported outside Europe and associated with a Juvenile phenotype rather than late infantile. Previous reports have documented four *CLN5* mutations in patients from Finland, Sweden and Holland [269, 303], all of them presenting a late infantile pheotype. Evaluation of the molecular effect

195

of this mutation should facilitate an understanding of disease mechanisms, including the formation of lysosomal deposits.

In Abdominal Epilepsy (AE) –Chapter 7- a large family presenting with a disorder that, although clinical characterization is underway, the most appropriate diagnosis is Abdominal epilepsy (AE). AE has never been reported as a familial form; however here we are reporting a multigenerational pedigree including 15 affected individuals. Both sexes are affected about equally (53% females) and clear autosomal dominant inheritance was observed. Having no clear candidate genes to evaluate a genome scan was conducted. Two-point analysis revealed a Zmax= 4.18 at marker locus D8S512. Fine mapping of the region resulted in a maximum multipoint lod score of 4.63 in the interval D8S1767-D8S1192. This region extends 1.3 Mb, which corresponds to 1.85 cM. Interestingly, this locus (i.e. AE) is overlapped by *FEB1* locus [215]. It could be the case that both disorders are due to mutations in the same gene, located in our candidate region. Despite the disorders are types of epilepsy they are also very different phenotypes. For example, none of the patients in our family presented febrile seizures, which is the phenotype linked to *FEB1*. In our candidate region there are three genes identified, DEPDC2, VEST1 and the hypothetical protein DKFZp547E186. Uncovering the mutant gene in this family could involve evluation of functional domains of other proteins implicated in epilepsy to see whether they have some degree of homology with DEPDC2, VEST1 or DKFZp547E186. If a level of homology is identified, then sequence of the gene implicated would reveal the mutation responsible for the disorder in this family. If no homology was identified, sequencing of all the three genes would be the next step. This is the first locus reported for AE so far.

Finally, in Type 1 Diabetes [T1D] –Chapter 8- an extended pedigree with the disease was studied. Here, I have further evaluated a novel locus that was identified by Uribe and colleagues [382]. This family was analyzed for several candidate loci, including *IDDM1*, *IDDM2*, *IDDM4*, *IDDM7*, *IDDM12*, *IDDM13* and *DXS1068*. Additonally, markers across chromosomes 2 and 11 were typed at an average density of 10 cM (Linkage panel set v.2, Applied Biosystems) as part of the beginning of a genome scan. We obtained lod scores of either zero or exclusion in all candidate regions evaluated. However, one peak of Z= 2.31 at θ= 0 for telomeric marker D2S319 was obtained. AntiGAD antibodies were positive in two individuals (III:11 and III:12), which are a marker of autoimmunity in T1D. Here I have examined seven additional markers around D2S319. A maximum two-point lod score of 2 for marker loci D2S2393 was obtained (Table 8.4), suggesting that true linkage of 2p25 to T1D is present in this family. A haplotype characterized by alleles 3-3-2-2 at markers D2S323-D2S2393-D2S2245-D2S319 was identified (Figure 3). This candidate region is 3.24 Mb long and spans 6.28 cM. Based on the observation that *CTLA4* is associated with both T1D [357-360] and thyroid disorders (Graves disease [361] and Hashimoto thyroiditis [362]), the *TPO* gene which has been associated with thyroid disorders [384, 385] is a potential candidate gene since it lies in our candidate gene. Nevertheless, the candidate region can still be further narrowed down in the telomeric region. The identification of this locus represents the second evidence of oligogenic inheritance in T1D.

In conclusion, various different findings have been achieved in this thesis work. Despite having ascertained my population sample from a homogeneous population, evidence of locus and allelic heterogeneity was evident. Locus heterogeneity was clear in GEFS+, since the set of families were linked to different loci. Allelic heterogeneity

was evident in Parkinson disease, where different mutations explain the disease in some individuals and the great majority of sporadic cases, although linked to *PARK2* are presenting mutations that went undetected.

Considering that the families and individuals studied here are affected by high penetrant mutations and that they originate in a homogeneous population in addition to the rarity of the disorders (i.e. Early-onset Parkinson disease), a high level of heterogeneity was found. Therefore, if studying low penetrant mutations and quite prevalent disorders (complex traits) much more heterogeneity can be expected. This emphasises the importance of ascertaining population samples from homogeneous populations, specially when studying complex disorders, even for Mendelian forms when one single family lacks of the power to detect significant linkage on its own.

References

1. Haines, J. & Perick-Vance, M. Approaches to Gene Mapping in Complex Human Diseases (Wiley-Liss, New York, 1998).
2. Risch, N. & Merikangas, K. The future of genetic studies of complex human diseases. Science 273, 1516-7 (1996).
3. Lander, E. S. The new genomics: global views of biology. Science 274, 536-9 (1996).
4. Risch, N. Searching for genes in complex diseases: lessons from systemic lupus erythematosus. J Clin Invest 105, 1503-6 (2000).
5. Patil, N. et al. Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. Science 294, 1719-23 (2001).
6. Gabriel, S. B. et al. The structure of haplotype blocks in the human genome. Science 296, 2225-9 (2002).
7. Stephens, J. C. et al. Haplotype variation and linkage disequilibrium in 313 human genes. Science 293, 489-93 (2001).
8. Glatt, C. E. et al. Screening a large reference sample to identify very low frequency sequence variants: comparisons between two genes. Nat Genet 27, 435-8 (2001).
9. Botstein, D. & Risch, N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nat Genet 33 Suppl, 228-37 (2003).
10. Kerem, E. et al. Clinical and genetic comparisons of patients with cystic fibrosis, with or without meconium ileus. J Pediatr 114, 767-73 (1989).
11. Riordan, J. R. et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 245, 1066-73 (1989).
12. Lander, E. S. & Botstein, D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. Science 236, 1567-70 (1987).
13. Gschwend, M. et al. A locus for Fanconi anemia on 16q determined by homozygosity mapping. Am J Hum Genet 59, 377-84 (1996).
14. Saar, K. et al. Localisation of a Fanconi anaemia gene to chromosome 9p. Eur J Hum Genet 6, 501-8 (1998).
15. Waisfisz, Q. et al. The Fanconi anemia group E gene, FANCE, maps to chromosome 6p. Am J Hum Genet 64, 1400-5 (1999).
16. Weber, J. L. & May, P. E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44, 388-96 (1989).
17. Jeffreys, A. J., Royle, N. J., Wilson, V. & Wong, Z. Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. Nature 332, 278-81 (1988).
18. Jakupciak, J. P. & Wells, R. D. Genetic instabilities in (CTG.CAG) repeats occur by recombination. J Biol Chem 274, 23468-79 (1999).
19. Richard, G. F., Dujon, B. & Haber, J. E. Double-strand break repair can lead to high frequencies of deletions within short CAG/CTG trinucleotide repeats. Mol Gen Genet 261, 871-82 (1999).

20. Riley, D. E. & Krieger, J. N. Short tandem repeat (STR) replacements in UTRs and introns suggest an important role for certain STRs in gene expression and disease. Gene 344, 203-11 (2005).

21. La Spada, A. R., Wilson, E. M., Lubahn, D. B., Harding, A. E. & Fischbeck, K. H. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Nature 352, 77-9 (1991).

22. Huntington's Collaborative Research group, T. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. Cell 72, 971-83 (1993).

23. Weber, J. L. Human DNA polymorphisms and methods of analysis. Curr Opin Biotechnol 1, 166-71 (1990).

24. Bruford, M. W. & Wayne, R. K. Microsatellites and their application to population genetic studies. Curr Opin Genet Dev 3, 939-43 (1993).

25. Landegren, U., Nilsson, M. & Kwok, P. Y. Reading bits of genetic information: methods for single-nucleotide polymorphism analysis. Genome Res 8, 769-76 (1998).

26. Wilson, C., Tiwana, H. & Ebringer, A. Molecular mimicry between HLA-DR alleles associated with rheumatoid arthritis and Proteus mirabilis as the Aetiological basis for autoimmunity. Microbes Infect 2, 1489-96 (2000).

27. Eaves, I. A. et al. The genetically isolated populations of Finland and sardinia may not be a panacea for linkage disequilibrium mapping of common disease genes. Nat Genet 25, 320-3 (2000).

28. Shifman, S. & Darvasi, A. The value of isolated populations. Nat Genet 28, 309-10 (2001).

29. Vaessen, N. et al. A genome-wide search for linkage-disequilibrium with type 1 diabetes in a recent genetically isolated population from the Netherlands. Diabetes 51, 856-9 (2002).

30. Alvarez, V. Poblamiento y poblacion en el valle de Aburrá y Medellín, 1541-1951. In: JO (ed) Historia de Medellin, Suramerica, Medellin (1996).

31. Bravo, M., Valenzuela, C. & Arcos-Burgos, O. Polymorphisms and phyletic relationships of the Paisa community from Antioquia (Colombia). Gene Geogr 10, 11-17 (1996).

32. Carvajal-Carmona, L. G. et al. Strong Amerind/white sex bias and a possible Sephardic contribution among the founders of a population in northwest Colombia. Am J Hum Genet 67, 1287-95 (2000).

33. Carvajal-Carmona, L. G. et al. Genetic demography of Antioquia (Colombia) and the Central Valley of Costa Rica. Hum Genet 112, 534-41 (2003).

34. Lendon, C. et al. E280A PS-1 mutation causes Alzheimer's disease but age of onset is not modified by ApoE alleles. Hum Genet 10, 186-195 (1997).

35. Lopera, F. et al. Clinical features of early-onset Alzheimer disease in a large kindred with an E280A presenilin-1 mutation. Jama 277, 793-9 (1997).

36. Pineda-Trujillo, N. et al. A novel Cys212Tyr founder mutation in parkin and allelic heterogeneity of juvenile Parkinsonism in a population from North West Colombia. Neurosci Lett 298, 87-90 (2001).

37. Ott, J. Analysis of Human Genetic Linkage (The Johns Hopkins University Press, Baltimore, 1999).

38. Morton, N. E. Sequential tests for the detection of linkage. Am J Hum Genet 7, 277-318 (1955).

39. Haldane, J. & Smith, C. A new estimate of the linkage between the genes for colour-blindness and haemohilia in man. Ann Eugen 14, 10-31 (1947).

40. Wald, A. Sequential Analysis (Wiley, New York, 1947).

41. Barnard, G. Statistical inference. J Roy Statist Soc, 115-139 (1949).

42. Morton, N. E. Significance levels in complex inheritance. Am J Hum Genet 62, 690-7 (1998).

43. Terwilliger, J. & Ott, J. Handbook of Human Genetic Linkage (Johns Hopkins University Press, Baltimore, 1994).

44. Lathrop, G. M., Lalouel, J. M., Julier, C. & Ott, J. Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci U S A 81, 3443-6 (1984).

45. Elston, R. C. & Stewart, J. A general model for the genetic analysis of pedigree data. Hum Hered 21, 523-42 (1971).

46. Boehnke, M. Allele frequency estimation from data on relatives. Am J Hum Genet 48, 22-5 (1991).

47. Kruglyak, L., Daly, M. J., Reeve-Daly, M. P. & Lander, E. S. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58, 1347-63 (1996).

48. Kong, A. & Cox, N. J. Allele-sharing models: LOD scores and accurate linkage tests. Am J Hum Genet 61, 1179-88 (1997).

49. Sobel, E. & Lange, K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 58, 1323-37 (1996).

50. Mukhopadhyay, N., Almasy, L., Schroeder, M., Mulvihill, W. P. & Weeks, D. E. Mega2: data-handling for facilitating genetic linkage and association analyses. Bioinformatics 21, 2556-7 (2005).

51. Thiele, H. & Nurnberg, P. HaploPainter: a tool for drawing pedigrees with complex haplotypes. Bioinformatics 21, 1730-2 (2005).

52. O'Connell, J. R. & Weeks, D. E. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 63, 259-66 (1998).

53. Lathrop, G. M. & Lalouel, J. M. Easy calculations of lod scores and genetic risks on small computers. Am J Hum Genet 36, 460-5 (1984).

54. Stringham, H. M. & Boehnke, M. Identifying marker typing incompatibilities in linkage analysis. Am J Hum Genet 59, 946-50 (1996).

55. Ott, J. Computer-simulation methods in human linkage analysis. Proc Natl Acad Sci U S A 86, 4175-8 (1989).

56. Boehnke, M. Estimating the power of a proposed linkage study: a practical computer simulation approach. Am J Hum Genet 39, 513-27 (1986).

57. Ott, S., Klingholz, F., Willich, N. & Kastenbauer, E. [Assessing the quality of the speaking voice after therapy of T1 and T2 vocal cord cancers]. Laryngorhinootologie 71, 236-41 (1992).

58. Parkinson, J. An Essay on the Shaking Palsy. (ed. Sherwood) (Neely and Jones, London, 1817).

59. Ishikawa, A. & Tsuji, S. Clinical analysis of 17 patients in 12 Japanese families with autosomal-recessive type juvenile parkinsonism. Neurology 47, 160-6 (1996).

60. de Rijk, M. C. et al. Prevalence of Parkinson's disease in the elderly: the Rotterdam Study. Neurology 45, 2143-6 (1995).

61. Lazzarini, A. M. et al. A clinical genetic study of Parkinson's disease: evidence for dominant transmission. Neurology 44, 499-506 (1994).

62. Pollanen, M. S., Dickson, D. W. & and Bergeron, C. Pathology and biology of the Lewy body. J. Neuropathol. Exp. Neurol 52, 183-191 (1993).

63. Shimura, H. et al. Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. Nat Genet 25, 302-5 (2000).

64. Shimura, H. et al. Ubiquitination of a new form of alpha-synuclein by parkin from human brain: implications for Parkinson's disease. Science 293, 263-9 (2001).

65. Farrer, M. et al. Lewy bodies and parkinsonism in families with parkin mutations. Ann Neurol 50, 293-300 (2001).

66. Tanaka, Y. et al. Inducible expression of mutant alpha-synuclein decreases proteasome activity and increases sensitivity to mitochondira-dependent apoptosis. Hum Mol Genet 10, 919-926 (2001).

67. Poskanser, D. a. S., RS. Studies in the epidemiology of Parkinson's disease predicting its disappearance as a major clinical enity by 1980. Trans. Am. Neurol. Assoc. 86, 234-245 (1961).

68. Langston, W., Ballard, P., Tetrud, JW. and Irwin, I. Chronic Parkinsonism in humans due to a product of meperidine-analog synthesis. Science 219, 979-980 (1983).

69. Polymeropoulus MH, L. C., Leroy E, Ide SE, Dehejia A, Dutra A, et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. Science 276, 2045-2047 (1997).

70. Kitada, T. et al. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. Nature 392, 605-8 (1998).

71. Leroy, E. et al. The Ubiquitin pathway in Parkinson's disease. Nature 395, 451-452 (1998).

72. Bonifati, V. et al. Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. Science 299, 256-9 (2003).

73. Valente, E. M. et al. Hereditary early-onset Parkinson's disease caused by mutations in PINK1. Science 304, 1158-60 (2004).

74. Spillantini, M. G., Crowther, R. A., Kamphorst, W., Heutink, P. & van Swieten, J. C. Tau pathology in two Dutch families with mutations in the microtubule-binding region of tau. Am J Pathol 153, 1359-63 (1998).

75. Gasser, T. et al. A susceptibility locus for Parkinson's disease maps to chromosome 2p13. Nat Genet 18, 262-5 (1998).

76. Farrer, M. et al. A chromosome 4p haplotype segregating with Parkinson's disease and postural tremor. Hum Mol Genet 8, 81-5 (1999).

77. Funayama, M. et al. A new locus for Parkinson's disease (PARK8) maps to chromosome 12p11.2-q13.1. Ann Neurol 51, 296-301 (2002).

78. Williams, D. R., Hadeed, A., Al-Din, A. S., Wreikat, A. L. & Lees, A. J. Kufor Rakeb Disease: Autosomal recessive, levodopa-responsive parkinsonism with pyramidal degeneration, supranuclear gaze palsy, and dementia. Mov Disord (2005).

79. Hicks, A. A. et al. A susceptibility gene for late-onset idiopathic Parkinson's disease. Ann Neurol 52, 549-55 (2002).

80. Pankratz, N. et al. Significant linkage of Parkinson disease to chromosome 2q36-37. Am J Hum Genet 72, 1053-7 (2003).

81. Wood, N. Genes and parkinsonism. J Neurol Neurosurg Psychiatry 62, 305-9 (1997).

82. Chase, B. A., Katechalidou, L., Wszolek, ZK., Markopoulou, K. Early-onset but not late onset sporadic Parkinson's disease in the Greek population is associated with the G209A mutation at the alpha-synuclein gene. Neurology 52(suppl), A221 (1999).

83. Lucking, C. B. et al. Association between early-onset Parkinson's disease and mutations in the parkin gene. French Parkinson's Disease Genetics Study Group. N Engl J Med 342, 1560-7 (2000).
84. Li, Y. J. et al. Age at onset in two common neurodegenerative diseases is genetically controlled. Am J Hum Genet 70, 985-93 (2002).
85. Polymeropoulus MH, H. J., Golbe LI, Johnson WG, et al. Mapping of a gene for Parkinson's disease to chromosome 4q21-q23. Science 274, 1197-1199 (1996).
86. Matsumine, H. et al. Localization of a gene for an autosomal recessive form of juvenile Parkinsonism to chromosome 6q25.2-27. Am J Hum Genet 60, 588-96 (1997).
87. Jones, A. C. et al. Autosomal recessive juvenile parkinsonism maps to 6q25.2-q27 in four ethnic groups: detailed genetic mapping of the linked region. Am J Hum Genet 63, 80-7 (1998).
88. Hutton, M. et al. Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. Nature 393, 702-5 (1998).
89. Valente, E. M. et al. Localization of a novel locus for autosomal recessive early-onset parkinsonism, PARK6, on human chromosome 1p35-p36. Am J Hum Genet 68, 895-900 (2001).
90. van Duijn, C. M. et al. Park7, a novel locus for autosomal recessive early-onset parkinsonism, on chromosome 1p36. Am J Hum Genet 69, 629-34 (2001).
91. Polymeropoulos MH, H. J., Golbe LI, Johnson WG, et al. Mapping of a gene for Parkinson's disease to chromosome 4q21-q23. Science 274, 1197-1199 (1996).
92. Jensen PH, S. E., Petersen TE, Gliemann J, Rasmussen LK. Residues in the synuclein consensus motif of the alpha-synuclein fragment, NAC, participate in transglutaminase-catalized cross-linking to Alzheimer-disease amyloid beta A4 peptide. Biochem J 310, 91-94 (1995).
93. Spillantini, M. G. et al. Alpha-synuclein in Lewy bodies. Nature 388, 839-40 (1997).
94. Kruger, R. et al. Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease. Nat Genet 18, 106-8 (1998).
95. Singleton, A. B. et al. alpha-Synuclein locus triplication causes Parkinson's disease. Science 302, 841 (2003).
96. Zarranz, J. J. et al. The new mutation, E46K, of alpha-synuclein causes Parkinson and Lewy body dementia. Ann. Neurol 55, 164-173 (2004).
97. Wilkinson, K., et al. Science 246, 670-673 (1989).
98. Larsen, C. N., Krantz, B. A. & Wilkinson, K. D. Substrate specificity of deubiquitinating enzymes: ubiquitin C-terminal hydrolases. Biochemistry 37, 3358-68 (1998).
99. Lowe, J., McDermott, H., Landon, M., Mayer, RJ., Wilkinson, KD. Ubiquitin conjugate immunoreactivity in the brains of scrapie infected mice. J. Pathol 161, 153-160 (1990).
100. Leroy, E., Anastasopoulos, D., Konitsiotis, S., Lavedan, C. & Polymeropoulos, M. H. Deletions in the Parkin gene and genetic heterogeneity in a Greek family with early onset Parkinson's disease. Hum Genet 103, 424-7 (1998).

101. Foster, N. L. et al. Frontotemporal dementia and parkinsonism linked to chromosome 17: a consensus conference. Conference Participants. Ann Neurol 41, 706-15 (1997).

102. Guidot, D. M., McCord, J. M., Wright, R. M. & Repine, J. E. Absence of electron transport (Rho 0 state) restores growth of a manganese-superoxide dismutase-deficient Saccharomyces cerevisiae in hyperoxia. Evidence for electron transport as a major source of superoxide generation in vivo. J Biol Chem 268, 26699-703 (1993).

103. Schapira, A. H. et al. Mitochondrial complex I deficiency in Parkinson's disease. Lancet 1, 1269 (1989).

104. Hattori, N., Ikebe, S., Tanaka, M., Ozawa, T. & Mizuno, Y. Immunohistochemical studies on complexes I, II, III, and IV of mitochondria in Parkinson's disease. Adv Neurol 60, 292-6 (1993).

105. Hassan, H. Biosynthesys and regulation of superoxide dismutases. Free Rad Biol Med 5, 377-385 (1988).

106. Saggu, H. et al. A selective increase in particulate superoxide dismutase activity in parkinsonian substantia nigra. J Neurochem 53, 692-7 (1989).

107. Abbas, N. et al. A wide variety of mutations in the parkin gene are responsible for autosomal recessive parkinsonism in Europe. French Parkinson's Disease Genetics Study Group and the European Consortium on Genetic Susceptibility in Parkinson's Disease. Hum Mol Genet 8, 567-74 (1999).

108. Hattori, N. et al. Molecular genetic analysis of a novel Parkin gene in Japanese families with autosomal recessive juvenile parkinsonism: evidence for variable homozygous deletions in the Parkin gene in affected individuals. Ann Neurol 44, 935-41 (1998).

109. Lucking, C. B. et al. Homozygous deletions in parkin gene in European and North African families with autosomal recessive juvenile parkinsonism. The European Consortium on Genetic Susceptibility in Parkinson's Disease and the French Parkinson's Disease Genetics Study Group. Lancet 352, 1355-6 (1998).

110. West, A. et al. Complex relationship between Parkin mutations and Parkinson disease. Am J Med Genet 114, 584-91 (2002).

111. Kann, M. et al. Role of parkin mutations in 111 community-based patients with early-onset parkinsonism. Ann Neurol 51, 621-5 (2002).

112. Nichols, W. C. et al. Linkage stratification and mutation analysis at the Parkin locus identifies mutation positive Parkinson's disease families. J Med Genet 39, 489-92 (2002).

113. Hedrich, K. et al. Evaluation of 50 probands with early-onset Parkinson's disease for Parkin mutations. Neurology 58, 1239-46 (2002).

114. Nisipeanu, P. et al. Parkin gene causing benign autosomal recessive juvenile parkinsonism. Neurology 56, 1573-5 (2001).

115. Ujike, H. et al. Prevalence of homozygous deletions of the parkin gene in a cohort of patients with sporadic and familial Parkinson's disease. Mov Disord 16, 111-3 (2001).

116. Hedrich, K. et al. The importance of gene dosage studies: mutational analysis of the parkin gene in early-onset parkinsonism. Hum Mol Genet 10, 1649-56 (2001).

117. Hoenicka, J. et al. Molecular findings in familial Parkinson disease in Spain. Arch Neurol 59, 966-70 (2002).

118. Maruyama, M. et al. Novel mutations, pseudo-dominant inheritance, and possible familial affects in patients with autosomal recessive juvenile parkinsonism. Ann Neurol 48, 245-50 (2000).

119. Oliveira, S. A. et al. Parkin mutations and susceptibility alleles in late-onset Parkinson's disease. Ann Neurol 53, 624-9 (2003).

120. Foroud, T. et al. Heterozygosity for a mutation in the parkin gene leads to later onset Parkinson disease. Neurology 60, 796-801 (2003).

121. Hattori, N. et al. Point mutations (Thr240Arg and Gln311Stop) [correction of Thr240Arg and Ala311Stop] in the Parkin gene. Biochem Biophys Res Commun 249, 754-8 (1998).

122. Periquet, M. et al. Parkin mutations are frequent in patients with isolated early-onset parkinsonism. Brain 126, 1271-8 (2003).

123. Wang, T. et al. [A novel point mutation in parkin gene was identified in an early-onset case of Parkinson's disease]. Zhonghua Yi Xue Yi Chuan Xue Za Zhi 20, 111-3 (2003).

124. Alvarez, V., Guisasola, L. M., Moreira, V. G., Lahoz, C. H. & Coto, E. Early-onset Parkinson's disease associated with a new parkin mutation in a Spanish family. Neurosci Lett 313, 108-10 (2001).

125. Rawal, N. et al. New parkin mutations and atypical phenotypes in families with autosomal recessive parkinsonism. Neurology 60, 1378-81 (2003).

126. Munoz, E. et al. Relative high frequency of the c.255delA parkin gene mutation in Spanish patients with autosomal recessive parkinsonism. J Neurol Neurosurg Psychiatry 73, 582-4 (2002).

127. Hedrich, K. et al. Distribution, type, and origin of Parkin mutations: review and case studies. Mov Disord 19, 1146-57 (2004).

128. Klein, C. et al. Parkin deletions in a family with adult-onset, tremor-dominant parkinsonism: expanding the phenotype. Ann Neurol 48, 65-71 (2000).

129. Hilker, R. et al. Positron emission tomographic analysis of the nigrostriatal dopaminergic system in familial parkinsonism associated with mutations in the parkin gene. Ann Neurol 49, 367-76 (2001).

130. McKeigue, P. M. Prospects for admixture mapping of complex traits. Am J Hum Genet 76, 1-7 (2005).

131. Schneider, S. K., J-M
Roessli, D
Excoffier, L. (2000).

132. Swaiman, K. F. Hallervorden-Spatz syndrome and brain iron metabolism. Arch Neurol 48, 1285-93 (1991).

133. Hayflick, S. J. et al. Genetic, clinical, and radiographic delineation of Hallervorden-Spatz syndrome. N Engl J Med 348, 33-40 (2003).

134. Zhou, B. et al. A novel pantothenate kinase gene (PANK2) is defective in Hallervorden-Spatz syndrome. Nat Genet 28, 345-9 (2001).

135. Taylor, T. D. et al. Homozygosity mapping of Hallervorden-Spatz syndrome to chromosome 20p12.3-p13. Nat Genet 14, 479-81 (1996).

136. Abiko, Y. Investigations on Patothenic acid and its related compounds. IX biochemical studies. Separation and substrate specificity of Pantothenate kinase and phosphopantothenoylcysteine. J Biochem 61, 290-299 (1967).

137. Gouider-Khouja, N., Miladi, N., Belal, S. & Hentati, F. Intrafamilial phenotypic variability of Hallervorden-Spatz syndrome in a Tunisian family. Parkinsonism Relat Disord 6, 175-179 (2000).

138. Grimes, D. A., Lang, A. E. & Bergeron, C. Late adult onset chorea with typical pathology of Hallervorden-Spatz syndrome. J Neurol Neurosurg Psychiatry 69, 392-5 (2000).

139. Cooper, G. E., Rizzo, M. & Jones, R. D. Adult-onset Hallervorden-Spatz syndrome presenting as cortical dementia. Alzheimer Dis Assoc Disord 14, 120-6 (2000).

140. Hayflick, S. J. Pantothenate kinase-associated neurodegeneration (formerly Hallervorden-Spatz syndrome). J Neurol Sci 207, 106-7 (2003).

141. Jankovic, J., Kirkpatrick, J. B., Blomquist, K. A., Langlais, P. J. & Bird, E. D. Late-onset Hallervorden-Spatz disease presenting as familial parkinsonism. Neurology 35, 227-34 (1985).

142. Gusella, J. F. et al. A polymorphic DNA marker genetically linked to Huntington's disease. Nature 306, 234-8 (1983).

143. Duyao, M. et al. Trinucleotide repeat length instability and age of onset in Huntington's disease. Nat Genet 4, 387-92 (1993).

144. Benjamin, C. M. et al. Proceed with care: direct predictive testing for Huntington disease. Am J Hum Genet 55, 606-17 (1994).

145. Leeflang, E. P. et al. Single sperm analysis of the trinucleotide repeats in the Huntington's disease gene: quantification of the mutation frequency spectrum. Hum Mol Genet 4, 1519-26 (1995).

146. Chong, S. S. et al. Contribution of DNA sequence and CAG size to mutation frequencies of intermediate alleles for Huntington disease: evidence from single sperm analyses. Hum Mol Genet 6, 301-9 (1997).

147. Kremer, B. et al. Sex-dependent mechanisms for expansions and contractions of the CAG repeat on affected Huntington disease chromosomes. Am J Hum Genet 57, 343-50 (1995).

148. Telenius, H. et al. Somatic mosaicism in sperm is associated with intergenerational (CAG)n changes in Huntington disease. Hum Mol Genet 4, 189-95 (1995).

149. Falush, D., Almqvist, E. W., Brinkmann, R. R., Iwasa, Y. & Hayden, M. R. Measurement of Mutational Flow Implies Both a High New-Mutation Rate for Huntington Disease and Substantial Underascertainment of Late-Onset Cases. Am J Hum Genet 68, 373-385 (2000).

150. Legius, E. et al. Limited expansion of the (CAG)n repeat of the Huntington gene: a premutation. Eur J Hum Genet 2, 44-50 (1994).

151. McNeil, S. M. et al. Reduced penetrance of the Huntington's disease mutation. Hum Mol Genet 6, 775-9 (1997).

152. Li, S. H., Lam, S., Cheng, A. L. & Li, X. J. Intranuclear huntingtin increases the expression of caspase-1 and induces apoptosis. Hum Mol Genet 9, 2859-67 (2000).

153. Zuccato, C. et al. Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes. Nat Genet 35, 76-83 (2003).

154. Humbert, S. et al. The IGF-1/Akt pathway is neuroprotective in Huntington's disease and involves Huntingtin phosphorylation by Akt. Dev Cell 2, 831-7 (2002).

155. Kalchman, M. A. et al. Huntingtin is ubiquitinated and interacts with a specific ubiquitin-conjugating enzyme. J Biol Chem 271, 19385-94 (1996).

156. Steffan, J. S. et al. SUMO modification of Huntingtin and Huntington's disease pathology. Science 304, 100-4 (2004).

157. Lunkes, A. et al. Proteases acting on mutant huntingtin generate cleaved products that differentially build up cytoplasmic and nuclear inclusions. Mol Cell 10, 259-69 (2002).

158. Kato, J. et al. A mutation, in the iron-responsive element of H ferritin mRNA, causing autosomal dominant iron overload. Am J Hum Genet 69, 191-7 (2001).

159. Curtis, A. R. et al. Mutation in the gene encoding ferritin light polypeptide causes dominant adult-onset basal ganglia disease. Nat Genet 28, 350-4 (2001).

160. Gerlach, M., Ben-Schachar, D., Riederer, P. & Youdim, M. B. Altered brain metabolism of iron as a cause of neurodenerative diseases? J. Neurochem 63, 793-807 (1994).

161. Ghetti, V. et al. Intracellular ferritin accumulation in neural and extaneural tissue characterizes a neurodegenerative disease associated with a mutation in the ferritin light polypeptide gene. J Neuropathol Exp 63, 363-380 (2004).

162. Mancuso, M. et al. Hereditary ferrinopathy: a novel mutation, its cellular pathology, and pathogenetic insgiths. J Neuropathol Exp 64, 280-294 (2005).

163. Beaumont, C. et al. Mutation in the iron responsive element of the L ferritin mRNA in a family with dominant hyperferritinaemia and cataract. Nat Genet 11, 444-6 (1995).

164. Aguilar-Martinez, P., Biron, C., Masmejean, C., Jeanjean, P. & Schved, J. F. A novel mutation in the iron responsive element of ferritin L-subunit gene as a cause for hereditary hyperferritinemia-cataract syndrome. Blood 88, 1895 (1996).

165. Cazzola, M. et al. Hereditary hyperferritinemia-cataract syndrome: relationship between phenotypes and specific mutations in the iron-responsive element of ferritin light-chain mRNA. Blood 90, 814-21 (1997).

166. Girelli, D. et al. Clinical, biochemical and molecular findings in a series of families with hereditary hyperferritinaemia-cataract syndrome. Br J Haematol 115, 334-40 (2001).

167. Camaschella, C. et al. A new mutation (G51C) in the iron-responsive element (IRE) of L-ferritin associated with hyperferritinaemia-cataract syndrome decreases the binding affinity of the mutated IRE for iron-regulatory proteins. Br J Haematol 108, 480-2 (2000).

168. Piperno, A. & Pulucchi, S. Report of a mutation in FTL. Hum Genet 114, 406 (2004).

169. Burmester, T., Weich, B., Reinhardt, S. & Hankeln, T. A vertebrate globin expressed in the brain. Nature 407, 520-3 (2000).

170. Hendrick, L. & Van Broeckhoven, C. Eur J Biochem 237, 6-15 (1996).

171. Moens, L. & Dewilde, S. Globins in the brain. Nature 407, 461-2 (2000).

172. Scheuer, M. L. & Pedley, T. A. The evaluation and treatment of seizures. N Engl J Med 323, 1468-74 (1990).

173. DeLorenzo, R. J. The challenging genetics of epilepsy. Epilepsy Res Suppl 4, 3-17 (1991).

174. ILAE, I. L. A. E. Proposal for Revised Classification of Epilepsies and Epileptic Syndromes: Commission on Classification and Terminology of the International League Against Epilepsy. Epilepsia 30, 389-399 (1989).

175. Bate, L. & Gardiner, M. Molecular genetics of human epilepsies. Expert Rev Mol Med 1999, 1-22 (1999).

176. Shorvon, S. D. The epidemiology and treatment of chronic and refractory epilepsy. Epilepsia 37 Suppl 2, S1-S3 (1996).

177. Zuluaga, L. et al. Prevalencia de epilepsia en Medellín, Colombia. Boletin de la oficina Sanitaria Panamericana 104, 331-344 (1983).

178. Annegers, J. F., Rocca, W. A. & Hauser, W. A. Causes of epilepsy: contributions of the Rochester epidemiology project. Mayo Clin Proc 71, 570-5 (1996).

179. Noebls, J. Modeling Human Epilepsies in Mice. Epilepsia. Epilepsia 42, 11-15 (2001).

180. Berkovic, S. F., Howell, R. A., Hay, D. A. & Hopper, J. L. Epilepsies in twins: genetics of the major epilepsy syndromes. Ann Neurol 43, 435-45 (1998).
181. Beck-Mannagetta, G. Genetics of the Epilepsies (Berlin, 1989).
182. Steinlein, O. K. et al. A missense mutation in the neuronal nicotinic acetylcholine receptor alpha 4 subunit is associated with autosomal dominant nocturnal frontal lobe epilepsy. Nat Genet 11, 201-3 (1995).
183. Gambardella, A. et al. A new locus for autosomal dominant nocturnal frontal lobe epilepsy maps to chromosome 1. Neurology 55, 1467-71 (2000).
184. Ottman, R. et al. Localization of a gene for partial epilepsy to chromosome 10q. Nat Genet 10, 56-60 (1995).
185. Kalachikov, S. et al. Mutations in LGI1 cause autosomal-dominant partial epilepsy with auditory features. Nat Genet 30, 335-41 (2002).
186. Neubauer, B. A. et al. Centrotemporal spikes in families with rolandic epilepsy: linkage to chromosome 15q14. Neurology 51, 1608-12 (1998).
187. Virtaneva, K. et al. Progressive myoclonus epilepsy EPM1 locus maps to a 175-kb interval in distal 21q. Am J Hum Genet 58, 1247-53 (1996).
188. Pennacchio, L. A. et al. Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1). Science 271, 1731-4 (1996).
189. Lalioti, M. D. et al. Dodecamer repeat expansion in cystatin B gene in progressive myoclonus epilepsy. Nature 386, 847-51 (1997).
190. Shoffner, J. M. et al. Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA(Lys) mutation. Cell 61, 931-7 (1990).
191. IBDC, I. B. D. C. Isolation of a novel gene underlying Batten disease, CLN3. The International Batten Disease Consortium. Cell 82, 949-57 (1995).
192. Munroe, P. B. et al. Spectrum of mutations in the Batten disease gene, CLN3. Am J Hum Genet 61, 310-6 (1997).
193. Mitchison, H. M. et al. Genomic structure and complete nucleotide sequence of the Batten disease gene, CLN3. Genomics 40, 346-50 (1997).
194. Minassian, B. A. et al. Mutations in a gene encoding a novel protein tyrosine phosphatase cause progressive myoclonus epilepsy. Nat Genet 20, 171-4 (1998).
195. Serratosa, J. M. et al. A novel protein tyrosine phosphatase gene is mutated in progressive myoclonus epilepsy of the Lafora type (EPM2). Hum Mol Genet 8, 345-52 (1999).
196. Tahvanainen, E. et al. The gene for a recessively inherited human childhood progressive epilepsy with mental retardation maps to the distal short arm of chromosome 8. Proc Natl Acad Sci U S A 91, 7267-70 (1994).
197. Ranta, S. et al. The neuronal ceroid lipofuscinoses in human EPMR and mnd mutant mice are associated with mutations in CLN8. Nat Genet 23, 233-6 (1999).
198. Liu, A. W. et al. Juvenile myoclonic epilepsy locus in chromosome 6p21.2-p11: linkage to convulsions and electroencephalography trait. Am J Hum Genet 57, 368-81 (1995).
199. Greenberg, D. A., Durner, M., Resor, S., Rosenbaum, D. & Shinnar, S. The genetics of idiopathic generalized epilepsies of adolescent onset: differences between juvenile myoclonic epilepsy and epilepsy with random grand mal and with awakening grand mal. Neurology 45, 942-6 (1995).
200. Greenberg, D. A. et al. Association of HLA class II alleles in patients with juvenile myoclonic epilepsy compared with patients with other forms of adolescent-onset generalized epilepsy. Neurology 47, 750-5 (1996).

201. Serratosa, J. M. et al. Clinical and genetic analysis of a large pedigree with juvenile myoclonic epilepsy. Ann Neurol 39, 187-95 (1996).
202. Elmslie, F. V. et al. Genetic mapping of a major susceptibility locus for juvenile myoclonic epilepsy on chromosome 15q. Hum Mol Genet 6, 1329-34 (1997).
203. Greenberg, D. A. et al. Reproducibility and complications in gene searches: linkage on chromosome 6, heterogeneity, association, and maternal inheritance in juvenile myoclonic epilepsy. Am J Hum Genet 66, 508-16 (2000).
204. Cossette, P. et al. Mutation of GABRA1 in an autosomal dominant form of juvenile myoclonic epilepsy. Nat Genet 31, 184-9 (2002).
205. Zara, F. et al. Mapping of genes predisposing to idiopathic generalized epilepsy. Hum Mol Genet 4, 1201-7 (1995).
206. Haug, K. et al. Mutations in CLCN2 encoding a voltage-gated chloride channel are associated with idiopathic generalized epilepsies. Nat Genet 33, 527-32 (2003).
207. Zara, F. et al. Mapping of a locus for a familial autosomal recessive idiopathic myoclonic epilepsy of infancy to chromosome 16p13. Am J Hum Genet 66, 1552-7 (2000).
208. Durner, M. et al. Evidence for linkage of adolescent-onset idiopathic generalized epilepsies to chromosome 8-and genetic heterogeneity. Am J Hum Genet 64, 1411-9 (1999).
209. Zara, F. et al. Unusual EEG pattern linked to chromosome 3p in a family with idiopathic generalized epilepsy. Neurology 51, 493-8 (1998).
210. Leppert, M. et al. Benign familial neonatal convulsions linked to genetic markers on chromosome 20. Nature 337, 647-8 (1989).
211. Biervert, C. et al. A potassium channel mutation in neonatal human epilepsy. Science 279, 403-6 (1998).
212. Singh, N. A. et al. A novel potassium channel gene, KCNQ2, is mutated in an inherited epilepsy of newborns. Nat Genet 18, 25-9 (1998).
213. Lewis, T. B., Leach, R. J., Ward, K., O'Connell, P. & Ryan, S. G. Genetic heterogeneity in benign familial neonatal convulsions: identification of a new locus on chromosome 8q. Am J Hum Genet 53, 670-5 (1993).
214. Charlier, C. et al. A pore mutation in a novel KQT-like potassium channel gene in an idiopathic epilepsy family. Nat Genet 18, 53-5 (1998).
215. Wallace, R. H., Berkovic, S. F., Howell, R. A., Sutherland, G. R. & Mulley, J. C. Suggestion of a major gene for familial febrile convulsions mapping to 8q13-21. J Med Genet 33, 308-12 (1996).
216. Johnson, E. W. et al. Evidence for a novel gene for familial febrile convulsions, FEB2, linked to chromosome 19p in an extended family from the Midwest. Hum Mol Genet 7, 63-7 (1998).
217. Nakayama, J. et al. Significant evidence for linkage of febrile seizures to chromosome 5q14-q15. Hum Mol Genet 9, 87-91 (2000).
218. Nakayama, J. et al. A nonsense mutation of the MASS1 gene in a family with febrile and afebrile seizures. Ann Neurol 52, 654-7 (2002).
219. Wallace, R. H. et al. Febrile seizures and generalized epilepsy associated with a mutation in the Na+-channel beta1 subunit gene SCN1B. Nat Genet 19, 366-70 (1998).
220. Escayg, A. et al. Mutations of SCN1A, encoding a neuronal sodium channel, in two families with GEFS+2. Nat Genet 24, 343-5 (2000).

221. Escayg, A. et al. A novel SCN1A mutation associated with generalized epilepsy with febrile seizures plus--and prevalence of variants in patients with epilepsy. Am J Hum Genet 68, 866-73 (2001).

222. Wallace, R. H. et al. Neuronal sodium-channel alpha1-subunit mutations in generalized epilepsy with febrile seizures plus. Am J Hum Genet 68, 859-65 (2001).

223. Sugawara, T. et al. A missense mutation of the Na+ channel alpha II subunit gene Na(v)1.2 in a patient with febrile and afebrile seizures causes channel dysfunction. Proc Natl Acad Sci U S A 98, 6384-9 (2001).

224. Baulac, S. et al. Evidence for digenic inheritance in a family with both febrile convulsions and temporal lobe epilepsy implicating chromosomes 18qter and 1q25-q31. Ann Neurol 49, 786-92 (2001).

225. Berkovic, S. F. & Scheffer, I. E. Genetics of the epilepsies. Epilepsia 42 Suppl 5, 16-23 (2001).

226. Ottman, R., Annegers, J. F., Risch, N., Hauser, W. A. & Susser, M. Relations of genetic and environmental factors in the etiology of epilepsy. Ann Neurol 39, 442-9 (1996).

227. Scheffer, I. E. & Berkovic, S. F. Generalized epilepsy with febrile seizures plus. A genetic disorder with heterogeneous clinical phenotypes. Brain 120 (Pt 3), 479-90 (1997).

228. Singh, R., Scheffer, I. E., Crossland, K. & Berkovic, S. F. Generalized epilepsy with febrile seizures plus: a common childhood-onset genetic epilepsy syndrome. Ann Neurol 45, 75-81 (1999).

229. Wallace, R. H. et al. Generalized epilepsy with febrile seizures plus: mutation of the sodium channel subunit SCN1B. Neurology 58, 1426-9 (2002).

230. Baulac, S. et al. A second locus for familial generalized epilepsy with febrile seizures plus maps to chromosome 2q21-q33. Am J Hum Genet 65, 1078-85 (1999).

231. Moulard, B. et al. Identification of a new locus for generalized epilepsy with febrile seizures plus (GEFS+) on chromosome 2q24-q33. Am J Hum Genet 65, 1396-400 (1999).

232. Abou-Khalil, B. et al. Partial and generalized epilepsy with febrile seizures plus and a novel SCN1A mutation. Neurology 26, 2265-72 (2001).

233. Annesi, G. et al. Two novel SCN1A missense mutations in generalized epilepsy with febrile seizures plus. Epilepsia 44, 1257-8 (2003).

234. Sugawara, T. et al. Nav1.1 mutations cause febrile seizures associated with afebrile partial seizures. Neurology 57, 703-5 (2001).

235. Lossin, C. et al. Epilepsy-associated dysfunction in the voltage-gated neuronal sodium channel SCN1A. J Neurosci 23, 11289-95 (2003).

236. Claes, L. et al. De novo mutations in the sodium-channel gene SCN1A cause severe myoclonic epilepsy of infancy. Am J Hum Genet 68, 1327-32 (2001).

237. Claes, L. et al. De novo SCN1A mutations are a major cause of severe myoclonic epilepsy of infancy. Hum Mutat 21, 615-21 (2003).

238. Sugawara, T. et al. Frequent mutations of SCN1A in severe myoclonic epilepsy in infancy. Neurology 58, 1122-4 (2002).

239. Ohmori, I., Ouchida, M., Ohtsuka, Y., Oka, E. & Shimizu, K. Significant correlation of the SCN1A mutations and severe myoclonic epilepsy in infancy. Biochem Biophys Res Commun 295, 17-23 (2002).

240. Lossin, C., Wang, D. W., Rhodes, T. H., Vanoye, C. G. & George, A. L., Jr. Molecular basis of an inherited epilepsy. Neuron 34, 877-84 (2002).

241. Wilcox, A. S. et al. Human chromosomal localization of genes encoding the gamma 1 and gamma 2 subunits of the gamma-aminobutyric acid receptor indicates that members of this gene family are often clustered in the genome. Proc Natl Acad Sci U S A 89, 5857-61 (1992).

242. Kostrzewa, M. et al. Assignment of genes encoding GABAA receptor subunits alpha 1, alpha 6, beta 2, and gamma 2 to a YAC contig of 5q33. Eur J Hum Genet 4, 199-204 (1996).

243. Baulac, S. et al. First genetic evidence of GABA(A) receptor dysfunction in epilepsy: a mutation in the gamma2-subunit gene. Nat Genet 28, 46-8 (2001).

244. Kananura, C. et al. A splice-site mutation in GABRG2 associated with childhood absence epilepsy and febrile convulsions. Arch Neurol 59, 1137-41 (2002).

245. Harkin, L. A. et al. Truncation of the GABA(A)-receptor gamma2 subunit in a family with generalized epilepsy with febrile seizures plus. Am J Hum Genet 70, 530-6 (2002).

246. Wallace, R. H. et al. Mutant GABA(A) receptor gamma2-subunit in childhood absence epilepsy and febrile seizures. Nat Genet 28, 49-52 (2001).

247. Mehta, A. K. & Ticku, M. K. An update on GABAA receptors. Brain Res Brain Res Rev 29, 196-217 (1999).

248. Sieghart, W. et al. Structure and subunit composition of GABA(A) receptors. Neurochem Int 34, 379-85 (1999).

249. Skradski, S. L. et al. A novel gene causing a mendelian audiogenic mouse epilepsy. Neuron 31, 537-44 (2001).

250. Peiffer, A. et al. A locus for febrile seizures (FEB3) maps to chromosome 2q23-24. Ann Neurol 46, 671-8 (1999).

251. Gerard, F., Pereira, S., Robaglia-Schlupp, A., Genton, P. & Szepetowski, P. Clinical and genetic analysis of a new multigenerational pedigree with GEFS+ (Generalized Epilepsy with Febrile Seizures Plus). Epilepsia 43, 581-6 (2002).

252. Rich, S. S., Annegers, J. F., Hauser, W. A. & Anderson, V. E. Complex segregation analysis of febrile convulsions. Am J Hum Genet 41, 249-57 (1987).

253. Malacarne, M. et al. Lack of SCN1A mutations in familial febrile seizures. Epilepsia 43, 559-62 (2002).

254. Ebedes, J. & Datta, A. Multiple sequence alignment in parallel on a workstation cluster. Bioinformatics 20, 1193-5 (2004).

255. Dyken, P. R. Reconsideration of the classification of the neuronal ceroid-lipofuscinoses. Am J Med Genet Suppl 5, 69-84 (1988).

256. Santavuori, P. Neuronal ceroid-lipofuscinoses in childhood. Brain Dev 10, 80-3 (1988).

257. Bate, L. & Gardiner, M. Genetics of inherited epilepsies. Epileptic Disord 1, 7-19 (1999).

258. Zeman, W. Presidential address: Studies in the neuronal ceroid-lipofuscinoses. J Neuropathol Exp Neurol 33, 1-12 (1974).

259. Tyynela, J., Palmer, D. N., Baumann, M. & Haltia, M. Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis. FEBS Lett 330, 8-12 (1993).

260. Furst, W. & Sandhoff, K. Activator proteins and topology of lysosomal sphingolipid catabolism. Biochim Biophys Acta 1126, 1-16 (1992).

261. Tyynela, J., Baumann, M., Henseler, M., Sandhoff, K. & Haltia, M. Sphingolipid activator proteins in the neuronal ceroid-lipofuscinoses: an immunological study. Acta Neuropathol (Berl) 89, 391-8 (1995).

262. Hall, N. A., Lake, B. D., Dewji, N. N. & Patrick, A. D. Lysosomal storage of subunit c of mitochondrial ATP synthase in Batten's disease (ceroid-lipofuscinosis). Biochem J 275 (Pt 1), 269-72 (1991).

263. Palmer, D. N. et al. Mitochondrial ATP synthase subunit c storage in the ceroid-lipofuscinoses (Batten disease). Am J Med Genet 42, 561-7 (1992).

264. Winchester, B. & Fleet, G. W. Amino-sugar glycosidase inhibitors: versatile tools for glycobiologists. Glycobiology 2, 199-210 (1992).

265. van Diggelen, O. P. et al. Adult neuronal ceroid lipofuscinosis with palmitoyl-protein thioesterase deficiency: first adult-onset patients of a childhood disease. Ann Neurol 50, 269-72 (2001).

266. Nijssen, P. C. et al. Autosomal dominant adult neuronal ceroid lipofuscinosis: a novel form of NCL with granular osmiophilic deposits without palmitoyl protein thioesterase 1 deficiency. Brain Pathol 13, 574-81 (2003).

267. Vesa, J. et al. Mutations in the palmitoyl protein thioesterase gene causing infantile neuronal ceroid lipofuscinosis. Nature 376, 584-7 (1995).

268. Sleat, D. E. et al. Association of mutations in a lysosomal protein with classical late-infantile neuronal ceroid lipofuscinosis. Science 277, 1802-5 (1997).

269. Savukoski, M. et al. CLN5, a novel gene encoding a putative transmembrane protein mutated in Finnish variant late infantile neuronal ceroid lipofuscinosis. Nat Genet 19, 286-8 (1998).

270. Gao, H. et al. Mutations in a novel CLN6-encoded transmembrane protein cause variant neuronal ceroid lipofuscinosis in man and mouse. Am J Hum Genet 70, 324-35 (2002).

271. Wheeler, R. B. et al. The gene mutated in variant late-infantile neuronal ceroid lipofuscinosis (CLN6) and in nclf mutant mice encodes a novel predicted transmembrane protein. Am J Hum Genet 70, 537-42 (2002).

272. Das, A. K. et al. Molecular genetics of palmitoyl-protein thioesterase deficiency in the U.S. J Clin Invest 102, 361-70 (1998).

273. Wisniewski, K. E. et al. Palmitoyl-protein thioesterase deficiency in a novel granular variant of LINCL. Pediatr Neurol 18, 119-23 (1998).

274. Zhong, N. et al. Molecular screening of Batten disease: identification of a missense mutation (E295K) in the CLN3 gene. Hum Genet 102, 57-62 (1998).

275. Sharp, J. D. et al. Loci for classical and a variant late infantile neuronal ceroid lipofuscinosis map to chromosomes 11p15 and 15q21-23. Hum Mol Genet 6, 591-5 (1997).

276. Hellsten, E., Vesa, J., Olkkonen, V. M., Jalanko, A. & Peltonen, L. Human palmitoyl protein thioesterase: evidence for lysosomal targeting of the enzyme and disturbed cellular routing in infantile neuronal ceroid lipofuscinosis. Embo J 15, 5240-5 (1996).

277. Verkruyse, L. A. & Hofmann, S. L. Lysosomal targeting of palmitoyl-protein thioesterase. J Biol Chem 271, 15831-6 (1996).

278. Ahtiainen, L., Van Diggelen, O. P., Jalanko, A. & Kopra, O. Palmitoyl protein thioesterase 1 is targeted to the axons in neurons. J Comp Neurol 455, 368-77 (2003).

279. Chattopadhyay, S. & Pearce, D. A. Neural and extraneural expression of the neuronal ceroid lipofuscinoses genes CLN1, CLN2, and CLN3: functional implications for CLN3. Mol Genet Metab 71, 207-11 (2000).

280. Heinonen, O., Salonen, T., Jalanko, A., Peltonen, L. & Copp, A. CLN-1 and CLN-5, genes for infantile and variant late infantile neuronal ceroid lipofuscinoses, are expressed in the embryonic human brain. J Comp Neurol 426, 406-12 (2000).

281. Mole, S. E., Williams, R. E. & Goebel, H. H. Correlations between genotype, ultrastructural morphology and clinical phenotype in the neuronal ceroid lipofuscinoses. Neurogenetics (2005).

282. Klockars, T. et al. Efficient construction of a physical map by fiber-FISH of the CLN5 region: refined assignment and long-range contig covering the critical region on 13q22. Genomics 35, 71-8 (1996).

283. Mazzei, R. et al. A novel mutation in the CLN1 gene in a patient with juvenile neuronal ceroid lipofuscinosis. J Neurol 249, 1398-400 (2002).

284. Sleat, D. E. et al. Mutational analysis of the defective protease in classic late-infantile neuronal ceroid lipofuscinosis, a neurodegenerative lysosomal storage disorder. Am J Hum Genet 64, 1511-23 (1999).

285. Zhong, N. et al. Heterogeneity of late-infantile neuronal ceroid lipofuscinosis. Genet Med 2, 312-8 (2000).

286. Mole, S. & Gardiner, M. Molecular genetics of the neuronal ceroid lipofuscinoses. Epilepsia 40 Suppl 3, 29-32 (1999).

287. Ezaki, J., Takeda-Ezaki, M. & Kominami, E. Tripeptidyl peptidase I, the late infantile neuronal ceroid lipofuscinosis gene product, initiates the lysosomal degradation of subunit c of ATP synthase. J Biochem (Tokyo) 128, 509-16 (2000).

288. Junaid, M. A., Wu, G. & Pullarkat, R. K. Purification and characterization of bovine brain lysosomal pepstatin-insensitive proteinase, the gene product deficient in the human late-infantile neuronal ceroid lipofuscinosis. J Neurochem 74, 287-94 (2000).

289. Golabek, A. A. et al. Biosynthesis, glycosylation, and enzymatic processing in vivo of human tripeptidyl-peptidase I. J Biol Chem 278, 7135-45 (2003).

290. Kurachi, Y. et al. Distribution and development of CLN2 protein, the late-infantile neuronal ceroid lipofuscinosis gene product. Acta Neuropathol (Berl) 102, 20-6 (2001).

291. Young, E. P., Worthington, V. C., Jackson, M. & Winchester, B. G. Pre- and postnatal diagnosis of patients with CLN1 and CLN2 by assay of palmitoyl-protein thioesterase and tripeptidyl-peptidase I activities. Eur J Paediatr Neurol 5 Suppl A, 193-6 (2001).

292. Ezaki, J. et al. Characterization of Cln3p, the gene product responsible for juvenile neuronal ceroid lipofuscinosis, as a lysosomal integral membrane glycoprotein. J Neurochem 87, 1296-308 (2003).

293. Haskell, R. E., Carr, C. J., Pearce, D. A., Bennett, M. J. & Davidson, B. L. Batten disease: evaluation of CLN3 mutations on protein localization and function. Hum Mol Genet 9, 735-44 (2000).

294. Jarvela, I., Lehtovirta, M., Tikkanen, R., Kyttala, A. & Jalanko, A. Defective intracellular transport of CLN3 is the molecular basis of Batten disease (JNCL). Hum Mol Genet 8, 1091-8 (1999).

295. Pearce, D. A., Ferea, T., Nosel, S. A., Das, B. & Sherman, F. Action of BTN1, the yeast orthologue of the gene mutated in Batten disease. Nat Genet 22, 55-8 (1999).

296. Mitchison, H. M. et al. Targeted disruption of the Cln3 gene provides a mouse model for Batten disease. The Batten Mouse Model Consortium [corrected]. Neurobiol Dis 6, 321-34 (1999).

297. Cooper, E. J., Johnston, G. A. & Edwards, F. A. Effects of a naturally occurring neurosteroid on GABAA IPSCs during development in rat hippocampal or cerebellar slices. J Physiol 521 Pt 2, 437-49 (1999).

298. Chattopadhyay, S. et al. An autoantibody inhibitory to glutamic acid decarboxylase in the neurodegenerative disorder Batten disease. Hum Mol Genet 11, 1421-31 (2002).

299. Mitchison, H. M. et al. Batten disease gene, CLN3: linkage disequilibrium mapping in the Finnish population, and analysis of European haplotypes. Am J Hum Genet 56, 654-62 (1995).

300. Wisniewski, H. M. & Wrzolek, M. Pathogenesis of amyloid formation in Alzheimer's disease, Down's syndrome and scrapie. Ciba Found Symp 135, 224-38 (1988).

301. Isosomppi, J., Vesa, J., Jalanko, A. & Peltonen, L. Lysosomal localization of the neuronal ceroid lipofuscinosis CLN5 protein. Hum Mol Genet 11, 885-91 (2002).

302. Vesa, J. et al. Neuronal ceroid lipofuscinoses are connected at molecular level: interaction of CLN5 protein with CLN2 and CLN3. Mol Biol Cell 13, 2410-20 (2002).

303. Holmberg, V. et al. Phenotype-genotype correlation in eight patients with Finnish variant late infantile NCL (CLN5). Neurology 55, 579-81 (2000).

304. Lonka, L., Kyttala, A., Ranta, S., Jalanko, A. & Lehesjoki, A. E. The neuronal ceroid lipofuscinosis CLN8 membrane protein is a resident of the endoplasmic reticulum. Hum Mol Genet 9, 1691-7 (2000).

305. Schmid, C. & Maraia, R. Transcriptional regulation and transpositional selection of active SINE sequences. Curr Opin Genet Dev 2, 874-82 (1992).

306. Santavuori, P., Rapola, J., Sainio, K. & Raitta, C. A variant of Jansky-Bielschowsky disease. Neuropediatrics 13, 135-41 (1982).

307. Tyynela, J., Suopanki, J., Santavuori, P., Baumann, M. & Haltia, M. Variant late infantile neuronal ceroid-lipofuscinosis: pathology and biochemistry. J Neuropathol Exp Neurol 56, 369-75 (1997).

308. Devinsky, O. Effects of Seizures on Autonomic and Cardiovascular Function. Epilepsy Currents 4, 43-46 (2004).

309. Baumgartner, C., Lurger, S. & Leutmezer, F. Autonomic Symptoms during epileptic Seizures. Epileptic Disord 1, 103-116 (2001).

310. Douglas, E. & White, P. Abdominal Epilepsy: A reappraisal. J Pediatr 78, 59-67 (1971).

311. Livingston, S. Abdominal pain as a manifestation of epilepsy (abdominal epilepsy) in children. J Pediatr 38, 687-695 (1958).

312. Van Buren, J. The abdominal aura: A study of abdominal sensations occuring in epilepsy and produced by depth stimulation. Electroenceph Clin Neurophysiol 15, 1-19 (1963).

313. Peppercorn, M. & Herzog, A. The spectrum of abdominal epilepsy in adults. Am J Gastroenterol 84, 1294-1296 (1989).

314. Garcia-Herrero, D., Fernandez-Torre, J. L., Barrasa, J., Calleja, J. & Pascual, J. Abdominal epilepsy in an adolescent with bilateral perisylvian polymicrogyria. Epilepsia 39, 1370-4 (1998).

315. Eschle, D., Siegel, A. & Wieser, H. Epilepsy with severe abdominal pain. Mayo Clin Proc 77, 1358-1360 (2002).

316. Moore, M. Paroxysmal abdominal pain: a form of focal symtomatic epilepsy. JAMA 124, 561-563 (1944).

317. Young, G. & Blume, W. Painful epileptic seizures. Brain 106, 537-554 (1983).

318. Mauguiere, F. & Courjon, J. Somatosensory epilepsy. A review of 127 cases. Brain 101, 307-332 (1978).

319. Zinkin, N. & Peppercorn, M. Abdominal epilepsy. Best Practice & Research Clinical Gastroenterology 19, 263-274 (2005).

320. Siegel, A. M., Williamson, P. D., Roberts, D. W., Thadani, V. M. & Darcey, T. M. Localized pain associated with seizures originating in the parietal lobe. Epilepsia 40, 845-55 (1999).

321. Mikami, M. et al. Localization of a gene for benign adult familial myoclonic epilepsy to chromosome 8q23.3-q24.1. Am J Hum Genet 65, 745-51 (1999).

322. Sugimoto, Y. et al. Childhood absence epilepsy in 8q24: refinement of candidate region and construction of physical map. Genomics 68, 264-72 (2000).

323. Rosenfeldt, H., Vazquez-Prado, J. & Gutkind, J. S. P-REX2, a novel PI-3-kinase sensitive Rac exchange factor. FEBS Lett 572, 167-71 (2004).

324. Ota, T. et al. Complete sequencing and characterization of 21,243 full-length human cDNAs. Nat Genet 36, 40-5 (2004).

325. Mellitus, E. C. o. t. D. a. C. o. D. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. . Diabetes Care 1, S4-19 (2000).

326. Gerard, J. [MODY types of diabetes mellitus]. Rev Med Liege 60, 439-41 (2005).

327. Todd, J. A. Transcribing diabetes. Nature 384, 407-8 (1996).

328. LaPorte, R. E. et al. The Pittsburgh Insulin Dependent Diabetes Mellitus Registry: the relationship of insulin dependent diabetes mellitus incidence to social class. Am J Epidemiol 114, 379-84 (1981).

329. Risch, N., Ghosh, S. & Todd, J. A. Statistical evaluation of multiple-locus linkage data in experimental species and its relevance to human studies: application to nonobese diabetic (NOD) mouse and human insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 53, 702-14 (1993).

330. Barnett, A. H., Eff, C., Leslie, R. D. & Pyke, D. A. Diabetes in identical twins. A study of 200 pairs. Diabetologia 20, 87-93 (1981).

331. Roivainen, M. et al. Several different enterovirus serotypes can be associated with prediabetic autoimmune episodes and onset of overt IDDM. Childhood Diabetes in Finland (DiMe) Study Group. J Med Virol 56, 74-8 (1998).

332. Thomson, G. et al. Genetic heterogeneity, modes of inheritance, and risk estimates for a joint study of Caucasians with insulin-dependent diabetes mellitus. Am J Hum Genet 43, 799-816 (1988).

333. Tuomilehto, J., Podar, T., Tuomilehto-Wolf, E. & Virtala, E. Evidence for importance of gender and birth cohort for risk of IDDM in offspring of IDDM parents. Diabetologia 38, 975-82 (1995).

334. Field, L. L. Genetic linkage and association studies of Type I diabetes: challenges and rewards. Diabetologia 45, 21-35 (2002).

335. Pugliese, A. Unraveling the genetics of insulin-dependent type 1A diabetes: the search must go on. Diabetes Reviews 7, 77-92 (1999).

336. Todd, J. A., Bell, J. I. & McDevitt, H. O. HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature 329, 599-604 (1987).

337. Sheehy, M. J. et al. A diabetes-susceptible HLA haplotype is best defined by a combination of HLA-DR and -DQ alleles. J Clin Invest 83, 830-5 (1989).

215

338. Horn, G. T., Bugawan, T. L., Long, C. M. & Erlich, H. A. Allelic sequence variation of the HLA-DQ loci: relationship to serology and to insulin-dependent diabetes susceptibility. Proc Natl Acad Sci U S A 85, 6012-6 (1988).

339. Easteal, S., Viswanathan, M. & Serjeantson, S. W. HLA-DP, -DQ and -DR RFLP types in south Indian insulin-dependent diabetes mellitus patients. Tissue Antigens 35, 71-4 (1990).

340. Erlich, H. A. et al. Association of HLA-DPB1*0301 with IDDM in Mexican-Americans. Diabetes 45, 610-4 (1996).

341. Lie, B. A. et al. A gene telomeric of the HLA class I region is involved in predisposition to both type 1 diabetes and coeliac disease. Tissue Antigens 54, 162-8 (1999).

342. Noble, J. A., Valdes, A. M., Thomson, G. & Erlich, H. A. The HLA class II locus DPB1 can influence susceptibility to type 1 diabetes. Diabetes 49, 121-5 (2000).

343. Cucca, F. et al. A correlation between the relative predisposition of MHC class II alleles to type 1 diabetes and the structure of their proteins. Hum Mol Genet 10, 2025-37 (2001).

344. Nerup, J., Mandrup-Poulsen, T. & Molvig, J. The HLA-IDDM association: implications for etiology and pathogenesis of IDDM. Diabetes Metab Rev 3, 779-802 (1987).

345. Sakurami, T. et al. HLA-DR specifications in Japanese with juvenile-onset insulin-dependent diabetes mellitus. Diabetes 31, 105-6 (1982).

346. Noble, J. A. et al. The role of HLA class II genes in insulin-dependent diabetes mellitus: molecular analysis of 180 Caucasian, multiplex families. Am J Hum Genet 59, 1134-48 (1996).

347. Spielman, R. S., Baur, M. P. & Clerget-Darpoux, F. Genetic analysis of IDDM: summary of GAW5 IDDM results. Genet Epidemiol 6, 43-58 (1989).

348. Bell, P. M., Hayes, J. R. & Stout, R. W. Lipoproteins, insulin and glycaemic control in diabetes. Horm Metab Res 16, 262-3 (1984).

349. Owerbach, D. & Gabbay, K. H. Localization of a type I diabetes susceptibility locus to the variable tandem repeat region flanking the insulin gene. Diabetes 42, 1708-14 (1993).

350. Julier, C. et al. Multiple DNA variant association analysis: application to the insulin gene region in type I diabetes. Am J Hum Genet 55, 1247-54 (1994).

351. Pugliese, A. et al. The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes. Nat Genet 15, 293-7 (1997).

352. Vafiadis, P. et al. Insulin expression in human thymus is modulated by INS VNTR alleles at the IDDM2 locus. Nat Genet 15, 289-92 (1997).

353. Kennedy, G. C., German, M. S. & Rutter, W. J. The minisatellite in the diabetes susceptibility locus IDDM2 regulates insulin transcription. Nat Genet 9, 293-8 (1995).

354. Vafiadis, P. et al. Imprinted and genotype-specific expression of genes at the IDDM2 locus in pancreas and leucocytes. J Autoimmun 9, 397-403 (1996).

355. Julier, C. et al. Insulin-IGF2 region on chromosome 11p encodes a gene implicated in HLA-DR4-dependent diabetes susceptibility. Nature 354, 155-9 (1991).

356. Risch, N. Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 46, 229-41 (1990).

357. Donner, H. et al. CTLA4 alanine-17 confers genetic susceptibility to Graves' disease and to type 1 diabetes mellitus. J Clin Endocrinol Metab 82, 143-6 (1997).

358. Abe, T. et al. CTLA4 gene polymorphism correlates with the mode of onset and presence of ICA512 Ab in Japanese type 1 diabetes. Diabetes Res Clin Pract 46, 169-75 (1999).

359. Kikuoka, N. et al. Cytotoxic T lymphocyte antigen 4 gene polymorphism confers susceptibility to type 1 diabetes in Japanese children: analysis of association with HLA genotypes and autoantibodies. Clin Endocrinol (Oxf) 55, 597-603 (2001).

360. Ide, A. et al. Association between IL-18 gene promoter polymorphisms and CTLA-4 gene 49A/G polymorphism in Japanese patients with type 1 diabetes. J Autoimmun 22, 73-8 (2004).

361. Naluai, A. T. et al. The CTLA4/CD28 gene region on chromosome 2q33 confers susceptibility to celiac disease in a way possibly distinct from that of type 1 diabetes and other chronic inflammatory disorders. Tissue Antigens 56, 350-5 (2000).

362. Chistiakov, D. A. & Turakulov, R. I. CTLA-4 and its role in autoimmune thyroid disease. J Mol Endocrinol 31, 21-36 (2003).

363. Linsley, P. S., Ledbetter, J., Peach, R. & Bajorath, J. CD28/CTLA-4 receptor structure, binding stoichiometry and aggregation during T-cell activation. Res Immunol 146, 130-40 (1995).

364. Field, L. L., Tobias, R., Thomson, G. & Plon, S. Susceptibility to insulin-dependent diabetes mellitus maps to a locus (IDDM11) on human chromosome 14q24.3-q31. Genomics 33, 1-8 (1996).

365. Todd, J. A. et al. Genetic analysis of autoimmune type 1 diabetes mellitus in mice. Nature 351, 542-7 (1991).

366. Ghosh, S. et al. Polygenic control of autoimmune diabetes in nonobese diabetic mice. Nat Genet 4, 404-9 (1993).

367. Cornall, R. J. Genetics of a multifactorial disease: autoimmune type 1 diabetes mellitus. Clin Sci (Lond) 84, 257-62 (1993).

368. Owerbach, D. & Gabbay, K. H. The HOXD8 locus (2q31) is linked to type I diabetes. Interaction with chromosome 6 and 11 disease susceptibility genes. Diabetes 44, 132-6 (1995).

369. Morahan, G., Huang, D., Tait, B. D., Colman, P. G. & Harrison, L. C. Markers on distal chromosome 2q linked to insulin-dependent diabetes mellitus. Science 272, 1811-3 (1996).

370. Esposito, L. et al. Genetic analysis of chromosome 2 in type 1 diabetes: analysis of putative loci IDDM7, IDDM12, and IDDM13 and candidate genes NRAMP1 and IA-2 and the interleukin-1 gene cluster. IMDIAB Group. Diabetes 47, 1797-9 (1998).

371. Awata, T. et al. Association of CTLA-4 gene A-G polymorphism (IDDM12 locus) with acute-onset and insulin-depleted IDDM as well as autoimmune thyroid disease (Graves' disease and Hashimoto's thyroiditis) in the Japanese population. Diabetes 47, 128-9 (1998).

372. Veijola, R. et al. Effect of genetic risk load defined by HLA-DQB1 polymorphism on clinical characteristics of IDDM in children. Eur J Clin Invest 25, 106-12 (1995).

373. Ikegami, H. et al. Identification of a new susceptibility locus for insulin-dependent diabetes mellitus by ancestral haplotype congenic mapping. J Clin Invest 96, 1936-42 (1995).

374. Colle, E., Guttmann, R. D., Seemayer, T. A. & Michel, F. Spontaneous diabetes mellitus syndrome in the rat. IV. Immunogenetic interactions of MHC and non-MHC components of the syndrome. Metabolism 32, 54-61 (1983).

375. Jacob, H. J. et al. Genetic dissection of autoimmune type I diabetes in the BB rat. Nat Genet 2, 56-60 (1992).

376. Kloting, I., Kovacs, P. & Kuttler, B. Phenotypic consequences after restoration of lymphopenia in the diabetes-prone BB/OK rat. Biochem Biophys Res Commun 239, 106-10 (1997).

377. Yokoi, N. et al. A non-MHC locus essential for autoimmune type I diabetes in the Komeda Diabetes-Prone rat. J Clin Invest 100, 2015-21 (1997).

378. Nagamine, K. et al. Positional cloning of the APECED gene. Nat Genet 17, 393-8 (1997).

379. Verge, C. F. et al. Evidence for oligogenic inheritance of type 1 diabetes in a large Bedouin Arab family. J Clin Invest 102, 1569-75 (1998).

380. Aschner, P., King, H., Traina-de-Torrado & Rodriguez, B. Glucose Intolerance in Colombia. A population-based survey in an urban community. Diabetes Care 16, 90-93 (1993).

381. Montoya, F. et al. Determinacion de Marcadores geneticos en pacientes con diabetes tipo 1 y poblacion sana. Acta Medica Colombiana 21 (1996).

382. Uribe F et al. Analisis de ligamiento genetico de la diabetes mellitus tipo 1 a marcadores de los cromosomas 2 y 11 en familias antioqueñas. IATREIA 17, 93-104 (2004).

383. Aaltonen, T. M., Valtonen, E. T. & Jokinen, E. I. Immunoreactivity of roach, Rutilus rutilus, following laboratory exposure to bleached pulp and paper mill effluents. Ecotoxicol Environ Saf 38, 266-71 (1997).

384. Chang, C. C., Huang, C. N. & Chuang, L. M. Autoantibodies to thyroid peroxidase in patients with type 1 diabetes in Taiwan. Eur J Endocrinol 139, 44-8 (1998).

385. Bikker, H., Vulsma, T., Baas, F. & de Vijlder, J. J. Identification of five novel inactivating mutations in the human thyroid peroxidase gene by denaturing gradient gel electrophoresis. Hum Mutat 6, 9-16 (1995).

386. Meloni, G. et al. Association of the ACP1 genotype with metabolic parameters upon initial diagnosis of type 1 diabetes. Med Sci Monit 9, CR105-8 (2003).

387. Johnson, W. G. et al. Pedigree analysis in families with febrile seizures. Am J Med Genet 61, 345-52 (1996).

Appendix 1. Markers typed by deCODE.

Locus	Marker	Chr	Genetic position (cM)		deCODE Distance	Cytogen etic band	Hetero- zygosity
			Marshfield	deCODE			
D1S468	AFM280we5	1	4.22	5.44	0	p36.32	72.7
D1S2870	AFMa052wg1	1	14.04	11.68	6.23	p36.31	80
D1S450	AFM247te9	1	20.61	16.99	5.32	p36.22	78.67
D1S2667	AFMa224wg9	1	24.68	20.36	3.37	p36.22	83.82
D1S434	AFM217yh8	1	29.93	21.8	1.44	p36.22	62.55
D1S2697	AFMa298yc5	1		29.07	7.27		69.22
D1S2620	AFMa114yd5	1	52.7	43.94	14.86	p36.12	41.51
D1S241	AFM211xa1	1	61.1	54.23	10.29	p35.1	56.61
D1S255	AFM260zg5	1	65.47	59.98	5.75	p34.3	73.69
D1S2861	AFM092xd11	1	72.59	68.18	8.21	p34.2	66.07
D1S2713	AFMa349yb5	1	73.81	69.43	1.25	p34.1	75.84
D1S2797	AFMb359wd1	1	75.66	70.48	1.06	p33	75.11
D1S197	AFM073xe9	1	76.27	73.67	3.19	p32.3	80.25
D1S2652	AFMa184xe1	1	80.77	78.87	5.2	p32.3	59.93
D1S476	AFM289yc1	1	85.68	80.34	1.47	p32.2	67.42
D1S2846	AFM345xb5	1	91.89	87.95	7.61	p31.3	51.87
D1S2788	AFMb351xc9	1	93.86	89.08	1.13	p31.3	81.69
D1S219	AFM161xb2	1	101.48	98.97	9.9	p31.1	79.75
D1S2841	AFM333zh9	1	106.45	105.86	6.88	p31.1	76.38
D1S500	AFM343vf9	1	107.56	106.4	0.54	p31.1	61.45
D1S430	AFM210vc9	1	109.04	107.37	0.97	p31.1	70.69
D1S2766	AFMb320yf1	1	118.14	112.11	4.74	p22.3	78.47
D1S2865	AFMa050ta5	1	120.28	114.17	2.06	p22.3	60.63
D1S435	AFM217zb2	1	125.51	117.67	3.5	p22.2	71.65
D1S206	AFM113xf6	1	134.2	124.95	7.28	p21.2	71.93
D1S495	AFM323ya5	1	136.88	126.35	1.4	p21.1	87.16
D1S2688	AFM186xg7	1	139.02	129.08	2.73	p21.1	60.8

D1S189	AFM036xc5	1	149.2	140.46	11.38	p13.1	79.82
D1S498	AFM336xb1	1	155.89	147.7	7.24	q21.3	80.49
D1S1653	CHLC.GATA43A0 4	1	164.09	154.54	6.83	q23.1	66.22
D1S484	AFM297wb9	1	169.68	160.45	5.91	q23.3	70.72
D1S2628	AFMa125ze5	1	177.86	168.68	8.23	q23.3	64.97
D1S452	AFM248wg5	1	188.85	176.39	7.7	q24.2	73.46
D1S2818	AFMc025xh9	1	198.3	187.03	10.64	q25.3	71.3
D1S413	AFM165xc9	1	212.44	198.52	11.49	q31.3	72.64
D1S2717	AFMb002zg5	1	218.46	208.15	9.62	q32.1	69.93
D1S249	AFM234wf6	1	220.65	210.82	2.67	q32.1	85.96
D1S2685	AFMa274wa5	1	222.84	213.64	2.82	q32.2	76.84
D1S245	AFM224xc1	1	227.81	216.77	3.14	q32.2	84.33
D1S205	AFM108ya3	1	229.13	217.22	0.44	q32.2	80.4
D1S237	AFM205xd8	1	232.81	222.18	4.96	q41	77.6
D1S2641	AFMA151YH5	1	238.52	230.38	8.2	q41	68.9
D1S2833	AFM321xe5	1	245.05	238.23	7.85	q42.2	84.82
D1S2709	AFMa342ye5	1	247.23	239.64	1.41	q42.2	67.72
D1S2850	AFM088xe5	1	256.26	251.42	11.78	q43	68.79
D1S2785	AFMb349xb9	1	266.27	262.68	11.26	q43	82.86
D2S323	AFM263WB5	2	5.4	4.61	0	p25.3	59.88
D2S319	AFM108XH8	2	7.6	7.81	3.2	p25.2	74.27
D2S2211	AFMA272YC5	2	15.61	18.24	10.42	p25.1	76.32
D2S398	AFMA127XB9	2	23.57	28.03	9.79	p25.1	69.57
D2S168	AFM240VF6	2	27.06	29.73	1.7	p25.1	83.1
D2S305	AFM073YA5	2	38.87	42.8	13.07	p24.1	75.72
D2S146	AFM198WC5	2	49.04	52.91	10.12	p23.2	50.93
D2S2328	AFMC014YC1	2	61	66.77	13.85	p22.1	89.12
D2S2156	AFMA119YE5	2	73.61	78.42	11.65	p16.3	74.39
D2S1364	GATA23H01	2	77.97	82.83	4.41	p16.1	61.32
D2S337	AFM275ZA9	2	80.69	86.05	3.22	p15	87.96
D2S2152	AFMA111YA9	2	87.62	94.62	8.57	p13.3	53.63
D2S2110	AFMA116ZH9	2	90.82	99.58	4.96	p13.2	78.05

D2S2116	AFMB313XD5	2	95.67	103.5	3.93	p12	71.66
D2S388	AFM333VH5	2	107.46	112.75	9.25	p11.2	74.38
D2S293	AFM205XA1	2	118.16	122.55	9.8	q12.2	83.49
D2S363	AFM303WC5	2	125.18	130.95	8.4	q14.1	67.79
D2S347	AFM289XB1	2	131.51	138.31	7.36	q14.3	77.04
D2S2271	AFMB285YH5	2	133.65	143.62	5.31	q14.3	80.92
D2S2256	AFM205xd4	2	141.62	149.54	5.92	q21.2	69.18
D2S2324	AFMC010ZA5	2	153.66	164.37	14.83	q23.3	68.83
D2S1353	ATA27H09	2	164.51	171.63	7.25	q24.1	82.78
D2S2330	AFMC015YD9	2	169.41	174.85	3.22	q24.3	83.89
D2S2188	AFMA203YE5	2	180.79	186.25	11.4	q31.1	73.75
D2S118	AFM066XC1	2	190	195.34	9.09	q32.2	78.74
D2S2321	AFMC009WH1	2	205.06	209.45	14.11	q33.3	74.21
D2S317	AFM094XF11	2	208.58	213.81	4.36	q34	58.97
D2S2361	AFMA052TC1	2	210.43	216.94	3.13	q35	73.85
D2S2382	AFMA082XC5	2	213.49	218.52	1.59	q35	80.79
D2S163	AFM234WA9	2	218.45	226.25	7.72	q35	80.73
D2S133	AFM165ZH8	2	224.33	232.12	5.88	q36.1	62.59
D2S427	GATA12H10	2	236.7	243.39	11.27	q37.1	72.98
D2S2202	AFMA239YC5	2	249.22	253.23	9.84	q37.2	68.46
D2S338	AFM276ZF5	2	250.54	253.84	0.61	q37.2	81.75
D2S395	AFM356TE5	2	261.34	262.12	8.28	q37.3	80.57
D2S140	AFM182YA5	2	263.56	265.28	3.16	q37.3	75.3
D3S4559	3PTEL25	3	1.32	1.08	0	p26.3	61.65
D3S1270	AFM122XB6	3	6.96	3.28	2.2	p26.3	72.24
D3S3630	AFMB296ZF5	3	10.7	6.22	2.95	p26.3	76.55
D3S3050	GATA88H04	3	14.46	10.68	4.46	p26.2	83.51
D3S2397	CHLC.GATA51A05	3		16.21	5.53	p26.1	66.67
D3S1515	UT440	3	18.38	19.89	3.68	p26.1	49.67
D3S3591	AFMA303YB1	3	24.89	23.9	4.02	p26.1	59.58
D3S1263	AFM079YG5	3	36.1	30.71	6.81	p25.3	87.76
D3S2385	GAAT3E04	3	38.54	33.55	2.84	p25.1	62.08

221

D3S1567	AFM234YA11	3	49.18	48.03	14.47	p24.2	49.17
D3S2335	AFM312YF5	3	50.94	50.4	2.37	p24.2	65.64
D3S1266	AFM095XC1	3	52.6	52.22	1.82	p24.1	75.89
D3S3547	AFMA108YC1	3	55.11	56.45	4.23	p24.1	71.79
D3S3521	AFM345YG5	3	63.12	66.4	9.95	p22.2	82.84
D3S3672	AFMB352XE1	3	72.21	76.71	10.31	p14.3	62.97
D3S1600	AFM308XC9	3	85.97	89.02	12.31	p14.2	75.84
D3S3551	AFMA116XA9	3	99.38	99.23	10.21	p13	87.58
D3S3653	AFMB326XB5	3	107.19	106.92	7.7	p12.3	66.72
D3S1271	AFM126ZC5	3	117.76	115.24	8.32	q12.2	67.37
D3S3045	GATA84B12	3	124.16	120.29	5.05	q13.12	81.72
D3S1558	AFM057XA5	3	133.93	129.21	8.92	q13.31	73.84
D3S1589	AFM290ZF1	3	141.79	136.43	7.22	q21.3	68.36
D3S3637	AFMB309XA9	3	149.8	146.26	9.83	q22.2	90.02
D3S1569	AFM240XB2	3	158.38	154.11	7.85	q24	77.64
D3S1308	AFM240VE1	3	165.85	160.67	6.56	q25.1	71.7
D3S1607	AFM319YB1	3	172.27	168.31	7.65	q25.31	79.88
D3S1564	AFM224YA3	3	180.8	175.21	6.89	q26.2	47.97
D3S3725	AFMA082ZD9	3	181.87	178.27	3.06	q26.31	85.51
D3S3041	CHLC.GATA81H05	3	188.29	185.92	7.65	q26.32	77.63
D3S3592	AFMA306XG9	3	198.68	195.05	9.13	q27.1	79.92
D3S1580	AFM270ZG9	3	207.73	206.43	11.38	q28	84.57
D3S3663	AFMB343ZF5	3	214.45	213.15	6.72	q28	0
D3S240	MFD30	3	218.66	219.97	6.82	q29	22.48
D3S1265	AFM087YB7	3	222.83	222.34	2.37	q29	85.98
D3S1311	AFM254VE1	3	224.88	225.05	2.71	q29	79.78
D4S412	AFM196XB6	4	4.74	4.57	0	p16.3	79.85
D4S2935	AFMA247XB5	4	13.96	14.07	9.51	p16.1	61.69
D4S394	AFM037YG1	4	16.01	15.28	1.21	p16.1	80.15
D4S2928	AFMA223WG1	4	23.17	23.84	8.56	p16.1	56.85
D4S3341	GGAT18G02	4	29.68	32.25	8.41	p15.32	66.3
D4S404	AFM158XC7	4	37.16	41	8.75	p15.2	74.06

222

D4S3022	AFMA045ZD5	4	39.3	43.89	2.9	p15.2	88.28
D4S391	AFM016XF3	4	43.59	48.95	5.06	p15.2	86.2
D4S2912	AFMA126XA9	4	47.58	48.97	0.01	p15.1	69.64
D4S1587	AFM288YB5	4	50.53	55.94	6.98	p15.1	73.26
D4S405	AFM161YF6	4	56.95	62.1	6.16	p14	84.28
D4S428	AFM255ZF1	4	64.24	71.03	8.93	q12	67.07
D4S2389	ATA22B11	4	79.69	84.72	13.69	q13.3	65.12
D4S1553	AFM210WH8	4	90.28	92.45	7.73	q21.22	52.84
D4S423	AFM224XD6	4	100.75	102.7	10.25	q22.1	80.97
D4S1572	AFM265VA9	4	107.95	110.14	7.44	q24	84.82
D4S406	AFM164TF6	4	117.06	117.48	7.35	q25	88.46
D4S1615	AFM336XH5	4	128.31	129.75	12.27	q28.1	75.8
D4S1575	AFM268YB1	4	132.05	134.67	4.92	q28.3	59.61
D4S424	AFM225XB10	4	144.56	142.05	7.38	q31.21	78.37
D4S1586	AFM288VB9	4	147.06	144.42	2.37	q31.21	76.22
D4S2962	AFMB292XB5	4	152.98	148.06	3.64	q31.23	76.76
D4S3046	AFMA071WG5	4	162.47	159.33	11.27	q32.2	75.41
D4S1539	AFM185XE1	4	176.19	173.75	14.42	q34.1	66.59
D4S1607	AFM310WF5	4	183.63	184.14	10.39	q34.3	81.32
D4S3041	AFMA062TC9	4	187.64	189.52	5.38	q35.1	68.29
D4S408	AFM165XC11	4	195.06	193.5	3.97	q35.1	76.91
D4S1540	AFM185XF8	4	199.93	195.66	2.16	q35.1	55.77
D4S2930	AFMA224XH1	4	208.07	209.27	13.61	q35.2	77.3
D5S1981	AFMA217ZH1	5	1.72	1.21	0	p15.33	74.25
D5S2088	AFMA050ZD9	5	9.41	12.05	10.84	p15.33	71.34
D5S2054	AFMB355WB1	5	14.3	16.64	4.59	p15.32	65.24
D5S635	AFM276YB9	5	14.91	18.11	1.46	p15.32	71.41
D5S1486	CHLC.GATA21F08	5	21.1	27.25	9.14	p15.2	66.23
D5S2081	AFM347TA5	5	24.48	31.94	4.69	p15.2	63.54
D5S2031	AFMB312XH5	5	36.25	42.23	10.29	p14.3	79.09
D5S674	AFM331ZE9	5	47.09	56.71	14.48	p13.3	79.93
D5S2021	AFMB285XC9	5	54.79	61.33	4.62	p13.2	63.96

223

D5S628	AFM254wf1	5	60.92	71.14	9.81	q11.2	46.78
D5S2072	AFM312ZH5	5	73.35	80.82	9.68	q12.3	73.84
D5S424	AFM214ZG9	5	81.95	93.23	12.41	q13.3	75.63
D5S1397	UT6142	5	85.25	98.14	4.91	q14.1	83.75
D5S1725	GATA89G08	5	97.82	106.88	8.74	q14.3	75.44
D5S409	AFM184YB6	5	109.63	114.89	8.01	q21.1	59.52
D5S2027	AFMB304XB5	5	119.5	120.95	6.06	q22.1	77.42
D5S2065	AFMC031XB5	5	122.01	122.49	1.54	q22.3	68.6
D5S490	AFM214YG1	5	133.65	134.52	12.03	q23.2	52.76
D5S2017	AFMB074XG1	5	145.21	145.35	10.83	q31.3	82.26
D5S2090	AFMA052TH5	5	150.34	153.72	8.37	q32	81.19
D5S410	AFM191xd8	5	156.47	162.45	8.73	q33.2	77.54
D5S2066	AFMC033XH1	5	165.13	171.45	9	q34	64.12
D5S400	AFM112YB6	5	174.8	181.34	9.89	q34	82.99
D5S1973	AFMA196XG1	5	175.34	184.16	2.83	q35.1	42.49
D5S2069	AFM302WC9	5	182.35	194.4	10.23	q35.2	73.29
D5S469	AFM137XF6	5		203.46	9.06	q35.3	48.87
D5S2006	AFMB005WF9	5	197.54	211.06	7.61	q35.3	71.8
D6S244	91232/91316	6		5.66	0	p25.3	51.97
D6S1617	AFM205xh4	6	7.02	12.18	6.52	p25.1	84.5
D6S309	AFM265ZH9	6	14.07	21.6	9.43	p24.3	84.9
D6S1279	CHLC.GATA82E04	6		30.35	8.75	p24.1	74.65
D6S422	AFM234XA3	6	35.66	44.68	14.33	p22.3	77.6
D6S1660	AFMb355wg5	6	40.14	48.37	3.69	p22.3	77.37
D6S273	AFM142XH6	6	44.96	53.25	4.87	p21.33	73.72
D6S1575	AFM172XB6	6	60.44	63.69	10.45	p21.1	82.89
D6S459	AFM312XC5	6	69.66	72.59	8.9	p21.1	58.17
D6S452	AFM301WB9	6	73.13	74.58	1.99	p12.3	82.52
D6S1557	AFMA191ZH5	6	82.59	87.06	12.48	q13	71.09
D6S460	AFM312YB9	6	89.83	93.4	6.34	q14.1	80.72
D6S458	AFM311WA5	6	100.91	102.77	9.38	q16.1	56.39
D6S268	AFM115XH2	6	114.93	115.03	12.25	q21	76.37

D6S1712	AFMA074WF5	6	122.51	123.79	8.77	q22.31	72.83
D6S1656	AFMB338ZF9	6	131.07	134.76	10.97	q23.2	78.54
D6S270	AFM127XB2	6	135.47	138.78	4.02	q23.2	73.96
D6S1009	GATA32B03	6	137.74	142.71	3.93	q23.3	78.06
D6S1637	AFMB314YD5	6	147.13	154.79	12.07	q24.3	76.97
D6S441	AFM269ZE1	6	154.1	164.5	9.71	q25.2	86.08
D6S419	AFM218XB6	6	159.98	170.85	6.35	q25.3	76.11
D6S1581	AFMA247ZD9	6	164.78	174.43	3.58	q26	68.13
D6S305	AFM242ZG5	6	166.39	176.93	2.5	q26	84.65
D6S1599	AFMA342VB5	6	169.95	179.13	2.2	q26	72.78
D6S1277	GATA81B01	6	173.31	180.81	1.68	q27	77.08
D6S297	AFM212YF6	6	182.11	187.65	6.84	q27	60.94
D6S503	GGAA8D08	6	184.51	190.17	2.53	q27	69.85
D6S446	AFM290XF5	6	189	194.22	4.05	q27	70.01
D7S2474	AFMB017YH1	7	3.13	3.71	0	p22.3	78
D7S2201	GATA61G06	7	10.68	12.21	8.5	p22.1	60.24
D7S513	AFM217YC5	7	17.74	23.24	11.03	p21.3	85.38
D7S664	AFM281VC9	7	20.84	26.72	3.49	p21.3	72.05
D7S1795	GATA21B01	7	36.03	38.61	11.88	p15.3	62.58
D7S2463	AFM044XB8	7	38.48	40.51	1.9	p15.3	72.5
D7S516	AFM224XG5	7	41.69	45.36	4.85	p15.1	72.18
D7S2250	AFM098xc1	7	54.11	55.9	10.54	p14.2	80.87
D7S2541	AFMA043XB5	7	61.53	63.45	7.55	p14.1	77.19
D7S691	AFM350VA9	7	63.67	64.93	1.49	p14.1	76.61
D7S519	AFM238VB12	7	69.03	70.24	5.31	p13	82.09
D7S2429	AFMA153TF5	7	76.71	80.22	9.98	q11.21	71.78
D7S502	AFM199VH8	7	78.65	80.93	0.7	q11.22	85.87
D7S2500	AFMB335WD9	7	80.42	83.71	2.78	q11.22	66.82
D7S672	AFM289VE9	7	84.52	86.43	2.72	q11.22	86.02
D7S2443	AFMA222WA5	7	93.1	94.47	8.04	q21.11	76.77
D7S2485	AFM051xb12	7		100.34	5.88	q21.11	61.57
D7S657	AFM263YD9	7	104.86	105.59	5.24	q21.3	78.58
D7S554	AFM248TE5	7	109.66	109.67	4.08	q21.3	72.82

225

D7S2509	AFM273VG5	7	112.85	114.48	4.81	q22.1	62.98
D7S2459	AFMA305YE9	7	119.81	120.05	5.57	q22.3	76.49
D7S2418	AFMA124XB9	7	122.48	122.15	2.1	q31.1	73.43
D7S530	AFM249XF9	7	134.55	133.2	11.05	q32.2	75.57
D7S2533	AFM087ZF11	7	138.42	141.74	8.54	q33	51.65
D7S684	AFM312WB5	7	147.22	149.78	8.04	q34	82.72
D7S1824	GATA32C12	7	149.9	152.03	2.25	q34	85.46
D7S661	AFM277ZF5	7	155.1	155.04	3.01	q35	73.53
D7S1805	GATA4H10	7	161.21	165.55	10.51	q36.1	78.28
D7S2447	AFM175YG1	7	175.77	179.62	14.07	q36.2	66.62
D7S2423	AFMA133ZC9	7	181.97	192.24	12.62	q36.3	68.87
D8S264	AFM143XD8	8	0.73	3.54	0	p23.3	83.21
D8S1819	AFMA050ZD5	8	9.96	16.91	13.37	p23.1	79.82
D8S520	AFM234YE5	8	20.61	21.48	4.58	p23.1	78.69
D8S516	AFM214ZB4	8	17	22.5	1.02	p23.1	61.41
D8S258	AFM107XB6	8	41.55	36.13	13.63	p21.3	70.92
D8S1734	AFMA337ZH5	8	46.26	42.1	5.97	p21.2	67.95
D8S1820	AFMA051ZB1	8	56.04	49.17	7.07	p21.1	75.01
D8S505	AFM198TB2	8	60.87	55.98	6.81	p12	78.85
D8S1737	AFMA344WA5	8	67.27	69.34	13.36	q11.23	51.39
D8S1763	AFMB307XB9	8	78.78	73.88	4.54	q12.1	65.08
D8S512	AFM206XH4	8	81.68	78.06	4.17	q12.3	59.32
D8S279	AFM203WC1	8	91.46	86.52	8.46	q13.3	86.62
D8S1707	AFM147YB6	8	101.69	97.46	10.95	q21.3	67.04
D8S1778	AFMB327YH1	8	110.2	106.89	9.43	q22.2	84.49
D8S1762	AFMB307XB5	8	112.42	108.03	1.14	q22.2	75.83
D8S1470	GATA68H01	8		118.7	10.67	q23.3	67.43
D8S1799	AFM283XB5	8	133.64	131.69	12.99	q24.13	85.03
D8S1793	AFMC014XF5	8	137.92	135.13	3.43	q24.21	81.64
D8S263	AFM141XA5	8	141.68	141.47	6.35	q24.21	65.45
D8S1746	AFMB018WF5	8	149.46	150.11	8.64	q24.22	64.19
D8S1783	AFMB340XD5	8	154.02	152.34	2.23	q24.23	75.13
D8S274	AFM182XA3	8	154.02	153.61	1.27	q24.23	75.32

D8S1743	AFMB011XG1	8	162.94	160.11	6.5	q24.3	82.64
D9S1779	AFM026TG9	9	0	0	0	p24.3	57.69
D9S1686	AFM238ZC1	9	14.78	11.93	11.93	p24.1	58.54
D9S286	AFM344YC9	9	18.06	18.53	6.6	p24.1	87.91
D9S168	AFM158XF12	9	21.88	24.28	5.75	p23	79.28
D9S1808	AFMA297WE9	9	26.24	28.24	3.96	p23	59.84
D9S235	UT727	9	27.32	32.65	4.41	p22.3	59.21
D9S171	AFM186XC3	9	42.73	47.16	14.52	p21.3	78.13
D9S2149	GATA137A02	9	54.5	55.3	8.14	p21.1	77.37
D9S1777	AFMA119VG9	9	64.72	67.34	12.04	q21.11	49.89
D9S1876	AFMA054ZE1	9	67.93	70.42	3.08	q21.12	82.29
D9S175	AFM224ZH10	9	70.33	72.8	2.38	q21.13	82.23
D9S1843	AFMB321YF9	9	80.31	79.87	7.07	q21.31	77.62
D9S307	MFD308	9		89.23	9.36	q21.33	71.26
D9S283	AFM318XC9	9		95	5.77		78.68
D9S287	AFM347ZF5	9	103.42	101.54	6.54	q22.32	69.02
D9S261	AFM210ZE7	9	117.37	114.65	13.11	q31.2	59.92
D9S1675	AFM303TA9	9	120.04	117.44	2.8	q31.3	81.61
D9S1828	AFM205TG6	9	120.04	117.67	0.23	q31.3	79.88
D9S1824	AFMB025YE1	9	122.23	123.56	5.88	q32	79.47
D9S934	GATA64G07	9	127.98	127.72	4.16	q33.1	75.65
D9S1682	AFMA061XD9	9	132.09	132.09	4.37	q33.2	68.1
D9S1825	AFMB029XG1	9	136.47	136.35	4.26	q33.3	78.75
D9S1830	AFMB036XF5	9	145.21	149.36	13.01	q34.13	71.65
D9S1793	AFMA191YB5	9		150.42	1.07		82.09
D9S1838	AFMB303ZG9	9	163.84	164.1	13.68	q34.3	85.81
D10S249	AFM207WD12	10	2.13	1.19	0	p15.3	78.35
D10S602	AFM343VD9	10	4.32	5.71	4.52	p15.3	68.88
D10S1218	CHLC.ATA10G07	10	5.21	11.3	5.59	p15.2	11.62
D10S189	AFM063XF4	10	19	20.56	9.26	p14	69.37
D10S1649	AFM155ZG1	10	27.19	26.55	5.99	p14	85.62
D10S1430	GATA84C01	10	33.18	32.11	5.56	p13	73.78
D10S570	AFM268ZH1	10	33.48	32.73	0.62	p13	81.85

D10S191	AFM066XA1	10	37.9	37.42	4.69	p13	81.12
D10S1653	AFMA175YD1	10	40.36	38.92	1.5	p13	81.38
D10S1661	AFMA204XG9	10	42.5	41.4	2.47	p13	79.79
D10S600	AFM338TA5	10	54.23	54.33	12.93	p12.1	80.18
D10S1746	AFMC012XC5	10	65.97	65.42	11.1	q11.21	60.47
D10S196	AFM115XF2	10	70.23	71.95	6.53	q11.23	71.96
D10S609	CHLC.GATA2G08	10	80.19	82.33	10.38	q21.2	69.95
D10S537	AFM203XC5	10	91.13	91.46	9.12	q22.1	86.09
D10S580	AFM284VF5	10	96.72	97.94	6.48	q22.3	72.77
D10S1677	AFMA233WG5	10	100.08	100.72	2.78	q22.3	71.76
D10S1753	AFM287ZE1	10	112.58	113.42	12.7	q23.31	74.46
D10S185	AFM019TH6	10	116.34	116.07	2.65	q23.33	80.95
D10S1267	AFMA301XE1	10	125.41	125.15	9.08	q24.32	87.36
D10S597	AFM331XA9	10	128.73	129.57	4.42	q25.1	61.21
D10S1693	AFMA344XC1	10	137.39	141.5	11.94	q26.11	79.3
D10S1656	AFMA184XD9	10	149.25	154.8	13.3	q26.13	75.95
D10S575	AFM270XB1	10	153.78	160.19	5.39	q26.2	71.41
D10S1676	AFMA232YH9	10	160.04	164.6	4.41	q26.2	68.41
D10S1651	AFMA162XG1	10	168.77	176.24	11.64	q26.3	84.34
D10S212	AFM198ZB4	10	170.94	181.66	5.41	q26.3	65.13
D11S4046	AFMB042YF5	11	2.79	0	0	p15.5	85.5
D11S1760	AFM114YC9	11	11.67	8.79	8.79	p15.4	74.38
D11S1331	AFM277WG1	11	12.92	12.38	3.59	p15.4	72.31
D11S4170	AFMB341YE5	11	20.93	23.4	11.02	p15.2	78.51
D11S902	AFM072YD3	11	21.47	26.23	2.82	p15.1	82.46
D11S928	AFM234WH12	11	26.34	34.9	8.68	p15.1	65.69
D11S4080	AFMA109ZG9	11	33.02	43.29	8.38	p14.3	80.91
D11S1776	AFMA041ZE5	11	40.12	49.46	6.17	p13	84.56
D11S1360	AFM362TB9	11	50.88	58.18	8.72	p12	65.96
D11S4191	AFM338WC1	11	60.09	66.18	8	q12.1	84.63
D11S4087	AFMA152YH1	11	68.55	73.81	7.63	q13.2	46.13
D11S937	AFM256ZB5	11	79.98	85.42	11.61	q14.1	90.73
D11S1780	AFMA082WB9	11	90.29	94.12	8.7	q14.2	73.5

D11S1886	AFM347TE5	11	100.62	107.71	13.59	q22.3	51.77
D11S4206	AFMA102XF9	11	105.17	110.81	3.1	q22.3	71.16
D11S908	AFM120XE9	11	108.59	118.66	7.86	q23.3	72.37
D11S4089	AFMA162WF5	11	119.07	126.83	8.17	q23.3	75.93
D11S4151	AFMB290YH1	11	127.33	135.73	8.9	q24.2	80.56
D11S2367	GATA72A01	11		148.23	12.5		45.05
D11S969	AFM205VF10	11		154.38	6.15	q25	69.32
D12S352	AFM303XD9	12	0	0.68	0	p13.33	65.05
D12S1656	AFMB292WE5	12	2.79	4.23	3.55	p13.33	70.02
D12S372	GATA4H03	12	6.42	8.63	4.4	p13.32	74.88
D12S374	GATA7F09	12	14.23	16.84	8.21	p13.31	72.54
D12S336	AFM273ZC9	12	19.68	24.76	7.91	p13.31	66.1
D12S364	AFM345WE1	12	30.6	31.97	7.21	p13.1	86.98
D12S1591	AFMA119XE5	12	43.38	44.81	12.85	p12.1	80.68
D12S1337	UT7594	12		51.8	6.99	p11.22	57.54
D12S1704	AFMA047WF1	12	50.9	52.93	1.13	p11.22	65.23
D12S85	AFM122XF6	12	61.34	61.77	8.84	q13.11	62.93
D12S368	AFMA128YD5	12	66.03	67.92	6.15	q13.13	83.28
D12S83	AFM112YF4	12	75.17	75.48	7.56	q14.1	71.88
D12S1294	GATA73H09	12		83.69	8.21	q15	81.43
D12S326	AFM238WA1	12	86.4	92.97	9.28	q21.2	78.93
D12S351	AFM302WD9	12	95.56	103.1	10.13	q21.33	77.02
D12S95	AFM207VE1	12	96.09	104.42	1.32	q22	78.68
D12S1346	AFM211XF2	12	97.16	106.03	1.6	q22	56.02
D12S2081	GATA7A02	12	101.98	110.59	4.56	q23.1	65.46
D12S1636	AFMA336XC1	12	114.28	120.17	9.57	q23.3	74.71
D12S1613	AFMA205YG9	12	116.08	124.08	3.91	q23.3	53.89
D12S1583	AFMA106WH5	12	119.55	127.22	3.14	q24.11	84.82
D12S354	AFM304WH5	12	123.77	133.66	6.44	q24.21	75.19
D12S369	AFMA142ZC5	12	125.31	134.33	0.67	q24.21	75.22
D12S366	AFM351TB9	12	133.33	140.32	5.99	q24.23	81.03
D12S2073	GATA10C07	12	139.61	144.14	3.81	q24.31	70.13
D12S2078	CHLC.GATA32F0	12		155.99	11.85	q24.32	76.5

	5						
D12S1609	AFMA197ZD9	12	153.33	160.72	4.73	q24.32	63.01
D12S1723	AFMA082ZE9	12	164.63	173.29	12.57	q24.33	66.84
D13S175	AFM249XB1	13	6.03	3.24	0	q12.11	70.85
D13S232	MFD299	13	6.99	8.98	5.74	q12.12	75.45
D13S1243	AFMA217YB5	13	9.79	11.57	2.59	q12.12	75.4
D13S221	AFM248WC1	13	12.91	16.88	5.31	q12.13	82.06
D13S217	AFM205XH12	13	17.21	22.89	6.01	q12.3	74.81
D13S289	AFM321XB1	13	21.51	28.14	5.26	q12.3	72.69
D13S171	AFM255ZE9	13	25.08	31.99	3.84	q13.1	68.07
D13S219	AFM225XE5	13	28.87	36.49	4.51	q13.3	58.18
D13S218	AFM210ZB2	13	32.9	40.44	3.95	q13.3	63.56
D13S272	AFM120XA3	13	45.55	54.31	13.87	q14.2	69.83
D13S279	AFM284ZA9	13	53.17	65.83	11.52	q21.33	77.14
D13S271	AFM115XA11	13	64.97	80.33	14.5	q31.1	79.32
D13S1241	AFMA204YH9	13	76.26	91.34	11.01	q32.1	80.43
D13S1256	AFMB291WF1	13	84.87	100.66	9.32	q33.1	69.06
D13S1809	GATA135E01	13	90.27	107.62	6.96	q33.2	63.48
D13S1315	AFMA058XD5	13	102.73	119	11.38	q34	76.47
D13S293	AFMA127XH5	13	114.98	131.92	12.92	q34	52.7
D14S261	AFM238YD6	14	6.46	9.53	0	q11.2	69.27
D14S1043	AFM324TC9	14	9.36	9.83	0.29	q11.2	49.45
D14S1280	GATA31B09	14	25.87	23.04	13.21	q12	68.03
D14S262	AFM240VC5	14	28.01	25.73	2.7	q12	63.23
D14S1071	AFMA102ZD9	14	31.75	29.19	3.46	q12	71.84
D14S741	GATA74A05	14	36.76	35	5.81	q13.1	68.97
D14S70	AFM191VE1	14	40.11	37.63	2.63	q13.1	76.88
D14S75	AFM214YG5	14	44.06	43.2	5.56	q13.3	76.45
D14S976	AFMA116ZB5	14	50.5	46.8	3.61	q21.3	82.31
D14S978	AFMA122YA5	14	53.19	51.44	4.63	q22.1	84.3
D14S276	AFM292WE1	14	56.36	57.33	5.89	q22.3	74.37
D14S1011	AFMB027YC9	14	74.96	70.01	12.68	q24.1	76.96
D14S258	AFM224ZF12	14	76.28	70.4	0.39	q24.2	76.67

D14S1433	GATA169E06	14	84.69	74.94	4.54	q24.3	56.52
D14S616	GATA70B06	14	92.69	85.45	10.5	q31.3	67.65
D14S67	AFM137XH12	14	95.89	88.68	3.24	q31.3	89.54
D14S1044	AFM324VG5	14	99.88	92.86	4.17	q32.11	63.11
D14S81	AFM260XB1	14	108.22	99.25	6.4	q32.12	79.1
D14S1054	AFM361TA9	14	113.17	102.52	3.27	q32.13	75.93
D14S987	AFM161YD12	14	114.81	106.22	3.7	q32.2	74.44
D14S1019	AFMB323YE9	14	118.68	109.31	3.09	q32.2	41.91
D14S1426	GATA136B01	14	125.88	117.81	8.5	q32.2	77.37
D14S542	TF66(GT)/TF67(CA)	14		126.06	8.26	q32.33	76.92
D15S128	AFM273YF9	15	6.11	6.06	0	q11.2	80.48
D15S975	AFMA216ZC9	15	13.06	13.46	7.4	q12	62.78
D15S1019	AFMB336YF1	15	19.12	21.44	7.98	q13.1	58.38
D15S165	AFM248VC5	15	20.24	23.51	2.07	q13.3	70.77
D15S231	GTAT1B2	15	24.06	27.43	3.92	q13.3	48.9
D15S118	AFM112XA1	15	32.58	35.23	7.8	q14	73.96
D15S1012	AFMB298WH9	15	35.95	39.39	4.16	q15.1	72.74
D15S146	AFM070XD7	15	39.72	43.14	3.75	q15.1	69.07
D15S1016	AFMB324YH9	15	47.29	52.71	9.57	q21.2	90.73
D15S117	AFM098YG1	15	51.21	58.09	5.38	q21.3	79.64
D15S988	AFMA301ZE9	15	66.9	70.39	12.3	q22.33	53.9
D15S114	AFM019TC9	15	72.94	82.51	12.13	q24.3	64.6
D15S1005	AFMB080ZH9	15	75.27	86.82	4.31	q25.1	80.6
D15S158	AFM234VF12	15	86.81	100.89	14.07	q26.1	69.46
D15S1004	AFMB077YD5	15	98.44	110.96	10.07	q26.2	69.96
D15S130	AFM072YB11	15	100.59	111.55	0.59	q26.2	73.77
D15S816	GATA73F01	15	100.59	114.36	2.81	q26.2	68.74
D15S157	AFM217ZG1	15	102.21	116.14	1.78	q26.2	41.86
D15S212	AFM331VB5	15	109.29	124.08	7.94	q26.2	74.81
D15S120	AFM164ZC9	15	112.58	130.4	6.32	q26.3	61.39
D16S521	AFMA139WG1	16	1.08	1.15	0	p13.3	70.21
D16S3065	AFMA323YF5	16	8.16	10.64	9.49	p13.3	76.48

D16S423	AFM249YC5	16	10.36	14.97	4.33	p13.3	76.76
D16S418	AFM225XD2	16	14.77	20.61	5.64	p13.2	85.25
D16S3062	AFMA305XB1	16	27.05	32.07	11.46	p13.12	75.61
D16S500	AFM112XG5	16	28.3	33.31	1.25	p13.12	81.48
D16S410	AFM165YB6	16	35.44	41.83	8.52	p12.3	58.97
D16S3068	AFMA349ZH9	16	48.53	51.27	9.43	p12.1	79.43
D16S3080	AFMB068ZB9	16	59.68	61.1	9.83	q12.1	76.89
D16S3034	AFMA184XG1	16	69.05	68.27	7.17	q12.2	65.16
D16S3057	AFMA299YF1	16	77.13	76.36	8.08	q13	71.67
D16S514	AFM330VD9	16	81.15	83.1	6.74	q21	81.28
D16S503	AFM274YA5	16	83.55	84.26	1.16	q21	81.5
D16S515	AFM340YE5	16	92.1	94.69	10.43	q23.1	84.86
D16S516	AFM350VD1	16	100.39	101.24	6.55	q23.1	74.81
D16S505	AFM296TB1	16	108.96	108.07	6.83	q23.2	74.54
D16S763	GATA2F07	16		120.63	12.56	q24.1	69.79
D16S2621	GATA71F09	16	130.41	131.46	10.83	q24.2	80.66
D17S849	AFM234WG3	17	0.63	0.63	0	p13.3	70.77
D17S831	AFM058xf4	17	6.6	7.47	6.84	p13.3	84.63
D17S1832	AFMB022WF9	17	13.07	16.79	9.32	p13.2	71.88
D17S804	AFM225ZC1	17	21.01	28.96	12.17	p13.1	61.78
D17S799	AFM192YH2	17	31.96	37.94	8.98	p12	77.07
D17S922	AFM197XH6	17	35.55	41.55	3.61	p12	62.36
D17S839	AFM200YB12	17	37.8	43.77	2.21	p12	58.81
D17S1824	AFMB002YC1	17	49.67	53.22	9.46	q11.2	82.15
D17S2194	GATA169F02	17	54.17	58.94	5.72	q11.2	71.27
D17S1867	AFM296TB5	17	59.32	66.96	8.02	q12	70.63
D17S1299	GATA25A04	17	62.01	71.13	4.17	q21.2	74.53
D17S1868	AFM301YE5	17	64.16	76.3	5.17	q21.32	77.36
D17S1795	AFMA154ZA9	17	68.44	77.3	1	q21.33	74.83
D17S957	AFM323WD9	17	80.38	89.64	12.33	q23.2	53.83
D17S794	AFM168XD12	17	83.4	94.36	4.73	q23.3	57.69
D17S944	AFM277VG9	17	82.56	95.11	0.75	q23.3	76.77
D17S940	AFM268YD5	17	89.89	102.6	7.49	q24.2	54.83

D17S1351	AFMA346XG5	17	95.99	108.83	6.23	q24.3	76.5
D17S1603	AFMA135XD5	17	102.99	117.82	8.99	q25.1	67.8
D17S1847	AFMB310YF5	17	111.22	126.71	8.89	q25.3	59.73
D17S836	AFM163YG1	17	112.92	128.03	1.32	q25.3	60.66
D17S784	AFM044XG3	17	116.86	132.74	4.71	q25.3	76.83
D18S1105	AFM182XG5	18	2.84	3.97	0	p11.32	69.85
D18S63	AFM205TD6	18	8.3	9.84	5.87	p11.31	77.5
D18S967	GATA116D12	18	19.64	21.74	11.91	p11.31	54.01
D18S464	AFM259VH9	18	31.17	33.37	11.63	p11.22	50.22
D18S1107	AFMA289WE1	18	51.21	47.29	13.92	q11.2	76.38
D18S877	GATA64H04	18	54.4	54.03	6.75	q12.1	71
D18S1102	AFMA224WB1	18	62.84	59.89	5.86	q12.2	78.34
D18S474	AFM295XH1	18	71.32	68.99	9.1	q21.1	80.68
D18S450	AFM191VC7	18	68.91	70.78	1.79	q21.1	77.65
D18S64	AFM212XG5	18	84.8	77.54	6.76	q21.32	72.18
D18S1134	AFMB330ZC5	18	88.62	87.2	9.65	q21.32	72.15
D18S1147	AFMA049ZE5	18	90.6	88.5	1.31	q21.33	85.35
D18S465	AFM260YH1	18	100.11	92.91	4.4	q22.1	76.77
D18S469	AFM116YG11	18	109.18	106.38	13.47	q22.3	64.95
D18S554	AFM296WD5	18	119.44	117.94	11.57	q23	84.59
D19S886	AFMA310WD9	19	0	0.9	0	p13.3	67.84
D19S591	GATA44F10	19	9.84	10.59	9.69	p13.3	74.16
D19S427	MFD319	19		20.75	10.17	p13.3	76.64
D19S584	ATA19D12	19	34.25	33.17	12.41	p13.2	41.37
D19S226	AFM256YC9	19	42.28	37.98	4.81	p13.12	85.86
D19S593	GATA64E12	19	45.48	43.02	5.04	p13.11	64.62
D19S414	AFM295XG9	19	54.01	55.81	12.78	q12	74.2
D19S208	AFM116XC7	19	59.36	63.1	7.3	q13.12	70.2
D19S903	AFMB039YE9	19	69.5	73.21	10.11	q13.31	77.09
D19S867	AFMa115wg5	19	77.54	84.6	11.39	q13.33	80.9
DG19S186	Build31.chr19.527 92623-L	19		92.14	7.54	q13.41	43.2

D19S572	AFMA083WB5	19	88.85	99.88	7.74	q13.42	84.61
D19S418	AFM319VB5	19	92.56	107.71	7.83	q13.42	61.29
D19S605	AFMB005WH1	19	95.26	108.45	0.74	q13.42	58.69
D19S573	AFMA054YF1	19	100.38	113.23	4.78	q13.43	42.69
D20S113	AFM205TH8	20	8.97	7.68	0	p13	45.5
D20S882	AFMB290WH5	20	15.05	18.3	10.62	p12.3	71.12
D20S115	AFM218YG3	20	21.15	25.35	7.05	p12.3	67.43
D20S189	AFM292XB5	20	30.56	34.88	9.53	p12.2	71.96
D20S904	AFM291WE5	20	37.65	42.54	7.66	p12.1	78.72
D20S912	AFMA070XG5	20	46.71	52.19	9.64	p11.23	80.72
D20S859	AFM189XB6	20	51.36	60.73	8.55	q11.23	74.15
D20S107	AFM142XH4	20	55.74	63.19	2.46	q12	78.4
D20S108	AFM163YH8	20	57.38	65.64	2.45	q12	73.75
D20S481	GATA47F05	20	62.32	70.97	5.33	q13.12	84.13
D20S838	AFMA139WF5	20	64.88	71.23	0.27	q13.12	64.54
D20S178	AFM240vd6	20	66.16	75.47	4.23	q13.13	80.15
D20S1083	GAAT1F1	20	77.75	83.33	7.86	q13.2	60.46
D20S100	AFM057XA3	20	84.78	90.85	7.52	q13.31	74.49
D20S171	AFM046XF6	20	95.7	100.84	9.99	q13.32	80.1
D21S1904	AFMB356WG1	21	2.13	7.23	0	q21.1	53.02
D21S1432	GATA11C12	21	2.99	8.75	1.53	q21.1	69.89
D21S1899	AFMB321WF5	21	9.72	16.2	7.45	q21.1	85.08
D21S1905	AFMB361YE9	21	11.3	17.29	1.09	q21.1	77.67
D21S1902	AFM268XG1	21	13.05	19.54	2.25	q21.1	62.08
D21S272	AFM206XG5	21	19.39	23.96	4.42	q21.2	77.96
D21S1442	GATA24H09	21	24.73	29.5	5.53	q21.3	83.89
D21S1909	AFM288YG5	21	28.48	33.81	4.31	q22.11	79.96
D21S1898	AFMB308XE5	21	31.26	38.62	4.81	q22.11	78.64
D21S1895	AFMB280XD9	21	33.84	41.37	2.75	q22.12	84.97
D21S1252	AFM261zg1	21	35.45	44.04	2.67	q22.13	84
D21S1919	AFMA085ZA5	21	38.65	45.79	1.75	q22.13	83.87
D21S270	AFM031XC5	21	38.08	46.25	0.46	q22.13	86.61
D21S1255	AFM283XH9	21	39.22	48.28	2.04	q22.2	82.92

234

D21S266	AFM234XG9	21	45.87	57.33	9.05	q22.3	55.56
D22S420	AFM217XF4	22	4.06	5.72	0	q11.1	72.79
D22S427	AFM288we5	22	8.32	6	0.28	q11.21	64.29
D22S539	AFMA037ZD1	22	14.44	15.82	9.82	q11.22	53.4
D22S1174	AFM309WD5	22	19.32	20.03	4.21	q11.23	69.95
D22S315	AFM183xe9	22	21.47	23.91	3.89	q12.1	77.8
D22S1154	AFMA298YB5	22	23.37	25.73	1.81	q12.1	69.71
D22S531	UT5900	22	27.48	36.2	10.47	q12.2	58.41
D22S1265	ACT2A09	22	32.93	40.84	4.64	q12.3	68.22
D22S276	AFM165ZA5	22	47.31	51.21	10.38	q13.2	67.27
D22S928	AFMA048WA5	22	52.08	59.39	8.18	q13.31	74.66
D22S1170	AFM268yg1	22	55.26	68.42	9.02	q13.31	61.56
DXS9903	GATA164D10	X	13.5	9.65	0	p22.33	30.69
DXS1223	AFM309YC1	X	16.75	16.52	6.87	p22.31	46.92
DXS7108	AFMA184WF1	X	18.37	19.97	3.45	p22.22	42.25
DXS1224	AFM311VF5	X	21.23	26.38	6.41	p22.22	28.01
DXS1229	AFM337WD5	X	27.59	35.8	9.42	p22.12	21.11
DXS7593	AFMA346ZC1	X	25.97	38.41	2.61	p22.11	43.48
DXS1048	AFM151XG11	X	29.76	43.89	5.48	p21.3	39.41
DXS1061	AFM205YD2	X	30.3	44.23	0.33	p21.3	44.02
DXS1214	AFM283WG9	X	33.54	47.36	3.14	p21.2	43.36
DXS1049	AFM155ZE1	X	36.79	56.63	9.27	p21.1	29.95
DXS1069	AFM240WA9	X	37.33	62.13	5.5	p11.4	38.84
DXS8083	AFMC024XC5	X	46.54	71.55	9.42	p11.3	41.4
DXS1216	AFM287ZG1	X	53.58	85.3	13.76	q13.1	36.86
DXS1217	AFM288YE9	X	59.72	91.84	6.54	q21.31	35.59
DXS6799	GATA29G07	X	64.41	101.75	9.91	q21.33	33.94
DXS8020	AFMA162TC1	X	65.5	104.17	2.41	q22.1	43.62
DXS8110	AFMA086WG5	X	70.91	110.32	6.15	q23	9.61
DXS8055	AFMB291YE5	X	70.91	118.09	7.77	q23	34.19
DXS1001	AFM248WE5	X	75.79	123.76	5.67	q24	45.64
DXS8059	AFMB303XD1	X	78.5	127.15	3.39	q25	37.15
DXS8074	AFMB355XC9	X	82.84	141.05	13.9	q26.3	35.28

DXS984	AFM105XC5	X	85.55	149.84	8.79	q27.1	36.8
DXS8106	AFMA074TF9	X	91.64	158.66	8.82	q27.3	40.84
DXS8073	AFMB354WA9	X	94.22	164.42	5.76	q27.3	38.21
DXS1193	AFM199WC7	X	97.89	174.06	9.64	q28	45.54
DXS8103	AFMA065WD9	X	100.73	183.79	9.74	q28	46.12
DXS1073	AFM276XH9	X	102.35	193.73	9.94	q28	37.66

Appendix 2.

Two point lod score results for Multiform Movement Disorder (MMD) family in chapter four.

Locus	genetic map	Recombination fraction					Theta max	Z max
		0	0.1	0.2	0.3	0.4		
D1S468	5.44	-5.89	-1.35	-0.55	-0.19	-0.04	0.5	0
D1S2870	11.68	-7.79	-0.82	-0.29	-0.1	-0.02	0.5	0
D1S450	16.99	-7.49	-1	-0.4	-0.14	-0.03	0.5	0
D1S2667	20.36	-8.84	-1.63	-0.72	-0.28	-0.07	0.5	0
D1S434	21.8	-8.04	-1.36	-0.61	-0.24	-0.06	0.5	0
D1S2697	29.07	-2.6	-0.89	-0.41	-0.16	-0.04	0.5	0
D1S2620	43.94	-0.07	-0.04	-0.02	-0.01	0	0.5	0
D1S241	54.23	-7.56	-0.86	-0.3	-0.1	-0.02	0.5	0
D1S255	59.98	-6.91	-0.23	0.05	0.07	0.03	0.5	0.08
D1S2861	68.18	-7.09	-0.83	-0.26	-0.06	0	0.5	0
D1S2713	69.43	-0.34	-0.19	-0.09	-0.04	-0.01	0.5	0
D1S2797	70.48	0.29	0.21	0.13	0.06	0.02	0	0.29
D1S197	73.67	0.29	0.21	0.13	0.06	0.02	0	0.29
D1S2652	78.87	-10.44	-0.79	-0.24	-0.05	0	0.5	0
D1S476	80.34	-9.48	-0.14	0.13	0.14	0.06	0.5	0.16
D1S2846	87.95	-0.05	-0.03	-0.02	-0.01	0	0.5	0
D1S2788	89.08	-10.44	-0.79	-0.24	-0.05	0	0.5	0
D1S219	98.97							
D1S2841	105.86	-7.1	-0.79	-0.24	-0.05	0	0.5	0
D1S500	106.4	-0.4	-0.21	-0.11	-0.04	-0.01	0.5	0
D1S430	107.37	-0.42	-0.22	-0.11	-0.05	-0.01	0.5	0
D1S2766	112.11	-7.68	-0.79	-0.24	-0.05	0	0.5	0
D1S2865	114.17	-8.91	-0.79	-0.24	-0.05	0	0.5	0
D1S435	117.67	-0.39	-0.21	-0.1	-0.04	-0.01	0.5	0
D1S206	124.95	-1.27	0.17	0.36	0.29	0.11	0.213	0.36
D1S495	126.35	-1.27	0.17	0.36	0.29	0.11	0.213	0.36
D1S2688	129.08	0.26	0.78	0.67	0.42	0.14	0.101	0.78
D1S189	140.46	0.25	0.17	0.11	0.05	0.01	0	0.25
D1S498	147.7	-1.27	0.17	0.36	0.29	0.11	0.213	0.36
D1S1653	154.54	0.26	0.82	0.73	0.5	0.19	0.112	0.82
D1S484	160.45	0.46	0.38	0.27	0.15	0.04	0	0.46
D1S2628	168.68	-0.35	-0.19	-0.1	-0.04	-0.01	0.5	0
D1S452	176.39	1.69	1.34	0.96	0.55	0.18	0	**1.7**
D1S2818	187.03	-3.1	-0.73	-0.22	-0.04	0	0.5	0
D1S413	198.52	-3.1	-0.73	-0.22	-0.04	0	0.5	0
D1S2717	208.15	-2.02	-0.42	-0.09	0.01	0.01	0.5	0.01
D1S249	210.82	-2.16	-0.5	-0.12	-0.01	0.01	0.5	0
D1S2685	213.64	-1.23	-0.4	-0.18	-0.07	-0.02	0.5	0
D1S245	216.77	-7.83	-1.36	-0.63	-0.25	-0.06	0.5	0
D1S205	217.22	-1.23	-0.4	-0.18	-0.07	-0.02	0.5	0
D1S237	222.18	-8.62	-1.65	-0.73	-0.29	-0.07	0.5	0
D1S2641	230.38	-8.91	-1.65	-0.73	-0.29	-0.07	0.5	0
D1S2833	238.23	-0.42	-0.23	-0.12	-0.05	-0.01	0.5	0
D1S2709	239.64	-8.88	-1.55	-0.7	-0.27	-0.06	0.5	0
D1S2850	251.42	-7.77	-1.36	-0.61	-0.24	-0.06	0.5	0
D1S2785	262.68	-6.67	-1.62	-0.72	-0.28	-0.07	0.5	0

Table 4.6. Two-point lod scores for chromosome 1 in MMD family.

Locus	Genetic map	Recombination fraction					Theta max	Z max
		0	0.1	0.2	0.3	0.4		
D2S323	4.61	-6.78	-1.33	-0.59	-0.23	-0.06	0.5	0
D2S319	7.81	-9.35	-1.66	-0.73	-0.28	-0.07	0.5	0
D2S2211	18.24	-8.56	-1.41	-0.63	-0.25	-0.06	0.5	0
D2S398	28.03	-5.39	-0.83	-0.34	-0.13	-0.03	0.5	0
D2S168	29.73	-8.46	-1.46	-0.65	-0.26	-0.06	0.5	0
D2S305	42.8	-6.19	-1.5	-0.66	-0.26	-0.06	0.5	0
D2S146	52.91	-0.33	-0.18	-0.09	-0.04	-0.01	0.5	0
D2S2328	66.77	-5.79	-1.33	-0.54	-0.19	-0.04	0.5	0
D2S2156	78.42	-7.88	-1.32	-0.53	-0.19	-0.04	0.5	0
D2S1364	82.83	-5.79	-1.33	-0.54	-0.19	-0.04	0.5	0
D2S337	86.05	-7.88	-0.56	-0.14	0	0.02	0.5	0.22
D2S2152	94.62	-7.6	0.39	0.45	0.33	0.13	0.173	0.45
D2S2110	99.58	-7.59	0.16	0.26	0.2	0.08	0.192	0.26
D2S2116	103.5	-8.13	-0.34	-0.12	-0.04	-0.01	0.5	0
D2S388	112.75	-6.4	0.3	0.33	0.23	0.08	0.165	0.34
D2S293	122.55	-7.77	-0.39	-0.05	0.04	0.03	0.5	0.48
D2S363	130.95	-7.65	-0.14	0.14	0.16	0.07	0.5	0.072
D2S347	138.31	-6.74	-1.02	-0.46	-0.18	-0.04	0.5	0
D2S2271	143.62	-7.66	-0.79	-0.24	-0.05	0	0.5	0
D2S2256	149.54	-5.85	-1.08	-0.43	-0.15	-0.03	0.5	0
D2S2324	164.37	-0.36	-0.19	-0.1	-0.04	-0.01	0.5	0
D2S1353	171.63	-9.26	-1.45	-0.66	-0.27	-0.06	0.5	0
D2S2330	174.85	-0.47	-0.24	-0.12	-0.05	-0.01	0.5	0
D2S2188	186.25	-10.44	-1.73	-0.76	-0.3	-0.07	0.5	0
D2S118	195.34	-7.09	0.46	0.42	0.25	0.08	0.128	0.47
D2S2321	209.45	-4.02	-1.22	-0.49	-0.17	-0.03	0.5	0
D2S317	213.81	-4.13	-1.41	-0.63	-0.25	-0.06	0.5	0
D2S2361	216.94	-1.23	-0.4	-0.18	-0.07	-0.02	0.5	0
D2S2382	218.52	-5.27	-1.73	-0.76	-0.3	-0.07	0.5	0
D2S163	226.25	-6.06	-1.34	-0.61	-0.24	-0.06	0.5	0
D2S133	232.12	-5.95	-1.36	-0.62	-0.25	-0.06	0.5	0
D2S427	243.39	-7.22	-1.03	-0.41	-0.15	-0.03	0.5	0
D2S2202	253.23	-8.24	-1.65	-0.73	-0.29	-0.07	0.5	0
D2S338	253.84	-1.23	-0.4	-0.18	-0.07	-0.02	0.5	0
D2S395	262.12	-6.19	-1.59	-0.7	-0.27	-0.06	0.5	0
D2S140	265.28	-1.72	-0.55	-0.26	-0.1	-0.02	0.5	0

Table 4.7. Two-point lod scores for chromosome 2 in MMD family.

Order	Genetic map	Recombination fraction					Theta max	Zmax
		0	0.1	0.2	0.3	0.4		
D3S4559	1.08	-1.23	-0.4	-0.18	-0.07	-0.02	0.5	0
D3S1270	3.28	-7.67	-1.69	-0.75	-0.29	-0.07	0.5	0
D3S3630	6.22	-0.47	-0.24	-0.12	-0.05	-0.01	0.5	0
D3S3050	10.68	-7.1	-1.69	-0.75	-0.29	-0.07	0.5	0
D3S2397	16.21	-7.1	-1.69	-0.75	-0.29	-0.07	0.5	0
D3S1515	19.89	-7.04	-1.45	-0.64	-0.25	-0.06	0.5	0
D3S3591	23.9	-0.35	-0.19	-0.1	-0.04	-0.01	0.5	0
D3S1263	30.71	-8.78	-1.67	-0.73	-0.29	-0.07	0.5	0
D3S2385	33.55	-0.19	-0.13	-0.08	-0.04	-0.01	0.5	0
D3S1567	48.03	-0.05	-0.03	-0.02	-0.01	0	0.5	0
D3S2335	50.4	-7.64	-1.42	-0.63	-0.25	-0.06	0.5	0
D3S1266	52.22	-7.93	-1.69	-0.75	-0.29	-0.07	0.5	0
D3S3547	56.45	-0.95	-0.2	-0.05	-0.01	0	0.5	0
D3S3521	66.4	-7.28	-1.55	-0.69	-0.27	-0.06	0.5	0
D3S3672	76.71	-8.28	-1.68	-0.74	-0.29	-0.07	0.5	0
D3S1600	89.02	-6.95	-1.43	-0.65	-0.26	-0.06	0.5	0
D3S3551	99.23	-8.58	-1.08	-0.44	-0.16	-0.03	0.5	0
D3S3653	106.92	0.29	0.21	0.13	0.06	0.02	0	0.29
D3S1271	115.24	-7.84	-1.44	-0.64	-0.25	-0.06	0.5	0
D3S3045	120.29	-7.5	-1.46	-0.59	-0.21	-0.04	0.5	0
D3S1558	129.21	0.29	0.21	0.13	0.06	0.02	0	0.29
D3S1589	136.43	-0.36	-0.19	-0.1	-0.04	-0.01	0.5	0
D3S3637	146.26	-6.16	-0.88	-0.32	-0.12	-0.03	0.5	0
D3S1569	154.11	-3.1	-0.73	-0.2	-0.02	0.02	0.5	0.27
D3S1308	160.67	-3.1	-0.73	-0.2	-0.02	0.02	0.5	0.27
D3S1607	168.31	0.29	0.21	0.13	0.06	0.02	0	0.29
D3S1564	175.21	-0.32	-0.18	-0.09	-0.04	-0.01	0.5	0
D3S3725	178.27	-1.57	-0.09	0.16	0.16	0.07	0.5	0.77
D3S3041	185.92	-1.57	-0.09	0.15	0.15	0.06	0.5	0.17
D3S3592	195.05	-1.27	0.17	0.36	0.29	0.11	0.213	0.36
D3S1580	206.43	-1.27	0.17	0.36	0.29	0.11	0.213	0.36
D3S3663	213.15	-9.23	-0.49	-0.11	0	0.01	0.5	0
D3S240	219.97	-0.02	-0.01	-0.01	0	0	0.5	0
D3S1265	222.34	-8.91	-1.69	-0.75	-0.29	-0.07	0.5	0
D3S1311	225.05	-0.52	-0.26	-0.13	-0.05	-0.01	0.5	0

Table 4.8. Two-point lod scores for chromosome 3 in MMD family.

Order	Genetic map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D4S412	4.57	-9.21	-1.45	-0.65	-0.26	-0.06	0.5	0
D4S2935	14.07	-0.21	-0.12	-0.07	-0.03	-0.01	0.5	0
D4S394	15.28	-7.88	-1.66	-0.73	-0.28	-0.07	0.5	0
D4S2928	23.84	-0.41	-0.21	-0.11	-0.04	-0.01	0.5	0
D4S3341	32.25	-2.21	-0.52	-0.13	-0.01	0.01	0.5	0.02
D4S404	41	-0.68	-0.31	-0.15	-0.06	-0.01	0.5	0
D4S3022	43.89	-0.66	-0.3	-0.14	-0.06	-0.01	0.5	0
D4S391	48.95	-2.19	-0.51	-0.13	-0.01	0.01	0.5	0
D4S2912	48.97	-2	-0.42	-0.08	0.01	0.01	0.5	0
D4S1587	55.94	-2.02	-0.43	-0.09	0.01	0.01	0.5	0
D4S405	62.1	-2.12	-0.48	-0.12	0	0.01	0.5	0
D4S428	71.03	-0.34	-0.18	-0.09	-0.04	-0.01	0.5	0
D4S2389	84.72	-1.3	-0.47	-0.23	-0.1	-0.02	0.5	0
D4S1553	92.45	-7.08	-1.28	-0.57	-0.23	-0.05	0.5	0
D4S423	102.7	-8.24	-1.65	-0.73	-0.29	-0.07	0.5	0
D4S1572	110.14	-1.23	-0.4	-0.18	-0.07	-0.02	0.5	0
D4S406	117.48	-8.62	-1.69	-0.75	-0.29	-0.07	0.5	0
D4S1615	129.75	-7.93	-1.69	-0.75	-0.29	-0.07	0.5	0
D4S1575	134.67	-7.85	-1.65	-0.72	-0.28	-0.06	0.5	0
D4S424	142.05	-8.95	-1.42	-0.62	-0.24	-0.05	0.5	0
D4S1586	144.42	-0.45	-0.23	-0.11	-0.05	-0.01	0.5	0
D4S2962	148.06	-10.44	-1.69	-0.75	-0.29	-0.07	0.5	0
D4S3046	159.33	-9.07	-1.54	-0.69	-0.27	-0.06	0.5	0
D4S1539	173.75	-6.63	-0.38	-0.06	0.02	0.02	0.5	0.07
D4S1607	184.14	-7.96	-1.47	-0.66	-0.26	-0.06	0.5	0
D4S3041	189.52	-7.88	-1.61	-0.71	-0.28	-0.06	0.5	0
D4S408	193.5	-7.89	-1.49	-0.67	-0.26	-0.06	0.5	0
D4S1540	195.66	-8.24	-1.65	-0.73	-0.29	-0.07	0.5	0
D4S2930	209.27	-0.65	-0.3	-0.14	-0.06	-0.01	0.5	0

Table 4.9. Two-point lod scores for chromosome 4 in MMD family.

Locus	Genetic Map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D5S1981	1.21	-0.51	-0.25	-0.12	-0.05	-0.01	0.5	0
D5S2088	12.05	-8.16	-1.65	-0.73	-0.29	-0.07	0.5	0
D5S2054	16.64	-7.22	-0.59	-0.28	-0.12	-0.03	0.5	0
D5S635	18.11	-8.24	-1.65	-0.73	-0.29	-0.07	0.5	0
D5S1486	27.25	-0.4	-0.21	-0.11	-0.04	-0.01	0.5	0
D5S2081	31.94	-6.65	-0.75	-0.31	-0.12	-0.03	0.5	0
D5S2031	42.23	-6.59	-0.48	-0.11	0	0.01	0.5	0.01
D5S674	56.71	-6.18	-0.2	0.05	0.07	0.03	0.5	0.04
D5S2021	61.33	0.46	0.38	0.28	0.16	0.06	0	0.46
D5S628	71.14	-6.74	-0.74	-0.22	-0.04	0	0.5	0.014
D5S2072	80.82	-6.65	-0.99	-0.44	-0.18	-0.04	0.5	0
D5S424	93.23	-8.53	-0.49	-0.12	0	0.01	0.5	0
D5S1397	98.14	-6.71	-0.53	-0.14	-0.01	0	0.5	0
D5S1725	106.88	0.29	0.21	0.13	0.06	0.02	0	0.29
D5S409	114.89	0.26	0.78	0.67	0.42	0.14	0.1	0.78
D5S2027	120.95	0.55	1.02	0.84	0.52	0.17	0.088	**1.03**
D5S2065	122.49	0.55	1.02	0.84	0.52	0.17	0.088	**1.03**
D5S490	134.52	-0.33	-0.18	-0.09	-0.04	-0.01	0.5	0
D5S2017	145.35	-7.91	-0.53	-0.14	-0.01	0	0.5	0
D5S2090	153.72	-7.58	0.4	0.45	0.33	0.13	0.17	0.46
D5S410	162.45	-5.27	0.97	0.82	0.51	0.17	0.101	0.97
D5S2066	171.45	-6.3	-0.15	0	0.03	0.02	0.5	0.25
D5S400	181.34	-8.62	0.11	0.33	0.28	0.12	0.223	0.34
D5S1973	184.16	-8.62	0.11	0.33	0.28	0.12	0.223	0.34
D5S2069	194.4	0.36	0.26	0.16	0.08	0.02	0	0.36
D5S469	203.46	-6.48	-0.41	-0.07	0.01	0.01	0.5	0.01
D5S2006	211.06	-0.11	-0.06	-0.04	-0.02	0	0.5	0

Table 4.10. Two-point lod scores for chromosome 5 in MMD family.

Locus	Genetic Map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D6S244	5.66	-6.57	-0.47	-0.1	0	0.01	0.5	0.02
D6S1617	12.18	-8.84	-0.79	-0.24	-0.05	0	0.5	0
D6S309	21.6	-7.38	-1.62	-0.72	-0.28	-0.06	0.5	0
D6S1279	30.35	-6.19	-1.64	-0.73	-0.29	-0.07	0.5	0
D6S422	44.68	-9.4	-1.56	-0.7	-0.28	-0.07	0.5	0
D6S1660	48.37	-0.41	-0.21	-0.11	-0.04	-0.01	0.5	0
D6S273	53.25	0.29	0.21	0.13	0.06	0.02	0	0.29
D6S1575	63.69	-8.2	-1.41	-0.6	-0.23	-0.05	0.5	0
D6S459	72.59	-5.69	-1.46	-0.66	-0.26	-0.06	0.5	0
D6S452	74.58	-0.27	0.06	0.13	0.1	0.03	0.207	0.13
D6S1557	87.06	-5.54	-0.82	-0.36	-0.14	-0.03	0.5	0
D6S460	93.4	-2.17	-0.58	-0.22	-0.07	-0.01	0.5	0
D6S458	102.77	-0.04	-0.02	-0.01	-0.01	0	0.5	0
D6S268	115.03	0.32	0.23	0.14	0.07	0.02	0	0.32
D6S1712	123.79	-6.15	-1.32	-0.54	-0.19	-0.04	0.5	0
D6S1656	134.76	-5.9	-0.99	-0.39	-0.14	-0.03	0.5	0
D6S270	138.78	-8.21	-0.82	-0.3	-0.11	-0.03	0.5	0
D6S1009	142.71	-7.52	-1.31	-0.6	-0.24	-0.06	0.5	0
D6S1637	154.79	-7.03	0.07	0.24	0.17	0.05	0.204	0.25
D6S441	164.5	-7.66	-0.79	-0.24	-0.05	0	0.5	0
D6S419	170.85	-7.41	-0.4	-0.06	0.03	0.02	0.5	0.25
D6S1581	174.43	-5.22	-0.23	-0.06	-0.01	0	0.5	0
D6S305	176.93	-7.95	-0.74	-0.29	-0.09	-0.02	0.5	0
D6S1599	179.13	-6.68	-0.52	-0.13	-0.01	0	0.5	0
D6S1277	180.81	-1.23	-0.4	-0.18	-0.07	-0.02	0.5	0
D6S297	187.65	-8.62	-0.79	-0.24	-0.05	0	0.5	0
D6S503	190.17	-8.77	-0.79	-0.24	-0.05	0	0.5	0
D6S446	194.22	-6.7	-0.83	-0.34	-0.12	-0.03	0.5	0

Table 4.11. Two-point lod scores for chromosome 6 in MMD family.

Locus	Genetic Map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D7S2474	3.71	-2.13	-0.48	-0.12	0	0.01	0.5	0.02
D7S2201	12.21	-0.99	-0.22	-0.07	-0.02	0	0.5	0
D7S513	23.24	-9.32	-1.55	-0.7	-0.27	-0.06	0.5	0
D7S664	26.72	-9.25	-0.92	-0.37	-0.13	-0.03	0.5	0
D7S1795	38.61	-7.76	-0.43	-0.12	-0.03	-0.01	0.5	0
D7S2463	40.51	-8.82	-1.4	-0.63	-0.25	-0.06	0.5	0
D7S516	45.36	-0.14	-0.08	-0.04	-0.02	0	0.5	0
D7S2250	55.9	-8.08	-1.35	-0.62	-0.25	-0.06	0.5	0
D7S2541	63.45	-9.63	-1.64	-0.73	-0.28	-0.07	0.5	0
D7S691	64.93	-9.45	-1.58	-0.71	-0.28	-0.07	0.5	0
D7S519	70.24	-9.29	-1.62	-0.71	-0.28	-0.06	0.5	0
D7S2429	80.22	-8.36	-1.47	-0.66	-0.26	-0.06	0.5	0
D7S502	80.93	-7.56	-1.63	-0.72	-0.28	-0.07	0.5	0
D7S2500	83.71	-9.07	-1.03	-0.41	-0.15	-0.03	0.5	0
D7S672	86.43	-9.13	-1.58	-0.7	-0.28	-0.07	0.5	0
D7S2443	94.47	0.29	0.21	0.13	0.06	0.02	0	0.29
D7S2485	100.34	-5.74	-0.75	-0.28	-0.09	-0.01	0.5	0
D7S657	105.59	-4.93	-0.87	-0.27	-0.06	-0.01	0.5	0
D7S554	109.67	-5.95	-1.39	-0.61	-0.24	-0.06	0.5	0
D7S2509	114.48	-7.58	-0.91	-0.32	-0.1	-0.02	0.5	0
D7S2459	120.05	-4.46	0.71	0.62	0.37	0.11	0.113	0.72
D7S2418	122.15	-5.27	0.97	0.82	0.51	0.17	0.101	**0.97**
D7S530	133.2	-6.08	0.39	0.44	0.32	0.12	0.173	0.45
D7S2533	141.74	-0.31	-0.17	-0.09	-0.04	-0.01	0.5	0
D7S684	149.78	-6.15	0.11	0.33	0.28	0.11	0.222	0.34
D7S1824	152.03	-4.92	0.51	0.51	0.35	0.13	0.143	0.54
D7S661	155.04	-7.75	0.38	0.44	0.32	0.12	0.176	0.45
D7S1805	165.55	-5.3	0.31	0.41	0.31	0.12	0.187	0.41
D7S2447	179.62	-0.37	-0.2	-0.1	-0.04	-0.01	0.5	0
D7S2423	192.24	-9.24	-0.79	-0.24	-0.05	0	0.5	0

Table 4.12. Two-point lod scores for chromosome 7 in MMD family.

		Recombination fraction					Theta Max	Zmax
Locus	Genetic map	0	0.1	0.2	0.3	0.4		
D8S264	3.54	-9.45	-1.63	-0.72	-0.28	-0.07	0.5	0
D8S1819	16.91	-8.91	-1.46	-0.59	-0.21	-0.04	0.5	0
D8S520	21.48	-7.07	-1.12	-0.49	-0.19	-0.05	0.5	0
D8S516	22.5	-6.23	-0.21	-0.04	0	0.01	0.5	0.09
D8S258	36.13	0.29	0.21	0.13	0.06	0.02	0	0.29
D8S1734	42.1	-5.27	-0.86	-0.3	-0.1	-0.02	0.5	0
D8S1820	49.17	-6.77	-1.26	-0.51	-0.18	-0.03	0.5	0
D8S505	55.98	-7.28	-0.32	-0.02	0.05	0.03	0.5	0.25
D8S1737	69.34	-6.55	-0.47	-0.1	0.02	0.02	0.5	0.25
D8S1763	73.88	-3.57	-0.45	-0.09	0.02	0.02	0.5	0.23
D8S512	78.06	0.59	0.45	0.31	0.16	0.05	0	0.59
D8S279	86.52	0.34	0.24	0.15	0.07	0.02	0	0.34
D8S1707	97.46	0.24	0.17	0.1	0.05	0.01	0	0.24
D8S1778	106.89	0.26	0.18	0.11	0.05	0.01	0	0.26
D8S1762	108.03	-10.5	-1.67	-0.74	-0.29	-0.07	0.5	0
D8S1470	118.7	-8.26	-1.29	-0.57	-0.22	-0.05	0.5	0
D8S1799	131.69	-9.3	-1.46	-0.66	-0.27	-0.06	0.5	0
D8S1793	135.13	-8.81	-1.52	-0.69	-0.27	-0.07	0.5	0
D8S263	141.47	-6.93	-1.41	-0.63	-0.25	-0.06	0.5	0
D8S1746	150.11	-0.56	-0.27	-0.13	-0.05	-0.01	0.5	0
D8S1783	152.34	-5.43	-0.56	-0.15	-0.02	0	0.5	0
D8S274	153.61	-0.13	-0.08	-0.04	-0.02	0	0.5	0
D8S1743	160.11	-4.25	-0.56	-0.15	-0.02	0	0.5	0

Table 4.13 Two-point lod scores for chromosome 8 in MMD family.

		Recombination fraction						
Locus	Genetic Map	0	0.1	0.2	0.3	0.4	Theta Max	Zmax
D9S1779	0	-0.2	0.04	0.08	0.04	0.01	0.178	0.08
D9S1686	11.93	0.05	0.29	0.31	0.21	0.07	0.159	0.32
D9S286	18.53	-0.19	0.08	0.15	0.11	0.04	0.201	0.15
D9S168	24.28	0.07	0.04	0.02	0.01	0	0	0.07
D9S1808	28.24	-1.14	-0.65	-0.33	-0.14	-0.03	0.5	0
D9S235	32.65	-0.71	-0.39	-0.2	-0.08	-0.02	0.5	0
D9S171	47.16	-0.17	-0.02	0.01	0.02	0.01	0.5	0.02
D9S2149	55.3	-0.14	-0.09	-0.05	-0.02	-0.01	0.5	0
D9S1777	67.34	-0.64	-0.34	-0.17	-0.07	-0.02	0.5	0
D9S1876	70.42	0.05	0.02	0.01	0	0	0	0.05
D9S175	72.8	-0.59	-0.21	-0.04	0.02	0.01	0.5	0.24
D9S1843	79.87	0.5	0.53	0.43	0.26	0.09	0.064	0.54
D9S307	89.23	-0.16	0.1	0.15	0.11	0.04	0.202	0.15
D9S283	95	0.44	0.38	0.28	0.16	0.05	0	0.44
D9S287	101.54	-0.12	-0.07	-0.03	-0.01	0	0.5	0
D9S261	114.65	0.37	0.45	0.38	0.24	0.08	0.5	0.45
D9S1675	117.44	0.22	0.2	0.15	0.08	0.02	0.022	0.22
D9S1828	117.67	-0.07	0.25	0.29	0.21	0.07	0.173	0.3
D9S1824	123.56	-0.61	-0.24	-0.07	-0.01	0	0.5	0
D9S934	127.72	-0.67	-0.26	-0.08	-0.01	0	0.5	0
D9S1682	132.09	-0.54	-0.2	-0.05	0	0	0.5	0
D9S1825	136.35	-0.83	-0.34	-0.12	-0.03	0	0.5	0
D9S1830	149.36	-0.44	-0.15	-0.05	-0.02	-0.01	0.5	0
D9S1793	150.42	0.23	0.16	0.1	0.04	0.01	0	0.23
D9S1838	164.1	-0.4	-0.1	0.02	0.04	0.02	0.5	0.52

Table 4.14. Two-point lod scores for chromosome 9 in MMD family.

Locus	Genetic map	Recombination fraction					Theta max	Zmax
		0	0.1	0.2	0.3	0.4		
D10S249	1.19	-5.85	-0.73	-0.27	-0.08	-0.01	0.5	0
D10S602	5.71	-4.93	-0.91	-0.32	-0.11	-0.02	0.5	0
D10S1218	11.3	-3.15	-0.95	-0.43	-0.17	-0.04	0.5	0
D10S189	20.56	-5.55	-0.62	-0.17	-0.03	0	0.5	0
D10S1649	26.55	-8.91	0.01	0.22	0.16	0.04	0.21	0.22
D10S1430	32.11	-7.43	-0.87	-0.31	-0.1	-0.02	0.5	0
D10S570	32.73	0.32	0.23	0.14	0.07	0.02	0	0.32
D10S191	37.42	0.35	0.25	0.16	0.08	0.02	0	0.35
D10S1653	38.92	-6.95	-1.25	-0.5	-0.17	-0.03	0.5	0
D10S1661	41.4	-7.41	-1.48	-0.67	-0.26	-0.06	0.5	0
D10S600	54.33	-7.39	-1.47	-0.67	-0.26	-0.06	0.5	0
D10S1746	65.42	-6.96	-1.41	-0.63	-0.25	-0.06	0.5	0
D10S196	71.95	-7.19	-1.52	-0.68	-0.27	-0.06	0.5	0
D10S609	82.33	-0.06	-0.04	-0.02	-0.01	0	0.5	0
D10S537	91.46	-6.69	-0.38	-0.14	-0.05	-0.01	0.5	0
D10S580	97.94	-0.19	-0.11	-0.06	-0.03	-0.01	0.5	0
D10S1677	100.72	-0.25	-0.14	-0.07	-0.03	-0.01	0.5	0
D10S1753	113.42	-6.14	-0.86	-0.3	-0.1	-0.02	0.5	0
D10S185	116.07	-4.92	-0.23	0.03	0.05	0.01	0.5	0
D10S1267	125.15	-6.98	-0.87	-0.31	-0.11	-0.03	0.5	0
D10S597	129.57	-0.18	-0.11	-0.06	-0.02	-0.01	0.5	0
D10S1693	141.5	-6.15	-0.87	-0.31	-0.11	-0.03	0.5	0
D10S1656	154.8	-7.58	-1.15	-0.45	-0.15	-0.03	0.5	0
D10S575	160.19	-8.87	-1.57	-0.69	-0.27	-0.06	0.5	0
D10S1676	164.6	0.29	0.21	0.13	0.06	0.02	0	0.29
D10S1651	176.24	-8.62	-0.83	-0.3	-0.11	-0.03	0.5	0
D10S212	181.66	-8.62	-1.46	-0.59	-0.21	-0.04	0.5	0

Table 4.15. Two-point lod scores for chromosome 10 in MMD family

Locus	Genetic map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D11S4046	0	-0.48	-0.24	-0.12	-0.05	-0.01	0.5	0
D11S1760	8.79	-7.76	-1.03	-0.42	-0.15	-0.03	0.5	0
D11S1331	12.38	-6.09	-0.62	-0.22	-0.07	-0.01	0.5	0
D11S4170	23.4	-7.5	-1.46	-0.59	-0.21	-0.04	0.5	0
D11S902	26.23	-8.68	-1.6	-0.72	-0.28	-0.07	0.5	0
D11S928	34.9	-7.09	-1.46	-0.59	-0.21	-0.04	0.5	0
D11S4080	43.29	-6.38	-0.6	-0.2	-0.05	0	0.5	0
D11S1776	49.46	-5.27	0.07	0.25	0.19	0.06	0.207	0.25
D11S1360	58.18	0.32	0.23	0.14	0.07	0.02	0	0.32
D11S4191	66.18	-7.09	0.07	0.24	0.17	0.05	0.203	0.24
D11S4087	73.81	-4.12	0.06	0.23	0.16	0.05	0.204	0.23
D11S937	85.42	-5.27	0.97	0.82	0.51	0.17	0.101	**0.97**
D11S1780	94.12	-5.56	-0.3	-0.1	-0.03	-0.01	0.5	0
D11S1886	107.71	-7.01	0.97	0.82	0.51	0.17	0.101	**0.97**
D11S4206	110.81	0.34	0.25	0.15	0.08	0.02	0	0.34
D11S908	118.66	-0.13	-0.08	-0.04	-0.02	0	0.5	0
D11S4089	126.83	-6	0.43	0.47	0.33	0.12	0.164	0.48
D11S4151	135.73	-4.33	0.97	0.82	0.51	0.17	0.101	**0.97**
D11S2367	148.23	-4.17	0.43	0.47	0.33	0.12	0.163	0.48
D11S969	154.38	-5.74	-0.05	0.16	0.15	0.06	0.5	0.16

Table 4.16. Two-point lod scores for chromosome 11 in MMD family.

Locus	Genetic map	Recombination fraction					Theta Max	Zmax
		0.0	0.1	0.2	0.3	0.4		
D12S352	0.68	-0.35	-0.19	-0.1	-0.04	-0.01	0.5	0
D12S1656	4.23	0.29	0.21	0.13	0.06	0.02	0	0.29
D12S372	8.63	-6.29	-1.08	-0.44	-0.16	-0.04	0.5	0
D12S374	16.84	-9.17	-1.43	-0.64	-0.25	-0.06	0.5	0
D12S336	24.76	-7.85	-1.39	-0.63	-0.25	-0.06	0.5	0
D12S364	31.97	-0.49	-0.25	-0.12	-0.05	-0.01	0.5	0
D12S1591	44.81	-7.44	-1.39	-0.63	-0.25	-0.06	0.5	0
D12S1337	51.8	-0.06	-0.04	-0.02	-0.01	0	0.5	0
D12S1704	52.93	-0.11	-0.07	-0.04	-0.02	0	0.5	0
D12S85	61.77	-1.98	-0.41	-0.08	0.01	0.01	0.5	0
D12S368	67.92	-1.27	0.12	0.27	0.18	0.05	0.193	0.27
D12S83	75.48	-1.27	0.12	0.27	0.18	0.05	0.195	0.27
D12S1294	83.69	0.12	0.06	0.03	0.01	0	0	0.12
D12S326	92.97	-0.19	0.48	0.49	0.34	0.12	0.149	0.51
D12S351	103.1	-6.14	0.47	0.44	0.28	0.09	0.13	0.49
D12S95	104.42	-8	-0.61	-0.17	-0.02	0	0.5	0
D12S1346	106.03	-7.62	-0.14	0.13	0.15	0.06	0.5	0.22
D12S2081	110.59	-7.65	-0.42	-0.08	0.01	0.01	0.5	0
D12S1636	120.17	-7.35	-0.48	-0.11	0	0.01	0.5	0.04
D12S1613	124.08	-0.23	-0.13	-0.07	-0.03	-0.01	0.5	0
D12S1583	127.22	-8.08	-0.61	-0.17	-0.03	0	0.5	0
D12S354	133.66	-0.15	-0.09	-0.05	-0.02	0	0.5	0
D12S369	134.33	-9.42	-0.95	-0.39	-0.14	-0.03	0.5	0
D12S366	140.32	-1.57	-0.09	0.15	0.15	0.06	0.5	0.24
D12S2073	144.14	-7.89	0.35	0.43	0.31	0.12	0.177	0.43
D12S2078	155.99	-0.16	0.49	0.49	0.34	0.13	0.146	0.52
D12S1609	160.72	-1.27	0.17	0.36	0.29	0.11	0.213	0.36
D12S1723	173.29	-8.78	-1.62	-0.71	-0.28	-0.06	0.5	0

Table 4.17. Two-point lod scores for chromosome 12 in MMD family.

Locus	Genetic Map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D13S175	3.24	-7.62	-1.63	-0.72	-0.28	-0.07	0.5	0
D13S232	8.98	-6.26	-1.6	-0.71	-0.28	-0.07	0.5	0
D13S1243	11.57	-1.57	-0.09	0.15	0.15	0.06	0.5	0.17
D13S221	16.88	-8.91	-0.79	-0.24	-0.05	0	0.5	0
D13S217	22.89	-8.14	-1.21	-0.52	-0.2	-0.05	0.5	0
D13S289	28.14	-9.29	-1.56	-0.69	-0.27	-0.06	0.5	0
D13S171	31.99	-9.32	-1.46	-0.65	-0.26	-0.06	0.5	0
D13S219	36.49	-8.64	-0.96	-0.43	-0.17	-0.04	0.5	0
D13S218	40.44	-1.48	-0.54	-0.25	-0.1	-0.02	0.5	0
D13S272	54.31	-8.78	-1.62	-0.71	-0.28	-0.06	0.5	0
D13S279	65.83	-7.26	-1.38	-0.63	-0.25	-0.06	0.5	0
D13S271	80.33	-9.31	-1.47	-0.67	-0.27	-0.06	0.5	0
D13S1241	91.34	-0.45	-0.23	-0.12	-0.05	-0.01	0.5	0
D13S1256	100.66	-8.77	-1.46	-0.67	-0.27	-0.06	0.5	0
D13S1809	107.62	-4.86	-0.64	-0.24	-0.08	-0.02	0.5	0
D13S1315	119	-1.26	-0.41	-0.18	-0.07	-0.02	0.5	0
D13S293	131.92	-2.47	-0.39	-0.12	-0.02	0	0.5	0.06

Table 4.18. Two-point lod scores for chromosome 13 in MMD family.

Locus	Genetic map	Recombination fraction					Theta Max	Z max
		0	0.1	0.2	0.3	0.4		
D14S261	9.53	-0.15	-0.09	-0.05	-0.02	-0.01	0.5	0
D14S1043	9.83	-0.04	-0.03	-0.01	-0.01	0	0.5	0
D14S1280	23.04	-9.16	-1.4	-0.64	-0.26	-0.06	0.5	0
D14S262	25.73	-1.23	-0.4	-0.18	-0.07	-0.02	0.5	0
D14S1071	29.19	-8.91	-1.69	-0.75	-0.29	-0.07	0.5	0
D14S741	35	-9.21	-0.78	-0.32	-0.12	-0.03	0.5	0
D14S70	37.63	-9.25	-0.8	-0.33	-0.12	-0.03	0.5	0
D14S75	43.2	-9.19	-1.42	-0.65	-0.26	-0.06	0.5	0
D14S976	46.8	-1.57	-0.08	0.16	0.16	0.06	0.5	0.18
D14S978	51.44	-1.24	-0.36	-0.13	-0.04	-0.01	0.5	0
D14S276	57.33	-3.1	-0.73	-0.22	-0.04	0	0.5	0
D14S1011	70.01	-0.47	-0.24	-0.12	-0.05	-0.01	0.5	0
D14S258	70.4	-1.59	-0.14	0.08	0.08	0.03	0.5	0.03
D14S1433	74.94	-0.01	0.59	0.55	0.37	0.13	0.126	0.6
D14S616	85.45	-3.45	0.97	0.82	0.51	0.17	0.101	0.97
D14S67	88.68	0.29	0.21	0.13	0.06	0.02	0	0.29
D14S1044	92.86	-0.12	-0.07	-0.04	-0.02	0	0.5	0
D14S81	99.25	-7.09	0.07	0.24	0.17	0.05	0.203	0.24
D14S1054	102.52	1.18	0.91	0.62	0.33	0.1	0.001	**1.18**
D14S987	106.22	-7.49	-0.78	-0.23	-0.03	0.01	0.5	0.24
D14S1019	109.31	0.33	0.24	0.15	0.07	0.02	0.001	0.33
D14S1426	117.81	-8.41	-0.72	-0.21	-0.04	0	0.5	0
D14S542	126.06	0.29	0.21	0.13	0.06	0.02	0.001	0.29

Table 4.19. Two-point lod scores for chromosome 14 in MMD family.

Locus	Genetic map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D15S128	6.06	-6.46	-1.46	-0.59	-0.2	-0.04	0.5	0
D15S975	13.46	0.34	0.25	0.15	0.08	0.02	0	0.34
D15S1019	21.44	-0.09	-0.06	-0.03	-0.01	0	0.5	0
D15S165	23.51	-5.27	0.07	0.24	0.17	0.05	0.203	0.24
D15S231	27.43	0.29	0.21	0.13	0.06	0.02	0	0.29
D15S118	35.23	0.34	0.25	0.16	0.08	0.02	0	0.34
D15S1012	39.39	-7.02	-1.26	-0.56	-0.22	-0.05	0.5	0
D15S146	43.14	-0.42	-0.22	-0.11	-0.05	-0.01	0.5	0
D15S1016	52.71	0.3	0.21	0.13	0.06	0.02	0	0.29
D15S117	58.09	-7.39	-0.87	-0.31	-0.1	-0.02	0.5	0
D15S988	70.39	-7.85	-1.25	-0.51	-0.18	-0.03	0.5	0
D15S114	82.51	-0.33	-0.18	-0.09	-0.04	-0.01	0.5	0
D15S1005	86.82	-8.51	-1.48	-0.66	-0.26	-0.06	0.5	0
D15S158	100.89	0.59	0.45	0.31	0.16	0.05	0	0.59
D15S1004	110.96	0.33	0.24	0.15	0.07	0.02	0	0.33
D15S130	111.55	0.59	0.45	0.31	0.16	0.05	0	0.59
D15S816	114.36	-3.45	0.01	0.22	0.17	0.05	0.212	0.22
D15S157	116.14	-4.63	0.01	0.22	0.17	0.05	0.212	0.22
D15S212	124.08	0.59	0.45	0.31	0.16	0.05	0	0.59
D15S120	130.4	-0.15	-0.09	-0.05	-0.02	-0.01	0.5	0

Table 4.20. Two-point lod scores for chromosome 15 in MMD family.

Locus	Genetic map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D16S521	1.15	-6.94	-1.41	-0.63	-0.25	-0.06	0.5	0
D16S3065	10.64	-1.57	-0.6	-0.29	-0.12	-0.03	0.5	0
D16S423	14.97	-8.91	-0.79	-0.24	-0.05	0	0.5	0
D16S418	20.61	-7.65	-0.14	0.13	0.14	0.06	0.5	0.16
D16S3062	32.07	-7.85	-0.49	-0.12	-0.01	0.01	0.5	0
D16S500	33.31	-1.23	-0.4	-0.18	-0.07	-0.02	0.5	0
D16S410	41.83	0.35	0.25	0.16	0.08	0.02	0	0.34
D16S3068	51.27	-7.09	0.07	0.25	0.19	0.06	0.207	0.25
D16S3080	61.1	-8.04	-0.58	-0.15	0	0.02	0.5	0.24
D16S3034	68.27	-7.83	-0.43	-0.12	-0.03	-0.01	0.5	0
D16S3057	76.36	-0.11	-0.06	-0.04	-0.02	0	0.5	0
D16S514	83.1	-7.82	0.07	0.25	0.17	0.05	0.204	0.25
D16S503	84.26	0.28	0.2	0.12	0.06	0.02	0	0.28
D16S515	94.69	-1.47	-0.13	0.07	0.06	0.01	0.5	0
D16S516	101.24	-1.27	0.17	0.36	0.29	0.11	0.213	0.36
D16S505	108.07	0.55	1.07	0.93	0.64	0.26	0.098	1.07
D16S763	120.63	1.23	1.07	0.84	0.54	0.2	0	**1.23**
D16S2621	131.46	-6.16	0.11	0.33	0.28	0.11	0.222	0.34

Table 4.21. Two-point lod scores for chromosome 16 in MMD family.

251

Locus	Genetic map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D17S849	0.63	-8.78	-1.26	-0.54	-0.21	-0.05	0.5	0
D17S831	7.47	-7.52	-1.19	-0.52	-0.2	-0.05	0.5	0
D17S1832	16.79	0.29	0.21	0.13	0.06	0.02	0	0.29
D17S804	28.96	-6.15	0.11	0.33	0.28	0.12	0.223	0.34
D17S799	37.94	-8.44	0.11	0.33	0.28	0.12	0.223	0.34
D17S922	41.55	-8.82	-0.84	-0.26	-0.06	0	0.5	0
D17S839	43.77	-6.74	-1.04	-0.39	-0.11	-0.01	0.5	0.05
D17S1824	53.22	-7.3	-0.65	-0.22	-0.05	0	0.5	0.06
D17S2194	58.94	0.59	0.45	0.31	0.16	0.05	0	0.59
D17S1867	66.96	-7.67	-0.72	-0.26	-0.07	-0.01	0.5	0.06
D17S1299	71.13	-7.09	-0.83	-0.25	-0.04	0.01	0.5	0.22
D17S1868	76.3	-5.56	-0.62	-0.16	-0.01	0.02	0.5	0.22
D17S1795	77.3	-7.09	-0.83	-0.25	-0.04	0.01	0.5	0.22
D17S957	89.64	0.37	0.27	0.17	0.08	0.02	0	0.37
D17S794	94.36	-5.47	-0.58	-0.15	0	0.02	0.5	0.22
D17S944	95.11	0.27	0.19	0.12	0.06	0.02	0	0.27
D17S940	102.6	-0.04	-0.03	-0.02	-0.01	0	0.5	0
D17S1351	108.83	0.3	0.22	0.13	0.06	0.02	0	0.3
D17S1603	117.82	0.35	0.25	0.16	0.08	0.02	0	0.35
D17S1847	126.71	-3.8	-1.27	-0.58	-0.23	-0.05	0.5	0
D17S836	128.03	-3.1	-0.77	-0.27	-0.09	-0.02	0.5	0
D17S784	132.74	-4.16	-0.67	-0.23	-0.06	0	0.5	0.05

Table 4.22. Two-point lod scores for chromosome 17 in MMD family.

Locus	Genetic map	Recombination fraction					Theta max	Zmax
		0	0.1	0.2	0.3	0.4		
D18S1105	3.97	-0.37	-0.2	-0.1	-0.04	-0.01	0.5	0
D18S63	9.84	-0.46	-0.23	-0.12	-0.05	-0.01	0.5	0
D18S967	21.74	-6.7	-1.18	-0.5	-0.19	-0.04	0.5	0
D18S464	33.37	-8.24	-1.65	-0.73	-0.29	-0.07	0.5	0
D18S1107	47.29	-6.29	-1.53	-0.69	-0.28	-0.07	0.5	0
D18S877	54.03	-0.5	-0.25	-0.12	-0.05	-0.01	0.5	0
D18S1102	59.89	-9.35	-1.66	-0.73	-0.28	-0.07	0.5	0
D18S474	68.99	-7.84	-0.49	-0.11	0	0.01	0.5	0
D18S450	70.78	-6.76	0.08	0.22	0.18	0.07	0.218	0.22
D18S64	77.54	-7.68	-0.79	-0.24	-0.05	0	0.5	0
D18S1134	87.2	-8.91	-0.79	-0.24	-0.05	0	0.5	0
D18S1147	88.5	-8.91	-0.79	-0.24	-0.05	0	0.5	0
D18S465	92.91	-6.69	-0.71	-0.27	-0.09	-0.01	0.5	0
D18S469	106.38	-8.74	-0.79	-0.24	-0.05	0	0.5	0
D18S554	117.94	-7.01	-1.45	-0.65	-0.26	-0.06	0.5	0

Table 4.23. Two-point lod scores for chromosome 18 in MMD family.

Locus	Genetic map	Recombination fraction					Theta max	Zmax
		0	0.1	0.2	0.3	0.4		
D19S886	0.9	-7.08	-1.46	-0.59	-0.21	-0.04	0.5	0
D19S591	10.59	0.74	0.62	0.46	0.27	0.09	0	0.74
D19S427	20.75	-9.38	-0.84	-0.26	-0.06	0	0.5	0
D19S584	33.17	-7.91	0.01	0.22	0.16	0.05	0.21	0.22
D19S226	37.98	-9.26	-1.08	-0.44	-0.16	-0.03	0.5	0
D19S593	43.02	-5.86	-0.73	-0.22	-0.04	0	0.5	0
D19S414	55.81	-7.66	-0.24	0.03	0.05	0.01	0.5	0
D19S208	63.1	-4.63	0.01	0.22	0.16	0.05	0.211	0.22
D19S903	73.21	-9.37	-0.84	-0.26	-0.06	0	0.5	0
D19S867	84.6	-5.68	-0.68	-0.2	-0.04	0	0.5	0
DG19S186	92.14	-0.06	-0.04	-0.02	-0.01	0	0.5	0
D19S572	99.88	-5.27	0.11	0.33	0.28	0.11	0.222	0.34
D19S418	107.71	-6.16	0.11	0.33	0.28	0.11	0.222	0.34
D19S605	108.45	-5.27	0.32	0.42	0.31	0.12	0.185	0.42
D19S573	113.23	-0.07	-0.05	-0.02	-0.01	0	0.5	0

Table 4.24. Two-point lod scores for chromosome 19 in MMD family.

Locus	Genetic map	Recombination fraction					Theta max	Zmax
		0	0.1	0.2	0.3	0.4		
D20S113	7.68	-7.44	-0.2	0.01	0.05	0.02	0.5	0.04
D20S882	18.3	-6.21	-0.58	-0.16	-0.02	0	0.5	0
D20S115	25.35	-6.61	-0.73	-0.29	-0.11	-0.02	0.5	0
D20S189	34.88	-8.55	-1.41	-0.64	-0.25	-0.06	0.5	0
D20S904	42.54	-7.76	-1.03	-0.42	-0.15	-0.03	0.5	0
D20S912	52.19	-0.44	-0.23	-0.11	-0.05	-0.01	0.5	0
D20S859	60.73	0.29	0.21	0.13	0.06	0.02	0	0.29
D20S107	63.19	-7.74	-0.87	-0.31	-0.1	-0.02	0.5	0
D20S108	65.64	-6.16	-0.88	-0.32	-0.12	-0.03	0.5	0
D20S481	70.97	-7.78	-0.87	-0.31	-0.11	-0.02	0.5	0
D20S838	71.23	-8.82	-1.08	-0.44	-0.16	-0.04	0.5	0
D20S178	75.47	-7.03	-1.47	-0.67	-0.27	-0.06	0.5	0
D20S1083	83.33	-7.94	-0.3	-0.1	-0.03	-0.01	0.5	0
D20S100	90.85	-8.62	-1.46	-0.59	-0.2	-0.04	0.5	0
D20S171	100.84	-3.4	-0.98	-0.4	-0.14	-0.03	0.5	0

Table 4.25. Two-point lod scores for chromosome 20 in MMD family.

		Recombination fraction						
Locus	Genetic map	0	0.1	0.2	0.3	0.4	Theta max	Zmax
D21S1904	7.23	-6.85	-0.25	-0.08	-0.02	0	0.5	0
D21S1432	8.75	-7	-1.22	-0.49	-0.17	-0.03	0.5	0
D21S1899	16.2	-6.66	-1.21	-0.48	-0.17	-0.03	0.5	0
D21S1905	17.29	-7.66	-1.2	-0.48	-0.16	-0.03	0.5	0
D21S1902	19.54	0.23	0.18	0.14	0.09	0.03	0	0.23
D21S272	23.96	-7.29	-1.16	-0.46	-0.16	-0.03	0.5	0
D21S1442	29.5	-8.62	-1.69	-0.75	-0.29	-0.07	0.5	0
D21S1909	33.81	-0.94	-0.19	-0.04	0	0	0.5	0.06
D21S1898	38.62	-0.18	-0.11	-0.06	-0.02	-0.01	0.5	0
D21S1895	41.37	-0.41	-0.22	-0.11	-0.04	-0.01	0.5	0
D21S1252	44.04	-8.79	-1.66	-0.73	-0.28	-0.07	0.5	0
D21S1919	45.79	-7.92	-0.53	-0.14	-0.01	0.01	0.5	0
D21S270	46.25	-0.52	-0.26	-0.13	-0.05	-0.01	0.5	0
D21S1255	48.28	-7.92	-0.55	-0.14	-0.02	0	0.5	0
D21S266	57.33	-8.09	-0.61	-0.17	-0.02	0	0.5	0.01

Table 4.26. Two-point lod scores for chromosome 21 in MMD family.

		Recombination fraction						
Locus	Genetic map	0	0.1	0.2	0.3	0.4	Theta max	Zmax
D22S420	5.72	-8.36	-0.5	-0.12	-0.01	0.01	0.5	0
D22S427	6	-0.09	-0.06	-0.03	-0.01	0	0.5	0
D22S539	15.82	-3.7	-0.75	-0.28	-0.08	-0.01	0.5	0
D22S1174	20.03	-5.33	-0.5	-0.12	-0.01	0.01	0.5	0
D22S315	23.91	0.32	0.23	0.14	0.07	0.02	0	0.32
D22S1154	25.73	-9.63	-0.76	-0.24	-0.05	0	0.5	0
D22S531	36.2	-0.33	-0.18	-0.09	-0.04	-0.01	0.5	0
D22S1265	40.84	-5.32	-0.5	-0.12	-0.01	0	0.5	0
D22S276	51.21	-6.11	-0.58	-0.16	-0.02	0	0.5	0
D22S928	59.39	-6.09	-1.1	-0.49	-0.19	-0.04	0.5	0
D22S1170	68.42	-8.17	-1.15	-0.49	-0.19	-0.05	0.5	0

Table 4.27. Two-point lod scores for chromosome 22 in MMD family.

Appendix 3.

Two point lod score results for Abdominal Epilepsy (AE) family in chapter seven.

Locus	genetic map	Recombination fraction					Theta max	Z max
		0	0.1	0.2	0.3	0.4		
D1S468	5.44	-13.54	-3.78	-1.85	-0.89	-0.34	0.5	0.19
D1S2870	11.68	-4.48	-0.06	0.12	0.15	0.11	0.289	0.15
D1S450	16.99	-18.98	-4.42	-2.14	-0.99	-0.34	0.5	0.11
D1S2667	20.36	-12.92	-6.38	-3.53	-1.78	-0.69	0.5	0.45
D1S434	21.8	-7.55	-2.97	-1.44	-0.65	-0.21	0.5	0.03
D1S2697	29.07	-18.36	-5.13	-2.7	-1.38	-0.55	0.5	0.49
D1S2620	43.94	-5.36	-1.74	-0.82	-0.38	-0.13	0.5	0.04
D1S241	54.23	-8.24	-0.42	-0.17	-0.09	-0.03	0.5	
D1S255	59.98	-5.05	-0.7	-0.4	-0.22	-0.1	0.5	0.3
D1S2861	68.18	-8.25	-2.73	-1.1	-0.36	-0.03	0.463	0.02
D1S2713	69.43	0.12	0.1	0.07	0.05	0.03	0	0.12
D1S2797	70.48	0.12	0.1	0.07	0.05	0.03	0	0.12
D1S197	73.67	-18.99	-4.25	-1.99	-0.86	-0.26	0.5	0.02
D1S2652	78.87	-14	-2.44	-0.91	-0.26	0	0.449	0.03
D1S476	80.34	-8.09	-1.03	-0.34	-0.06	0.04	0.418	0.04
D1S2846	87.95	-13.05	-1.8	-0.79	-0.31	-0.08	0.5	0
D1S2788	89.08	-4.94	-0.61	-0.33	-0.18	-0.07	0.5	0.1
D1S219	98.97							
D1S2841	105.86	-19.51	-2.63	-1.07	-0.39	-0.09	0.5	0
D1S500	106.4	-13.13	-0.63	0.06	0.21	0.14	0.303	0.21
D1S430	107.37	-23.06	-2.36	-0.84	-0.24	-0.03	0.5	0
D1S2766	112.11	-18.15	-1.32	-0.33	0.01	0.06	0.384	0.07
D1S2865	114.17	-5.42	-0.67	0.03	0.2	0.14	0.312	0.2
D1S435	117.67	-22.99	-1.77	-0.55	-0.11	-0.01	0.5	0.02
D1S206	124.95	-12.39	-1.41	-0.24	0.13	0.14	0.353	0.17
D1S495	126.35	-17.84	-0.27	0.34	0.4	0.23	0.268	0.42
D1S2688	129.08	-10.64	-1.36	-0.37	-0.02	0.05	0.385	0.04
D1S189	140.46	-4.81	-0.35	-0.13	-0.03	0.01	0.432	0.43
D1S498	147.7	-17.45	0.21	0.57	0.5	0.23	0.225	0.58
D1S1653	154.54	-5.05	-0.7	-0.4	-0.22	-0.1	0.5	0.3
D1S484	160.45	-19.2	-0.76	0.13	0.33	0.19	0.5	0.03
D1S2628	168.68	-18.26	-3.3	-1.41	-0.58	-0.19	0.5	0.11
D1S452	176.39	-8.13	-0.73	-0.16	0.01	0.03	0.5	0.67
D1S2818	187.03	-13.65	-1.95	-0.65	-0.17	-0.05	0.5	0.53
D1S413	198.52	-14.16	-2.49	-1.18	-0.53	-0.17	0.5	0.02
D1S2717	208.15	-5.25	1.04	0.94	0.66	0.29	0.114	1.05
D1S249	210.82	-19.38	-3.11	-1.26	-0.48	-0.16	0.5	0.28
D1S2685	213.64	-19.2	-3.37	-1.47	-0.63	-0.24	0.5	0.5
D1S245	216.77	0	0	0	0	0	0.1	0
D1S205	217.22	-19.08	-4.07	-1.86	-0.85	-0.34	0.5	0.66
D1S237	222.18	-18.46	-3.45	-1.71	-0.88	-0.4	0.5	1.21
D1S2641	230.38	-19.36	-4.6	-2.28	-1.11	-0.42	0.5	0.23
D1S2833	238.23	-12.2	-1.43	-0.7	-0.34	-0.13	0.5	0.09
D1S2709	239.64	-14.46	-2.23	-0.83	-0.23	0.01	0.444	0.03
D1S2850	251.42	-11.66	-0.57	0.13	0.29	0.21	0.307	0.29
D1S2785	262.68	-18.15	-1.02	-0.06	0.22	0.19	0.338	0.23

Table 7.5. Two-point lod scores for chromosome 1 in AE family

Locus	Genetic map	Recombination fraction					Theta max	Z max
		0	0.1	0.2	0.3	0.4		
D2S323	4.61	-5.14	0.96	0.74	0.42	0.11	0.082	0.97
D2S319	7.81	-9.91	-1.35	-0.19	0.2	0.23	0.356	0.25
D2S2211	18.24	-11.64	-0.61	0.28	0.5	0.39	0.31	0.5
D2S398	28.03	-12.33	-1.05	0.08	0.43	0.37	0.333	0.45
D2S168	29.73	-12.93	-1.94	-0.75	-0.21	0.01	0.444	0.03
D2S305	42.8	-13.03	-2	-0.52	0.06	0.19	0.388	0.19
D2S146	52.91	-7.25	-1.51	-0.43	0	0.12	0.399	0.12
D2S2328	66.77	-13.55	-0.45	0.36	0.5	0.32	0.284	0.5
D2S2156	78.42	-4.4	-0.58	0.02	0.2	0.18	0.337	0.22
D2S1364	82.83	-3.96	0.57	0.62	0.5	0.3	0.167	0.62
D2S337	86.05	-12.22	0.7	1.15	0.99	0.54	0.212	1.15
D2S2152	94.62	-12.19	**2.24**	2.04	1.5	0.75	0.115	2.24
D2S2110	99.58	-8.8	-0.09	0.15	0.14	0.04	0.5	0.12
D2S2116	103.5	-5.71	-0.9	-0.39	-0.15	-0.03	0.5	0
D2S388	112.75	-12.26	-1.74	-0.47	0.02	0.13	0.391	0.13
D2S293	122.55	-17.56	-1.41	-0.23	0.15	0.18	0.361	0.2
D2S363	130.95	-8.83	-0.43	0.21	0.34	0.24	0.298	0.34
D2S347	138.31	-17.55	-3.95	-1.74	-0.68	-0.16	0.5	0
D2S2271	143.62	-13.86	-4.2	-1.94	-0.82	-0.24	0.5	0.01
D2S2256	149.54	-6.18	-1.7	-1.09	-0.65	-0.27	0.5	0.21
D2S2324	164.37	-7.36	-2.02	-0.98	-0.44	-0.13	0.5	0.01
D2S1353	171.63	-13.86	-3.31	-1.44	-0.59	-0.17	0.5	0.02
D2S2330	174.85	-6.18	-1.09	-0.39	-0.09	0.02	0.431	0.03
D2S2188	186.25	-13.56	-2	-0.52	0.06	0.19	0.388	0.19
D2S118	195.34	-13.42	-1.6	-0.39	0.05	0.15	0.384	0.15
D2S2321	209.45	-11.85	-1.3	-0.12	0.28	0.29	0.352	0.33
D2S317	213.81	-6.83	-0.78	-0.31	-0.09	-0.01	0.458	0
D2S2361	216.94	-6.25	-1.3	-0.42	-0.06	0.06	0.414	0.06
D2S2382	218.52	-12.33	-1.58	-0.33	0.14	0.21	0.375	0.22
D2S163	226.25	-13.66	-2.41	-0.81	-0.14	0.09	0.42	0.09
D2S133	232.12	-5.73	-1.73	-0.9	-0.46	-0.18	0.5	0.15
D2S427	243.39	-9.75	-2.47	-1.04	-0.38	-0.07	0.48	0
D2S2202	253.23	-18.93	-3.59	-1.59	-0.63	-0.15	0.5	0
D2S338	253.84	-19.08	-3.06	-1.23	-0.44	-0.1	0.5	0
D2S395	262.12	-19.01	-3.6	-1.64	-0.72	-0.25	0.5	0.1
D2S140	265.28	-9.23	-2.43	-1.16	-0.52	-0.17	0.5	0.02

Table 7.6. Two-point lod scores for chromosome 2 in AE family

Locus	Genetic map	Recombination fraction					Theta max	Zmax
		0	0.1	0.2	0.3	0.4		
D3S4559	1.08	-5.72	-1.59	-0.51	-0.08	0.06	0.413	0.06
D3S1270	3.28	-12.77	-4.24	-2.16	-0.99	-0.34	0.5	0
D3S3630	6.22	-12.63	-3.13	-1.44	-0.63	-0.19	0.5	0
D3S3050	10.68	-9.22	-0.13	0.22	0.28	0.2	0.29	0.29
D3S2397	16.21	-12.93	-1.49	-0.47	-0.09	0.03	0.431	0.03
D3S1515	19.89	0.3	0.25	0.2	0.14	0.08	0	0.29
D3S3591	23.9	-7.85	-3.58	-1.8	-0.86	-0.31	0.5	0.01
D3S1263	30.71	-14.95	-2.57	-1.17	-0.53	-0.18	0.5	0.04
D3S2385	33.55	-6.23	-1.24	-0.59	-0.3	-0.13	0.5	0.2
D3S1567	48.03	-10.17	-2.53	-1.4	-0.76	-0.32	0.5	0.53
D3S2335	50.4	-5.58	-1.7	-0.81	-0.39	-0.15	0.5	0.09
D3S1266	52.22	-6.17	-2.55	-0.99	-0.31	-0.02	0.46	0.02
D3S3547	56.45	-3.87	0.3	0.4	0.35	0.21	0.205	0.4
D3S3521	66.4	-17.37	-0.73	0.12	0.34	0.26	0.317	0.34
D3S3672	76.71	-5.15	-2.45	-0.88	-0.21	0.03	0.434	0.05
D3S1600	89.02	-4.06	0.78	0.75	0.45	0.1	0.138	0.82
D3S3551	99.23	-13.34	-0.05	0.54	0.56	0.31	0.251	0.6
D3S3653	106.92	**2.73**	2.28	1.77	1.18	0.52	0	2.73
D3S1271	115.24	-8.26	1.53	1.32	0.93	0.41	0.091	1.53
D3S3045	120.29	-8.01	1.79	1.53	1.07	0.49	0.085	1.79
D3S1558	129.21	-12.83	1.55	1.66	1.29	0.65	0.162	1.7
D3S1589	136.43	-7.69	0.58	0.88	0.74	0.35	0.205	0.9
D3S3637	146.26	-10.88	-0.74	-0.04	0.12	0.06	0.5	0.9
D3S1569	154.11	-11.54	-1.29	-0.38	-0.02	0.08	0.403	0.08
D3S1308	160.67	-5.05	-0.7	-0.4	-0.22	-0.1	0.5	0.3
D3S1607	168.31	-4.88	-0.79	-0.23	-0.04	0.01	0.5	0.02
D3S1564	175.21	-5.05	-0.7	-0.4	-0.22	-0.1	0.5	0.3
D3S3725	178.27	-12.63	-4.14	-2.18	-1.12	-0.44	0.5	0.4
D3S3041	185.92	0	0	0	0	0	0.1	0
D3S3592	195.05	-16.92	-3.06	-1.23	-0.44	-0.1	0.5	0
D3S1580	206.43	-11.05	-0.36	0.28	0.38	0.22	0.275	0.4
D3S3663	213.15	-5.05	-0.7	-0.4	-0.22	-0.1	0.5	0.3
D3S240	219.97	0.12	0.1	0.08	0.05	0.03	0	0.12
D3S1265	222.34	-12.63	-3.13	-1.56	-0.74	-0.26	0.5	0
D3S1311	225.05	-17.54	-3.96	-1.81	-0.8	-0.27	0.5	0

Table 7.7. Two-point lod scores for chromosome 3 in AE family

Locus	Genetic map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D4S412	4.57	-12.94	-1.99	-0.66	-0.12	0.06	0.418	0.06
D4S2935	14.07	-5.05	-0.71	-0.4	-0.23	-0.1	0.5	0.3
D4S394	15.28	-12.46	-1.37	-0.33	0.07	0.15	0.383	0.15
D4S2928	23.84	-5.05	-0.7	-0.4	-0.22	-0.1	0.5	0.3
D4S3341	32.25	-11.79	-0.62	0.19	0.37	0.27	0.308	0.37
D4S404	41	-11.73	-0.73	0.13	0.32	0.21	0.304	0.32
D4S3022	43.89	-13.69	-0.57	0.24	0.38	0.23	0.286	0.38
D4S391	48.95	-13.67	-1.61	-0.42	0	0.07	0.382	0.07
D4S2912	48.97	-10.11	-1.41	-0.58	-0.22	-0.05	0.495	0
D4S1587	55.94	-11.48	-1.8	-0.73	-0.29	-0.12	0.5	0.5
D4S405	62.1	-14.31	-1.95	-0.72	-0.22	-0.03	0.472	0
D4S428	71.03	-7.49	-2.27	-0.92	-0.32	-0.06	0.48	0
D4S2389	84.72	-9.24	-1.38	-0.43	-0.07	0.03	0.424	0.03
D4S1553	92.45	-4.98	-1.22	-0.52	-0.19	-0.03	0.481	0
D4S423	102.7	-12.64	-1.9	-0.68	-0.19	-0.01	0.459	0.01
D4S1572	110.14	-17.36	-2.96	-1.15	-0.34	0	0.451	0.04
D4S406	117.48	-12.93	-2.02	-0.71	-0.15	0.06	0.427	0.06
D4S1615	129.75	-6.97	-1.2	-0.49	-0.17	-0.02	0.469	0
D4S1575	134.67	-4.09	0.06	0.17	0.16	0.1	0.23	0.2
D4S424	142.05	-5.08	-0.33	-0.1	-0.01	0.02	0.404	0.02
D4S1586	144.42	-17.64	-2.46	-0.89	-0.19	0.06	0.428	0.08
D4S2962	148.06	-7.65	-0.11	0	0.01	0	0.5	0.01
D4S3046	159.33	-14.3	-1.64	-0.39	0.09	0.18	0.381	0.2
D4S1539	173.75	**1.17**	0.91	0.63	0.35	0.1	0	1.17
D4S1607	184.14	-13.49	-1.78	-0.68	-0.22	-0.03	0.47	0
D4S3041	189.52	0.11	0.07	0.03	0.01	0	0	0.11
D4S408	193.5	-5.09	-0.72	-0.4	-0.22	-0.1	0.5	0.24
D4S1540	195.66	0.12	0.1	0.08	0.05	0.03	0	0.12
D4S2930	209.27	-6.59	-3.13	-1.56	-0.74	-0.26	0.5	0.09

Table 7.8. Two-point lod scores for chromosome 4 in AE family

259

Locus	Genetic Map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D5S1981	1.21	-13.71	-3.34	-1.51	-0.67	-0.26	0.5	0.43
D5S2088	12.05	-7.37	-2.7	-1.37	-0.67	-0.25	0.5	0.1
D5S2054	16.64	-5.94	-2.19	-0.96	-0.38	-0.09	0.5	0
D5S635	18.11	-5.92	-2.17	-0.95	-0.37	-0.09	0.5	0
D5S1486	27.25	-13.56	-3.24	-1.33	-0.45	-0.06	0.47	0.01
D5S2081	31.94	-6.19	-1.31	-0.6	-0.27	-0.09	0.5	0.03
D5S2031	42.23	-5.54	-0.7	-0.3	-0.12	-0.03	0.5	0
D5S674	56.71	-6.26	-1.28	-0.38	-0.02	0.08	0.402	0.08
D5S2021	61.33	-6.23	-0.81	-0.17	0.07	0.11	0.373	0.11
D5S628	71.14	0	0	0	0	0	0.1	0
D5S2072	80.82	0.42	0.35	0.28	0.2	0.11	0	0.42
D5S424	93.23	-0.01	-0.01	0	0	0	0.1	0
D5S1397	98.14	-5.47	-0.45	0.12	0.27	0.22	0.321	0.28
D5S1725	106.88	0.23	0.2	0.16	0.11	0.05	0	0.23
D5S409	114.89	-6.95	-1.34	-0.31	0.07	0.14	0.377	0.14
D5S2027	120.95	-8.51	0.05	0.29	0.26	0.12	0.227	0.3
D5S2065	122.49	-17.84	-2.41	-0.87	-0.2	0.05	0.432	0.06
D5S490	134.52	0.3	0.25	0.2	0.15	0.08	0	0.3
D5S2017	145.35	-13.23	-0.48	0.15	0.28	0.2	0.305	0.28
D5S2090	153.72	-9.21	-1.14	-0.35	-0.03	0.06	0.408	0.07
D5S410	162.45	-14.16	-2.58	-1.05	-0.38	-0.09	0.5	0
D5S2066	171.45	-6.6	-3.42	-1.67	-0.8	-0.3	0.5	0.21
D5S400	181.34	-18.48	-4.36	-2.23	-1.09	-0.41	0.5	0.19
D5S1973	184.16	0.4	0.34	0.27	0.19	0.1	0	0.4
D5S2069	194.4	-7.25	-2.01	-0.83	-0.29	-0.04	0.469	0
D5S469	203.46	-5.72	-1.59	-0.51	-0.08	0.06	0.413	0.06
D5S2006	211.06	-19.13	-2	-0.74	-0.22	-0.03	0.472	0

Table 7.9. Two-point lod scores for chromosome 5 in AE family

Locus	Genetic Map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D6S244	5.66	-7.22	-1.53	-0.58	-0.16	0.01	0.446	0.02
D6S1617	12.18	-9.36	-0.96	-0.21	0.05	0.08	0.365	0.08
D6S309	21.6	-13.56	-3.51	-1.62	-0.75	-0.31	0.5	0.79
D6S1279	30.35	-4.94	0.33	0.45	0.32	0.13	0.184	0.45
D6S422	44.68	-11.2	-2.23	-0.76	-0.2	-0.05	0.5	0.41
D6S1660	48.37	-8.7	-1.27	-0.15	0.18	0.13	0.5	0.06
D6S273	53.25	-12.04	-1.62	-0.75	-0.32	-0.09	0.5	0
D6S1575	63.69	-13.85	-3.44	-1.57	-0.7	-0.25	0.5	0.24
D6S459	72.59	-6.64	-1.64	-0.86	-0.44	-0.17	0.5	0.14
D6S452	74.58	-6.62	-0.89	-0.42	-0.19	-0.06	0.5	0.02
D6S1557	87.06	-13.86	-3.15	-1.37	-0.55	-0.16	0.5	0.02
D6S460	93.4	-13.76	-3.85	-1.76	-0.78	-0.26	0.5	0.09
D6S458	102.77	-9.07	-4.3	-2.14	-1.04	-0.4	0.5	0.26
D6S268	115.03	-19.38	-2.92	-1.3	-0.54	-0.16	0.5	0.03
D6S1712	123.79	-6.23	-2.66	-1.34	-0.65	-0.24	0.5	0.09
D6S1656	134.76	-14.63	-1.57	-0.5	-0.11	-0.03	0.5	0.44
D6S270	138.78	0.42	0.35	0.28	0.2	0.11	0	0.42
D6S1009	142.71	-15.87	-1.27	-0.12	0.21	0.16	0.328	0.23
D6S1637	154.79	-18.96	-2.02	-0.85	-0.32	-0.07	0.5	0
D6S441	164.5	-19.37	-2.16	-0.66	-0.11	0.02	0.5	0.02
D6S419	170.85	-15.34	-2.19	-0.85	-0.3	-0.09	0.5	0.1
D6S1581	174.43	-11.62	-1.24	-0.46	-0.17	-0.07	0.5	0.25
D6S305	176.93	-19.2	-3.31	-1.44	-0.59	-0.17	0.5	0.02
D6S1599	179.13	-19.28	-4.95	-2.39	-1.09	-0.37	0.5	0.08
D6S1277	180.81	-19.38	-4.96	-2.44	-1.18	-0.45	0.5	0.32
D6S297	187.65	-5.36	-0.6	-0.32	-0.17	-0.07	0.5	0.11
D6S503	190.17	-13.95	-5.25	-2.65	-1.28	-0.47	0.5	0.19
D6S446	194.22	-10.71	-2.36	-1.09	-0.51	-0.2	0.5	0.1

Table 7.10. Two-point lod scores for chromosome 6 in AE family

Locus	Genetic Map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D7S2474	3.71	-11.82	-1.91	-0.62	-0.1	0.07	0.418	0.07
D7S2201	12.21	-5.26	-0.34	-0.11	-0.02	0.01	0.416	0.01
D7S513	23.24	-14.22	-0.85	-0.2	0.05	0.1	0.381	0.1
D7S664	26.72	-4.54	0.38	0.46	0.38	0.22	0.184	0.46
D7S1795	38.61	-7.06	-0.06	0.21	0.21	0.11	0.247	0.23
D7S2463	40.51	-8.58	-1.17	-0.36	-0.04	0.06	0.41	0.06
D7S516	45.36	-17.55	-2	-0.52	0.06	0.19	0.388	0.19
D7S2250	55.9	-14.09	-0.32	0.23	0.35	0.26	0.3	0.35
D7S2541	63.45	-13.86	-2.65	-1.02	-0.3	0	0.448	0.04
D7S691	64.93	-7.03	-0.92	-0.13	0.15	0.17	0.361	0.18
D7S519	70.24	-5.93	-1.27	-0.38	-0.02	0.08	0.401	0.08
D7S2429	80.22	-8.18	-3.6	-1.62	-0.66	-0.17	0.5	0
D7S502	80.93	-11.21	-1.65	-0.68	-0.24	-0.04	0.474	0
D7S2500	83.71	-8.56	-1.26	-0.43	-0.09	0.03	0.427	0.03
D7S672	86.43	-17.46	-2.96	-1.12	-0.31	0.02	0.444	0.05
D7S2443	94.47	-17.44	-5.24	-2.69	-1.32	-0.49	0.5	0.27
D7S2485	100.34	-13.37	-0.67	0.02	0.2	0.16	0.323	0.2
D7S657	105.59	-23.06	-3	-1.21	-0.43	-0.09	0.5	0
D7S554	109.67	-7.47	-0.74	-0.44	-0.26	-0.1	0.5	0.02
D7S2509	114.48	-14.07	-2.94	-1.33	-0.56	-0.16	0.5	0
D7S2459	120.05	-12.63	-1.21	-0.24	0.05	0.06	0.5	0.08
D7S2418	122.15	-11	-1.54	-0.5	-0.08	0.06	0.416	0.06
D7S530	133.2	0	0	0	0	0	0.1	0
D7S2533	141.74	-18.81	-1.86	-0.62	-0.15	-0.01	0.463	0
D7S684	149.78	-19.04	-1.18	-0.26	0.03	0.04	0.5	0.04
D7S1824	152.03	-14.47	-2.19	-0.96	-0.37	-0.08	0.5	0
D7S661	155.04	-11.37	-2.44	-1.17	-0.52	-0.17	0.5	0.02
D7S1805	165.55	-13.83	-0.56	0.08	0.22	0.13	0.298	0.22
D7S2447	179.62	-6.11	-1.17	-0.68	-0.41	-0.2	0.5	0.82
D7S2423	192.24	-14.16	-2.89	-1.27	-0.54	-0.17	0.5	0.25

Table 7.11. Two-point lod scores for chromosome 7 in AE family

Locus	Genetic map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D8S264	3.54	-10.46	-3.11	-1.25	-0.45	-0.1	0.5	0
D8S1819	16.91	-12.58	-2.81	-1.14	-0.38	-0.04	0.466	0.01
D8S520	21.48	-8.16	-1.6	-0.34	0.13	0.21	0.374	0.22
D8S516	22.5	-12.01	-1.81	-0.53	-0.04	0.1	0.404	0.1
D8S258	36.13	0.12	0.1	0.08	0.05	0.03	0	0.12
D8S1734	42.1	-2.83	1.17	1.1	0.85	0.48	0.118	1.18
D8S1820	49.17	-2.56	2.41	2.17	1.59	0.79	0.106	2.41
D8S505	55.98	-2.82	2.17	1.96	1.42	0.71	0.109	2.17
D8S1737	69.34	0.38	0.9	0.78	0.52	0.19	0.101	0.9
D8S1763	73.88	-2.23	2.05	1.74	1.22	0.56	0.078	2.06
D8S512	78.06	**4.18**	3.48	2.7	1.83	0.86	0	4.18
D8S279	86.52	-5.91	1.57	1.69	1.35	0.77	0.164	1.72
D8S1707	97.46	2.11	1.79	1.43	1.02	0.55	0	2.11
D8S1778	106.89	2.7	2.28	1.83	1.31	0.71	0	2.7
D8S1762	108.03	2.76	2.92	2.39	1.68	0.82	0.045	3.04
D8S1470	118.7	-3.01	1.36	1.23	0.93	0.52	0.104	1.36
D8S1799	131.69	-13.51	-0.23	0.44	0.54	0.36	0.278	0.54
D8S1793	135.13	-4.8	-1.4	-0.21	0.19	0.23	0.359	0.25
D8S263	141.47	-4.4	-0.87	-0.21	0.04	0.09	0.384	0.09
D8S1746	150.11	-10.46	-2.56	-0.96	-0.26	0.01	0.444	0.04
D8S1783	152.34	-10.17	-2.48	-0.9	-0.23	0.02	0.439	0.04
D8S274	153.61	-11.07	-1.28	-0.1	0.28	0.28	0.347	0.32
D8S1743	160.11	-6.81	-0.61	-0.11	0.02	0.02	0.5	0.37

Table 7.12. Two-point lod scores for chromosome 8 in AE family

Locus	Genetic Map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D9S1779	0	-17.54	-5.83	-3.02	-1.54	-0.63	0.5	0.97
D9S1686	11.93	0	0	0	0	0	0.1	0
D9S286	18.53	-14.18	-2.8	-1.22	-0.51	-0.17	0.5	0.05
D9S168	24.28	-13.86	-1.41	-0.24	0.12	0.09	0.5	0.5
D9S1808	28.24	-14.17	-0.6	-0.09	0.02	-0.02	0.5	0.63
D9S235	32.65	-5.05	-0.7	-0.4	-0.22	-0.1	0.5	0.29
D9S171	47.16	-7.68	-0.67	-0.08	0.13	0.13	0.354	0.15
D9S2149	55.3	-13.55	0.09	0.62	0.6	0.29	0.238	0.66
D9S1777	67.34	**2.14**	1.78	1.37	0.9	0.37	0	2.14
D9S1876	70.42	-12.83	0.9	1.14	0.92	0.45	0.189	1.14
D9S175	72.8	-6.66	-0.68	-0.09	0.12	0.13	0.357	0.15
D9S1843	79.87	-5.75	-1.73	-0.89	-0.45	-0.18	0.5	0.14
D9S307	89.23	-13.42	-0.28	0.47	0.54	0.29	0.5	0.04
D9S283	95	-5.79	-0.17	0.12	0.14	0.06	0.5	0.16
D9S287	101.54	-14.1	-2.43	-1.16	-0.52	-0.16	0.5	0.02
D9S261	114.65	-6.35	-2.79	-1.38	-0.67	-0.26	0.5	0.16
D9S1675	117.44	-17.35	-2	-0.87	-0.35	-0.1	0.5	0
D9S1828	117.67	-7.7	-0.85	0.01	0.25	0.2	0.328	0.26
D9S1824	123.56	-13.86	-2.76	-1.2	-0.49	-0.15	0.5	0.03
D9S934	127.72	-9.36	-2.2	-0.85	-0.25	0	0.45	0.03
D9S1682	132.09	-17.85	-3.16	-1.3	-0.44	-0.06	0.47	0.01
D9S1825	136.35	-11.11	-1.86	-0.58	-0.06	0.1	0.408	0.1
D9S1830	149.36	-5.74	-0.29	0.04	0.13	0.11	0.321	0.13
D9S1793	150.42	-13.62	-1.17	-0.18	0.17	0.2	0.36	0.21
D9S1838	164.1	-9.41	-1.78	-0.56	-0.09	0.06	0.42	0.06

Table 7.13. Two-point lod scores for chromosome 9 in AE family

Locus	Genetic map	Recombination fraction					Theta max	Zmax
		0	0.1	0.2	0.3	0.4		
D10S249	1.19	-5.4	-0.05	0.18	0.16	0.05	0.5	0.25
D10S602	5.71	-18.71	0.13	0.66	0.64	0.33	0.242	0.69
D10S1218	11.3	0	0	0	0	0	0.1	0
D10S189	20.56	-11.99	-1.86	-0.6	-0.09	0.07	0.415	0.07
D10S1649	26.55	-12.5	-3.37	-1.47	-0.57	-0.14	0.5	0
D10S1430	32.11	-14.06	-2.89	-1.24	-0.47	-0.11	0.5	0
D10S570	32.73	-17.56	-2.56	-1.01	-0.35	-0.07	0.5	0
D10S191	37.42	-17.66	-4.31	-2.09	-0.98	-0.35	0.5	0.16
D10S1653	38.92	-12.39	-4.38	-2.12	-0.99	-0.34	0.5	0.11
D10S1661	41.4	-14.06	-4.45	-2.2	-1.05	-0.38	0.5	0.15
D10S600	54.33	-13.56	-5.35	-2.72	-1.33	-0.49	0.5	0.21
D10S1746	65.42	-13.45	-5.07	-2.57	-1.26	-0.47	0.5	0.25
D10S196	71.95	-13.38	-3.93	-1.98	-0.98	-0.38	0.5	0.24
D10S609	82.33	0	0	0	0	0	0.1	0
D10S537	91.46	-7.6	-5.49	-3.02	-1.6	-0.66	0.5	0.78
D10S580	97.94	-13.56	-6.03	-3.18	-1.63	-0.65	0.5	0.57
D10S1677	100.72	-6.16	-4.42	-2.22	-1.11	-0.48	0.5	1.38
D10S1753	113.42	-3.7	0.32	0.5	0.39	0.17	0.199	0.5
D10S185	116.07	-13.95	-3.2	-1.38	-0.58	-0.23	0.5	0.9
D10S1267	125.15	-5.78	-1.16	-0.23	0.06	0.04	0.5	0.43
D10S597	129.57	0.89	0.76	0.61	0.44	0.24	0	0.89
D10S1693	141.5	-13.73	-0.86	0.04	0.26	0.19	0.316	0.26
D10S1656	154.8	-13.81	-2.09	-0.88	-0.33	-0.08	0.5	0
D10S575	160.19	-6.11	-0.19	0.11	0.14	0.05	0.5	0.14
D10S1676	164.6	-6.66	-1	-0.49	-0.23	-0.08	0.5	0.02
D10S1651	176.24	-19.25	-4.01	-1.82	-0.76	-0.21	0.5	0
D10S212	181.66	-8.75	-2.45	-0.95	-0.29	-0.02	0.459	0.02

Table 7.14. Two-point lod scores for chromosome 10 in AE family

Locus	Genetic map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D11S4046	0	-6.47	-1.61	-0.41	0.02	0.1	0.389	0.11
D11S1760	8.79	-4.45	-0.99	-0.36	-0.13	-0.04	0.5	0.06
D11S1331	12.38	-6.68	-1.19	-0.48	-0.16	-0.02	0.468	0
D11S4170	23.4	-11.85	-2.39	-1.13	-0.5	-0.15	0.5	0.02
D11S902	26.23	-6.17	-1.35	-0.22	0.13	0.12	0.341	0.16
D11S928	34.9	-4.34	-0.11	0.07	0.11	0.08	0.303	0.11
D11S4080	43.29	-18.89	-1.84	-0.62	-0.16	-0.03	0.5	0.03
D11S1776	49.46	-5.49	-0.61	-0.3	-0.14	-0.06	0.5	0.07
D11S1360	58.18	-6.93	-1.56	-0.79	-0.41	-0.16	0.5	0.07
D11S4191	66.18	-10.19	-1.53	-0.66	-0.3	-0.14	0.5	0.54
D11S4087	73.81	-4.87	-0.1	0.12	0.1	0.03	0.5	0.03
D11S937	85.42	-18.67	-2.23	-1.06	-0.52	-0.23	0.5	0.75
D11S1780	94.12	-9.74	-0.66	-0.19	0	0.05	0.395	0.05
D11S1886	107.71	-6.77	-0.44	-0.09	0	0.01	0.5	0.1
D11S4206	110.81	0.07	0.07	0.07	0.06	0.04	0.06	0.07
D11S908	118.66	-7.84	0.07	0.45	0.42	0.22	0.236	0.47
D11S4089	126.83	-6.17	-1.41	-0.24	0.13	0.14	0.353	0.17
D11S4151	135.73	-12.01	-1.04	-0.29	-0.05	-0.02	0.5	0.27
D11S2367	148.23	-10.13	-1.47	-0.57	-0.19	-0.04	0.5	0
D11S969	154.38	-13.73	-3.38	-1.43	-0.52	-0.09	0.481	0

Table 7.15. Two-point lod scores for chromosome 11 in AE family

Locus	Genetic map	Recombination fraction					Theta Max	Zmax
		0.0	0.1	0.2	0.3	0.4		
D12S352	0.68	-13.33	-0.8	-0.17	0.03	0.05	0.359	0.06
D12S1656	4.23	0	0	0	0	0	0.1	0
D12S372	8.63	-12.74	-1.66	-0.64	-0.22	-0.06	0.5	0.01
D12S374	16.84	-14.38	-1.07	-0.2	0.07	0.08	0.355	0.1
D12S336	24.76	-4.14	0.23	0.34	0.3	0.18	0.211	0.34
D12S364	31.97	-14.06	-1.1	-0.18	0.14	0.17	0.36	0.18
D12S1591	44.81	-12.93	-0.48	0.15	0.29	0.21	0.305	0.28
D12S1337	51.8	-6.96	-0.74	-0.25	-0.13	-0.08	0.5	0.11
D12S1704	52.93	-7.33	-0.24	0.2	0.28	0.2	0.29	0.28
D12S85	61.77	-8.44	-0.58	-0.04	0.14	0.13	0.343	0.15
D12S368	67.92	-23.07	-2.3	-0.79	-0.17	0.05	0.428	0.06
D12S83	75.48	-18.88	-2.3	-0.79	-0.17	0.05	0.428	0.06
D12S1294	83.69	-12.64	-3.32	-1.61	-0.75	-0.27	0.5	0.09
D12S326	92.97	-13.23	-5.8	-2.92	-1.41	-0.51	0.5	0.2
D12S351	103.1	-0.01	-0.01	-0.01	0	0	0.5	0
D12S95	104.42	-11.05	-3.58	-1.89	-0.96	-0.38	0.5	0.28
D12S1346	106.03	-10.65	-1.48	-0.6	-0.26	-0.1	0.5	0.05
D12S2081	110.59	-5.77	-3.18	-1.58	-0.75	-0.27	0.5	0.09
D12S1636	120.17	-19.02	-4.4	-2.28	-1.13	-0.42	0.5	0.21
D12S1613	124.08	-4.26	-0.06	0.09	0.11	0.08	0.282	0.11
D12S1583	127.22	-18.43	-3.92	-1.89	-0.88	-0.3	0.5	0.1
D12S354	133.66	-3.72	-0.89	-0.39	-0.15	-0.04	0.5	0
D12S369	134.33	-17.21	-3.93	-1.73	-0.68	-0.16	0.5	0
D12S366	140.32	-10.67	-2.6	-1.3	-0.62	-0.22	0.5	0.08
D12S2073	144.14	-6.81	-2.1	-0.92	-0.37	-0.1	0.5	0
D12S2078	155.99	-6.6	-1.18	-0.34	0	0.09	0.397	0.09
D12S1609	160.72	-7.49	-0.98	-0.34	-0.07	0.03	0.422	0.03
D12S1723	173.29	-12.62	-2.95	-1.19	-0.42	-0.09	0.5	0

Table 7.16. Two-point lod scores for chromosome 12 in AE family

Locus	Genetic Map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D13S175	3.24	-14.25	-3.85	-1.91	-0.92	-0.33	0.5	0.1
D13S232	8.98	-9.67	-2.88	-1.38	-0.66	-0.25	0.5	0.25
D13S1243	11.57	0.12	0.1	0.07	0.05	0.03	0	0.12
D13S221	16.88	-12.31	-2.11	-1.21	-0.71	-0.31	0.5	0.3
D13S217	22.89	-14.19	-1.28	-0.46	-0.15	-0.06	0.5	0.21
D13S289	28.14	-7.26	-2.33	-1.26	-0.67	-0.27	0.5	0.32
D13S171	31.99	-14.45	-5.32	-2.83	-1.42	-0.55	0.5	0.34
D13S219	36.49	-8.89	-4.07	-1.92	-0.86	-0.28	0.5	0.05
D13S218	40.44	-11.1	-4.5	-2.36	-1.14	-0.42	0.5	0.17
D13S272	54.31	-7.07	-4.16	-2.15	-1.12	-0.48	0.5	1.26
D13S279	65.83	-13.78	-3.51	-1.62	-0.74	-0.28	0.5	0.35
D13S271	80.33	-11.49	-5.86	-3.11	-1.61	-0.68	0.5	1.35
D13S1241	91.34	-11.36	-4.14	-2.18	-1.12	-0.44	0.5	0.36
D13S1256	100.66	-17.75	-3.72	-1.73	-0.75	-0.22	0.5	0.02
D13S1809	107.62	-18.73	-2.17	-0.97	-0.39	-0.1	0.5	0
D13S1315	119	-13.51	-3.05	-1.36	-0.55	-0.14	0.5	0
D13S293	131.92	**2.4**	2.04	1.63	1.16	0.63	0	2.4

Table 7.17. Two-point lod scores for chromosome 13 in AE family

Locus	Genetic map	Recombination fraction					Theta Max	Z max
		0	0.1	0.2	0.3	0.4		
D14S261	9.53	-11.3	-3.41	-1.5	-0.59	-0.14	0.5	0
D14S1043	9.83	-9.71	-0.97	-0.28	-0.05	0	0.5	0
D14S1280	23.04	-7.41	-0.41	0.27	0.44	0.33	0.301	0.44
D14S262	25.73	-7.68	-0.52	0.05	0.22	0.19	0.333	0.23
D14S1071	29.19	-13.06	-1.95	-0.49	0.07	0.2	0.386	0.2
D14S741	35	-13.56	-1.95	-0.49	0.07	0.2	0.383	0.2
D14S70	37.63	-11.31	-1.89	-0.89	-0.39	-0.12	0.5	0.01
D14S75	43.2	-12.75	-1.37	-0.33	0.07	0.15	0.385	0.15
D14S976	46.8	-7.54	0.22	0.55	0.51	0.3	0.228	0.56
D14S978	51.44	-17.59	-1.41	-0.28	0.07	0.09	0.355	0.11
D14S276	57.33	-13.45	-2.25	-0.79	-0.2	0.01	0.443	0.03
D14S1011	70.01	-13.26	-4.15	-1.97	-0.89	-0.3	0.5	0.06
D14S258	70.4	-13.38	-4.15	-1.97	-0.89	-0.3	0.5	0.06
D14S1433	74.94	-7.87	-1.64	-0.84	-0.46	-0.2	0.5	0.13
D14S616	85.45	-11.32	-1.94	-0.99	-0.49	-0.19	0.5	0.1
D14S67	88.68	-6.37	-2.76	-1.48	-0.78	-0.32	0.5	0.25
D14S1044	92.86	-16.95	-4.11	-2.18	-1.12	-0.44	0.5	0.33
D14S81	99.25	-17.07	-2.25	-0.99	-0.39	-0.09	0.5	0
D14S1054	102.52	-17.21	-2.26	-0.99	-0.39	-0.09	0.5	0
D14S987	106.22	-12.01	-1.54	-0.59	-0.17	0.01	0.446	0.02
D14S1019	109.31	-17.26	-2.26	-1	-0.39	-0.09	0.5	0
D14S1426	117.81	-5.05	-0.71	-0.4	-0.23	-0.1	0.5	0.3
D14S542	126.06	-19.38	-3.31	-1.41	-0.54	-0.13	0.5	0

Table 7.18. Two-point lod scores for chromosome 14 in AE family

268

Locus	Genetic map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D15S128	6.06	-11.7	-4.25	-2.04	-0.95	-0.35	0.5	0.2
D15S975	13.46	-13.73	-1.6	-0.69	-0.3	-0.11	0.5	0.14
D15S1019	21.44	-19.08	-3.11	-1.26	-0.48	-0.16	0.5	0.28
D15S165	23.51	-14.45	-1.53	-0.58	-0.16	0.01	0.5	0.02
D15S231	27.43	-7.29	-0.32	0	0.09	0.08	0.333	0.09
D15S118	35.23	-13.81	-2.18	-1	-0.47	-0.19	0.5	0.56
D15S1012	39.39	-9.31	-0.97	-0.02	0.26	0.23	0.338	0.28
D15S146	43.14	-12.23	0.07	0.6	0.6	0.37	0.247	0.64
D15S1016	52.71	-4.88	-0.77	-0.24	-0.02	0.04	0.408	0.04
D15S117	58.09	-11.94	-1.65	-0.41	0.05	0.15	0.386	0.15
D15S988	70.39	0.42	0.35	0.28	0.2	0.11	0.001	0.42
D15S114	82.51	-9.83	-1.31	-0.5	-0.17	-0.04	0.5	0
D15S1005	86.82	-9.25	-0.55	0.13	0.29	0.22	0.313	0.29
D15S158	100.89	-7.35	-0.79	-0.31	-0.1	-0.01	0.46	0
D15S1004	110.96	-13.56	-4.07	-1.85	-0.78	-0.22	0.5	0.01
D15S130	111.55	-14.46	-2.42	-1.03	-0.39	-0.09	0.5	0
D15S816	114.36	-7.35	-3.3	-1.69	-0.85	-0.33	0.5	0.32
D15S157	116.14	-9.08	-2.46	-1.03	-0.36	-0.06	0.5	0
D15S212	124.08	-12.84	-4.52	-2.35	-1.19	-0.46	0.5	0.3
D15S120	130.4	-13.98	-2.72	-1.13	-0.44	-0.13	0.5	0.07

Table 7.19. Two-point lod scores for chromosome 15 in AE family

Locus	Genetic map	0	0.1	0.2	0.3	0.4	Theta Max	Zmax
D16S521	1.15	-13.68	-3	-1.37	-0.61	-0.21	0.5	0.4
D16S3065	10.64	-8.31	-0.27	0.17	0.24	0.16	0.282	0.25
D16S423	14.97	-9.17	-3.44	-1.7	-0.8	-0.28	0.5	0.08
D16S418	20.61	-13.83	-3.94	-1.77	-0.72	-0.19	0.5	0
D16S3062	32.07	-8.83	-3.38	-1.45	-0.54	-0.1	0.5	0
D16S500	33.31	-5.93	-2.43	-1.16	-0.52	-0.17	0.5	0.02
D16S410	41.83	-13.81	-4.53	-2.25	-1.08	-0.39	0.5	0.19
D16S3068	51.27	-8.16	-3.31	-1.47	-0.61	-0.18	0.5	0.02
D16S3080	61.1	-5.79	-0.88	-0.42	-0.2	-0.08	0.5	0.08
D16S3034	68.27	-14.16	-4.37	-2.11	-0.96	-0.31	0.5	0.08
D16S3057	76.36	-17.55	-4.81	-2.3	-1.03	-0.33	0.5	0.05
D16S514	83.1	-14.16	-4.49	-2.22	-1.05	-0.38	0.5	0.16
D16S503	84.26	-13.96	-3.8	-1.86	-0.89	-0.32	0.5	0.12
D16S515	94.69	-10.31	-3.51	-1.73	-0.81	-0.28	0.5	0.08
D16S516	101.24	-7.35	-0.96	-0.32	-0.05	0.03	0.416	0.04
D16S505	108.07	-13.77	-2.7	-1.14	-0.42	-0.09	0.5	0
D16S763	120.63	-13.45	-0.93	-0.35	-0.12	-0.03	0.5	0
D16S2621	131.46	-13.45	-1.37	-0.55	-0.21	-0.05	0.5	0

Table 7.20. Two-point lod scores for chromosome 16 in AE family

Locus	Genetic map	Recombination fraction					Theta Max	Zmax
		0	0.1	0.2	0.3	0.4		
D17S849	0.63	-6.49	-0.84	-0.19	0.05	0.1	0.379	0.1
D17S831	7.47	-13.86	-2.38	-0.84	-0.18	0.05	0.429	0.06
D17S1832	16.79	-5.86	-0.68	-0.07	0.14	0.14	0.355	0.16
D17S804	28.96	-4.78	1.41	1.29	0.98	0.56	0.107	1.41
D17S799	37.94	-4.65	0.69	0.85	0.73	0.44	0.195	0.85
D17S922	41.55	-5.11	0.27	0.5	0.46	0.28	0.228	0.5
D17S839	43.77	-0.59	0.1	0.19	0.16	0.09	0.214	0.19
D17S1824	53.22	-4.68	0.01	0.17	0.18	0.12	0.267	0.2
D17S2194	58.94	0.83	0.7	0.51	0.3	0.12	0	0.83
D17S1867	66.96	-11.79	-3.1	-1.3	-0.51	-0.14	0.5	0.02
D17S1299	71.13	0	0	0	0	0	0.1	0
D17S1868	76.3	-13.85	-3.13	-1.56	-0.74	-0.26	0.5	0.08
D17S1795	77.3	-5.8	-3.12	-1.55	-0.74	-0.26	0.5	0.08
D17S957	89.64	-8.1	-1.18	-0.48	-0.16	-0.02	0.468	0
D17S794	94.36	-6.14	-2.18	-1.12	-0.55	-0.21	0.5	0.11
D17S944	95.11	-5.26	-2.25	-1.17	-0.59	-0.23	0.5	0.14
D17S940	102.6	-7.97	-2.15	-0.95	-0.37	-0.09	0.5	0
D17S1351	108.83	-5.93	-1.22	-0.35	-0.01	0.09	0.399	0.09
D17S1603	117.82	-14.77	-2.52	-0.95	-0.26	0.01	0.444	0.04
D17S1847	126.71	-9.67	-1.69	-0.61	-0.14	0.04	0.431	0.04
D17S836	128.03	0.12	0.1	0.08	0.05	0.03	0	0.12
D17S784	132.74	-8.08	-1	-0.21	0.08	0.13	0.37	0.13

Table 7.21. Two-point lod scores for chromosome 17 in AE family

Locus	Genetic map	Recombinaton fraction					Theta max	Zmax
		0	0.1	0.2	0.3	0.4		
D18S1105	3.97	-4.65	0.64	0.83	0.72	0.44	0.202	0.83
D18S63	9.84	-10.16	-0.73	0.13	0.34	0.26	0.314	0.35
D18S967	21.74	-13.91	-0.85	-0.02	0.21	0.18	0.334	0.22
D18S464	33.37	-6.12	-1.39	-0.58	-0.23	-0.07	0.5	0
D18S1107	47.29	-12.6	-3	-1.21	-0.43	-0.09	0.5	0
D18S877	54.03	-3.65	0.81	0.66	0.36	0.09	0.098	0.81
D18S1102	59.89	-14.39	-4.85	-2.44	-1.2	-0.46	0.5	0.36
D18S474	68.99	-12.12	-3.49	-1.85	-0.95	-0.38	0.5	0.32
D18S450	70.78	-17.38	-4.15	-1.99	-0.93	-0.34	0.5	0.2
D18S64	77.54	-12.33	-3.84	-1.96	-0.95	-0.34	0.5	0.13
D18S1134	87.2	-7.4	-3.46	-1.59	-0.71	-0.25	0.5	0.1
D18S1147	88.5	-13.68	-4.62	-2.31	-1.1	-0.39	0.5	0.13
D18S465	92.91	-13.86	-4.43	-2.23	-1.11	-0.43	0.5	0.32
D18S469	106.38	-8.7	-2.02	-0.86	-0.32	-0.06	0.485	0
D18S554	117.94	-12.03	-3.2	-1.39	-0.56	-0.17	0.5	0.02

Table 7.22. Two-point lod scores for chromosome 18 in AE family

Locus	Genetic map	Recombination fraction					Theta max	Zmax
		0	0.1	0.2	0.3	0.4		
D19S886	0.9	-17.55	-1.81	-0.53	-0.01	0.13	0.399	0.13
D19S591	10.59	-14.39	-2	-0.52	0.06	0.19	0.388	0.19
D19S427	20.75	-17.55	-2.52	-0.98	-0.32	-0.04	0.473	0
D19S584	33.17	-12.95	-0.77	-0.04	0.19	0.18	0.344	0.2
D19S226	37.98	-19.03	-2.47	-1.05	-0.43	-0.14	0.5	0.04
D19S593	43.02	-7.15	-2.81	-1.38	-0.68	-0.27	0.5	0.22
D19S414	55.81	-7.59	0.05	0.46	0.48	0.31	0.254	0.49
D19S208	63.1	-14.58	-1.58	-0.61	-0.24	-0.09	0.5	0.26
D19S903	73.21	-22.89	-3	-1.21	-0.43	-0.09	0.5	0
D19S867	84.6	-13.56	-3.12	-1.28	-0.44	-0.06	0.469	0.013
DG19S186	92.14	-11.42	-1.79	-0.83	-0.36	-0.11	0.5	0.015
D19S572	99.88	-11.39	-2.39	-1.28	-0.67	-0.27	0.5	0.31
D19S418	107.71	-14.4	-2.95	-1.12	-0.31	0.02	0.444	0.05
D19S605	108.45	-5.05	-0.7	-0.4	-0.22	-0.1	0.5	0.3
D19S573	113.23	0.06	0.05	0.04	0.03	0.01	0.001	0.06

Table 7.23. Two-point lod scores for chromosome 19 in AE family

Locus	Genetic map	Recombination fraction					Theta max	Zmax
		0	0.1	0.2	0.3	0.4		
D20S113	7.68	-5.11	-0.79	-0.31	-0.1	-0.01	0.463	0
D20S882	18.3	-12.03	-1.69	-0.66	-0.23	-0.05	0.5	0
D20S115	25.35	-12.99	-1.3	-0.3	0.07	0.14	0.376	0.14
D20S189	34.88	-0.01	-0.01	-0.01	0	0	0.471	0
D20S904	42.54	-13.18	-1.03	-0.22	0.06	0.11	0.375	0.11
D20S912	52.19	-14.16	-3.3	-1.41	-0.54	-0.13	0.5	0
D20S859	60.73	0.77	0.65	0.53	0.38	0.21	0	0.77
D20S107	63.19	-14.04	-2.92	-1.29	-0.53	-0.15	0.5	0
D20S108	65.64	-17.55	-2.96	-1.12	-0.31	0.02	0.44	0.05
D20S481	70.97	-13.86	-2.96	-1.12	-0.31	0.02	0.44	0.05
D20S838	71.23	-8.77	-1.88	-0.76	-0.26	-0.03	0.469	0
D20S178	75.47	-8.01	-0.58	0.02	0.2	0.18	0.337	0.21
D20S1083	83.33	-6.86	-0.27	0.06	0.14	0.1	0.311	0.14
D20S100	90.85	-12.59	-0.65	0.21	0.42	0.32	0.312	0.42
D20S171	100.84	-8.66	-1.36	-0.51	-0.17	-0.03	0.484	0

Table 7.24. Two-point lod scores for chromosome 20 in AE family

Locus	Genetic map	Recombination fraction					Theta max	Zmax
		0	0.1	0.2	0.3	0.4		
D21S1904	7.23	0	0	0	0	0	0.1	0
D21S1432	8.75	-5.06	1.15	1.27	0.99	0.52	0.168	1.29
D21S1899	16.2	-12.75	-2.56	-1.02	-0.37	-0.11	0.5	0.08
D21S1905	17.29	-7.55	-1.15	-0.25	0.03	0.05	0.358	0.06
D21S1902	19.54	-5.05	-0.7	-0.4	-0.22	-0.1	0.5	0.03
D21S272	23.96	0	0	0	0	0	0.1	0
D21S1442	29.5	-17.85	-3.26	-1.41	-0.58	-0.17	0.5	0.03
D21S1909	33.81	-13.69	-3.43	-1.78	-0.89	-0.34	0.5	0.22
D21S1898	38.62	-13.79	-3.26	-1.64	-0.8	-0.29	0.5	0.12
D21S1895	41.37	-13.6	-2.48	-1.18	-0.53	-0.17	0.5	0.02
D21S1252	44.04	-13.85	-3.19	-1.59	-0.76	-0.27	0.5	0.09
D21S1919	45.79	-13.85	-3.18	-1.58	-0.75	-0.27	0.5	0.09
D21S270	46.25	-13.17	-3.18	-1.58	-0.75	-0.27	0.5	0.09
D21S1255	48.28	0.3	0.25	0.2	0.14	0.08	0	0.31
D21S266	57.33	-8.16	-1.26	-0.53	-0.19	-0.03	0.477	0

Table 7.25. Two-point lod scores for chromosome 21 in AE family

Locus	Genetic map	Recombination fraction					Theta max	Zmax
		0	0.1	0.2	0.3	0.4		
D22S420	5.72	-14.35	-0.45	0.18	0.28	0.16	0.282	0.28
D22S427	6	0.29	0.25	0.2	0.14	0.08	0	0.29
D22S539	15.82	-10.46	-2.61	-1.01	-0.32	-0.03	0.461	0.02
D22S1174	20.03	-9.23	-2.3	-0.91	-0.31	-0.06	0.5	0
D22S315	23.91	-8.44	-3.61	-1.63	-0.66	-0.17	0.5	0
D22S1154	25.73	-13.68	-4.11	-1.9	-0.81	-0.23	0.5	0.01
D22S531	36.2	-6.68	-1.01	-0.5	-0.24	-0.09	0.5	0.04
D22S1265	40.84	-10.01	-2.83	-1.34	-0.59	-0.18	0.5	0.02
D22S276	51.21	-12.99	-2.18	-0.96	-0.37	-0.09	0.497	0
D22S928	59.39	-17.8	-1.48	-0.31	0.09	0.15	0.372	0.16
D22S1170	68.42	-4.68	0.02	0.18	0.18	0.12	0.251	0.19

Table 7.26. Two-point lod scores for chromosome 22 in AE family

Appendix 4.

Papers published

Here I am annexin three published papers in addition to a manuscript submitted to Am

J Med Genet.

The paper "A Novel Cys212Tyr founder in parkin and allelic heterogeneity of juvenile

parkinsonism in a population from North West Colombia" was not part of my PhD but

it is included as it was basic to the chapter on *PARK2* and early onset Parkinson

disease.

ABSTRACT

Background A parkin Cys212Tyr substitution due to a G736A mutation has been observed in two early onset PD families from the province of Antioquia in North West Colombia and (independently) in one family from Spain. Microsatellite haplotype analysis indicates that point mutations of *parkin* represent unique events, suggesting that sharing of these mutations across populations is due to migration or common ancestry.

Objective To assess the extent of the founder effect for the G736A mutation in Antioquia and the possible introduction of this mutation through migration from Spain.

Methods We ascertained 7 additional Antioquian early onset PD cases (from 5 nuclear families) and screened for *parkin* mutations with a restriction assay recognizing the G736A substitution, SSCP and exon dosage. We carried out a haplotype analysis using 11 microsatellites in the *parkin* region and estimated the time since introduction of the G736A mutation in Antioquia based on the frequency of recombinant haplotypes.

Results Genealogical investigations allowed the identification of recent connections between 6 of the 7 Colombian families examined. The 7 new juvenile PD cases identified carry the G736A mutation, 4 in homozygous form. The three heterozygous patients also carry a duplication of exon 3. All Colombian patients with the *parkin* G736A mutation share a core haplotype for markers <1cM away from the mutation. The estimated age for the G736A substitution in Antioquia is 15 generations. In the Spanish family, this mutation is present in a different microsatellite haplotype.

Conclusions An extensive founder effect for the G736A mutation has been documented in Antioquia. So far 13 carriers of the mutation (10 homozygotes) from 7 nuclear families have been confirmed. A second *parkin* mutation (a duplication of exon 3), is present in the three G736A heterozygote carriers. Haplotype analysis excludes a recent relationship between the Colombian and Spanish patients studied and is consistent with the introduction of this mutation in early colonial times.

Introduction

We previously reported the finding of a parkin Cys212Tyr substitution in JPD patients from Antioquia in North-West Colombia (Pineda-Trujillo et al. 2001). This aminoacid substitution is caused by a G to A transition at position 736 (G736A) of the *parkin* cDNA. Patients homozygous for this mutation were observed in two unrelated Antioquian families, pointing to the occurrence of a founder effect in this region. This same G736A mutation was subsequently reported in one Spanish family (Hoenicka et al. 2002). Since the population of Antioquia is mostly of Spanish ancestry (Carvajal-Carmona et al. 2003) and as previous surveys suggest that point mutations of *parkin* are unique events (Periquet et al. 2001), the identification of the G736A mutation in Spain suggests that this mutation could have been taken to Antioquia by Spanish migrants. This could have resulted in a founder effect due to the genetic isolation of Antioquia.

Here we report the molecular characterization of 7 additional Antioquian early onset PD cases (from 5 nuclear families) and of the single Spanish family available with the *parkin* G736A mutation. An extensive founder effect in Antioquia has been evidenced with a total of 14 carriers of the C736A mutation identified so far. The analysis of all available cases with the G736A substitution is consistent with the introduction of this mutation early in colonial times. A nuclear family with an apparent dominant transmission of early onset PD was identified but this was explained by the segregation of a second mutation (a duplication of exon 3) in this isolated population.

Methods

Patients

Patients presented with parkin sonism (bradikynesia, rigidity and resting tremor) before

45 years of age and had a good response to levodopa. In cases not available for

examination, affection status was inferred from a review of clinical records or family

interview. Diagnosis was confirmed by neuropathological analysis of one case (a detailed

description of clinical and pathological findings of the patients studied is given in

Moreno et al submitted).

Mutation screening

The G736A mutation was tested using a mismatched primer introducing a site for

restriction enzyme NdeI 5'TCTTTCAGGAATTTTTCTTTA**C**AT3' (Pineda-Trujillo et

al, 2001). The reverse primer was Exon 6R as described (Kitada et al. 1998). PCRs were

performed in a final volume of 25 microlitters with 30-100 ng/µL of genomic DNA, 0.33

µM of each primer, 1X of Qiagen HotStarTaq Master Mix Kit. PCR conditions included

94o 15' of initial denaturing followed by 10 cycles of 94o15", 48o15" and 72o30"; then

20 cycles of 89o15", 48o15" and 72o30", followed by a final elongation step of 72o10'.

Restriction digestion was performed overnight using 5-10 units per sample according to

manufacturer specifications (New England Biolabs). Restriction fragments were run in a

3.5% agarose ethidium-bromide gel. For single strand conformational polymorphism

(SSCP) screening all exons were amplified as reported by Kitada et al. 1998 and PCR

products run in 10% non-denaturing polyacrylamide gels followed by standard silver

staining. Exon dosage was carried out as described in(West et al. 2002).

Microsatellite marker typing

Eleven microsatellite markers (D6S255, D6S437, D6S1581, D6S1579, D6S1550,

D6S253, D6S305, D6S980, D6S1599, D6S1277, D6S386) were typed in the parkin gene

region (Figure 1). These were amplified in a final volume of 10 μL using 30-100 ng/□L

of genomic DNA, 0.33μM of each primer and 1x of Qiagen HotStarTaq Master Mix Kit.

PCR conditions were as follows: 94o15' followed by 5 cycles of 94o30" 58o30" and

72o30"; then 26 cycles of 89o15"56o15"72o30" and a final extension at 72o for 10 min.

Allelic sizes were determined using Genescan 2.1 and Genotyper 3.5 software (Applied

Biosystems).

Data analysis

Haplotypes were reconstructed using the Simwalk2 program (Sobel and Lange 1996).

The age since the introduction of the mutation was estimated based on the frequency of

recombinants (Reich and Goldstein 1999). According to this method, $G = -\ln(P)/M$ where

G is time in generations since the introduction of the mutation, P the proportion of non-

recombinant chromosomes (taking the most frequent haplotype as the ancestral) and M

the genetic distance between marker and mutation. The recombination rate on both sides

of the G736A mutation varies greatly (Figure 1). In the region 3' from the mutation (in

the sense of transcription) current genetic maps do not show evidence of recombination

between markers D6S1579 and D6S305, covering an interval of about 1.2Mb. Further

downstream, recombination rate is about average for the genome (~1Mb/cM). The

interval between markers D6S305 and D6S980, which overlaps the mutation, has an

estimated genetic distance of ~1.4 cM for a physical distance of ~461,336 bp

(~329,526bp/cM). The rate of recombination increases sharply between markers D6S980

and D6S1599 (upstream of the mutation), which are 182,972bp apart but separated by a

distance of 6.7cM on the genetic map (~27,309bp/cM).

Results

Genealogical investigations identified some common ancestors between 6 of the 7

Antioquian early onset PD families (Figure 2). In one of the nuclear families there

appears to be vertical transmission of PD as individual VI:2 is affected and her mother

(V:9) and two sibs of the mother (V:7 and V:8) are also affected (Figure 2).

Mutation screening identified the G736A mutation in all available affected

individuals (Figure 2). In all but three patients (V:1, V:12 and VI:2), the mutation is

present in homozygous form. One of these patients (VI:2) is a member of the family

showing vertical transmission of early onset PD. Carrier status for the G736A mutation

was confirmed in unaffected members of several families.

Microsatellite analysis identified a core haplotype consisting of alleles 3, 3, 1 at

markers D6S1550, D6S253 and D6S305 respectively (Figures 1 and 3). This core

haplotype is shared by all Antioquian PD patients, including individuals from the single

family with no documented connection to the other pedigrees (Figures 3 and 4). The

G736A mutation in the Spanish PD family is present in a microsatelllite haplotype that

differs from that observed in Antioquia at all the markers examined (Figure 4).

The three patients heterozygous for the G736A mutation (V:1 V:12 and VI:2)

share the microsatellite haplotype associated with this mutation and a second haplotype

characterized by alleles 5, 2, 3, 2, 4 at markers D6S253, D6S305, D6S980, D6S1599 and

D6S1277 (Figure 2). An unaffected individual (VI:4), currently aged 20, also carries

these two haplotypes. SSCP and sequencing of *parkin* exons did not identify other point

mutations but exon dosage detected an exon 3 duplication in these individuals.

The two microsatellite flanking markers nearest to the G736A mutation that show

evidence of recombination with the mutation are D6S1579 and D6S980 (Figures 1 and

3). Across the most recent generation of all the families examined, 7 recombinant

chromosomes are observed for marker D6S980 and 5 for D6S1579 (Figure 3). To

account for the large variation in rates of recombination of the *parkin* region (see

methods), we assumed a recombination rate of 1,000,000bp/cM (the genome average)

between the G736A mutation and marker D6S1579 and of 178,418bp/cM (the average

for the region D6S305 to D6S1599) between the mutation and marker D6S980 (figure 1).

The resulting estimates for the time since introduction of the mutation in Antioquia are 15

and 17 generations based on the data for markers D6S1579 and D6S980, respectively.

Comment

Our findings document the occurrence of an extensive founder effect for the

parkin G736A mutation in Antioquia. Fourteen individuals from seven nuclear families

carrying this mutation have now been identified in the region. Previous studies have

indicated the presence of PD-related symptomatology in some heterozygous carriers of

parkin mutations (Foroud et al. 2003; Khan et al. 2005). Interestingly, the father of the

Spanish family shown in figure 4 is heterozygous for the G736A mutation and presented

progressive supranuclear palsy (Morales et al. 2002). Three of the patients identified in

Antioquia are heterozygous for the G736A mutation, including one case in a family

showing vertical transmission of early onset PD. However, the analyses carried out here confirmed the presence of a second mutant haplotype of *parkin* (carrying an exon 3 duplication) in all heterozygous cases, thus excluding the possibility that heterozygosity of the G736A mutation is responsible for PD in these cases.

Although genealogical connections between six of the seven nuclear families have been documented, including ancestors common to several families 4-5 generations ago (Figure 2), no couple ancestral to all living affecteds (thus including the founder carrier of the mutation) was identified. These observations imply that the founding individual (the initial carrier of the G736A mutation) lived further back in time. Consistent with this scenario, the age estimate obtained for the mutation from the frequency of recombinant haplotypes suggests that it was introduced in Antioquia about 16 generations ago. Assuming a generation time of 25 years, this would place the appearance of the mutation in Antioquia some 400 years ago, or early in colonial times. This mutation could have been introduced because of migration from Spain or recurrent mutation in Antioquia.

The observation that the Spanish family with the G736A substitution carries a mutant microsatellite haplotype different from the one found in Antioquia, demonstrates that there is no recent connection between the families examined in both countries. This could indicate that the G736A substitution in Antioquia represents a recurrent mutation. However, only one Spanish chromosome was available for study and consequently it is not possible to evaluate the degree of association in Spain of the G736A mutation with alleles at the markers typed (as done for Antioquia). It is possible that the G736A mutation is of considerable age in Europe and is present in Spain on various haplotypic backgrounds, one of which is the one observed in Antioquia.

A previous survey using mostly the same microsatellite markers examined here demonstrated high disequilibrium for several point mutations in *parkin* but also documented their occurrence on various haplotypic backgrounds (Periquet et al. 2001). This observation points to a single origin for these mutations and also suggests that some of them are of considerable age. It therefore appears more likely that the G736A mutation was introduced to Antioquia by migration from Spain early in colonial times. The age estimate obtained for the mutation in Antioquia is also consistent with historical information documenting that the period of highest migration from Spain to the Americas was during the 16th century (Burkholder and Johnson 2003).

The early onset PD cases identified here represent the largest set of patients identified so far carrying an identical parkin mutation. The relative homogeneity of this patient group (particularly since they come from the same population) should facilitate the investigation of modifying factors influencing disease course and severity. Such studies should contribute to further our understanding of the pathogenesis of PD.

References

1. Burkholder MA, Johnson LL. 2003. *Colonial Latin America.*Oxford University Press.

2. Carvajal-Carmona LG, Ophoff R, Service S, Hartiala J, Molina J, Leon P, Ospina J, Bedoya G, Freimer N, and Ruiz-Linares A (2003) Genetic demography of Antioquia (Colombia) and the Central Valley of Costa Rica. *Hum.Genet.*

3. Foroud T, Uniacke SK, Liu L, Pankratz N, Rudolph A, Halter C, Shults C, Marder K, Conneally PM, and Nichols WC (2003) Heterozygosity for a mutation in the parkin gene leads to later onset Parkinson disease. *Neurology* 60 (5):796-801.

4. Hoenicka J, Vidal L, Morales B, Ampuero I, Jimenez-Jimenez FJ, Berciano J, del Ser T, Jimenez A, Ruiz PG, and de Yebenes JG (2002) Molecular findings in familial Parkinson disease in Spain. *Arch.Neurol.* 59 (6):966-970.

5. Khan NL, Horta W, Eunson L, Graham E, Johnson JO, Chang S, Davis M, Singleton A, Wood NW, and Lees AJ (2005) Parkin disease in a Brazilian kindred: Manifesting heterozygotes and clinical follow-up over 10 years. *Mov Disord.* 20 (4):479-484.

6. Morales B, Martinez A, Gonzalo I, Vidal L, Ros R, Gomez-Tortosa E, Rabano A, Ampuero I, Sanchez M, Hoenicka J, and Garcia DY (2002) Steele-Richardson-Olszewski syndrome in a patient with a single C212Y mutation in the parkin protein
1. *Mov Disord.* 17 (6):1374-1380.

7. Periquet M, Lucking C, Vaughan J, Bonifati V, Durr A, De Michele G, Horstink M, Farrer M, Illarioshkin SN, Pollak P, Borg M, Brefel-Courbon C, Denefle P, Meco G, Gasser T, Breteler MM, Wood N, Agid Y, and Brice A (2001) Origin of the mutations in the parkin gene in Europe: exon rearrangements are independent recurrent events, whereas point mutations may result from Founder effects. *Am.J.Hum.Genet.* 68 (3):617-626.

8. Pineda-Trujillo N, Carvajal-Carmona LG, Buritica O, Moreno S, Uribe C, Pineda D, Toro M, Garcia F, Arias W, Bedoya G, Lopera F, and Ruiz-Linares A (2001) A novel Cys212Tyr founder mutation in parkin and allelic heterogeneity of juvenile Parkinsonism in a population from North West Colombia. *Neurosci.Lett.* 298 (2):87-90.

9. Sobel E and Lange K (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am.J.Hum.Genet.* 58 (6):1323-1337.

10. West A, Periquet M, Lincoln S, Lucking CB, Nicholl D, Bonifati V, Rawal N, Gasser T, Lohmann E, Deleuze JF, Maraganore D, Levey A, Wood N, Durr A, Hardy J, Brice A, and Farrer M (2002) Complex relationship between Parkin mutations and Parkinson disease. *Am.J.Med.Genet.* 114 (5):584-591.
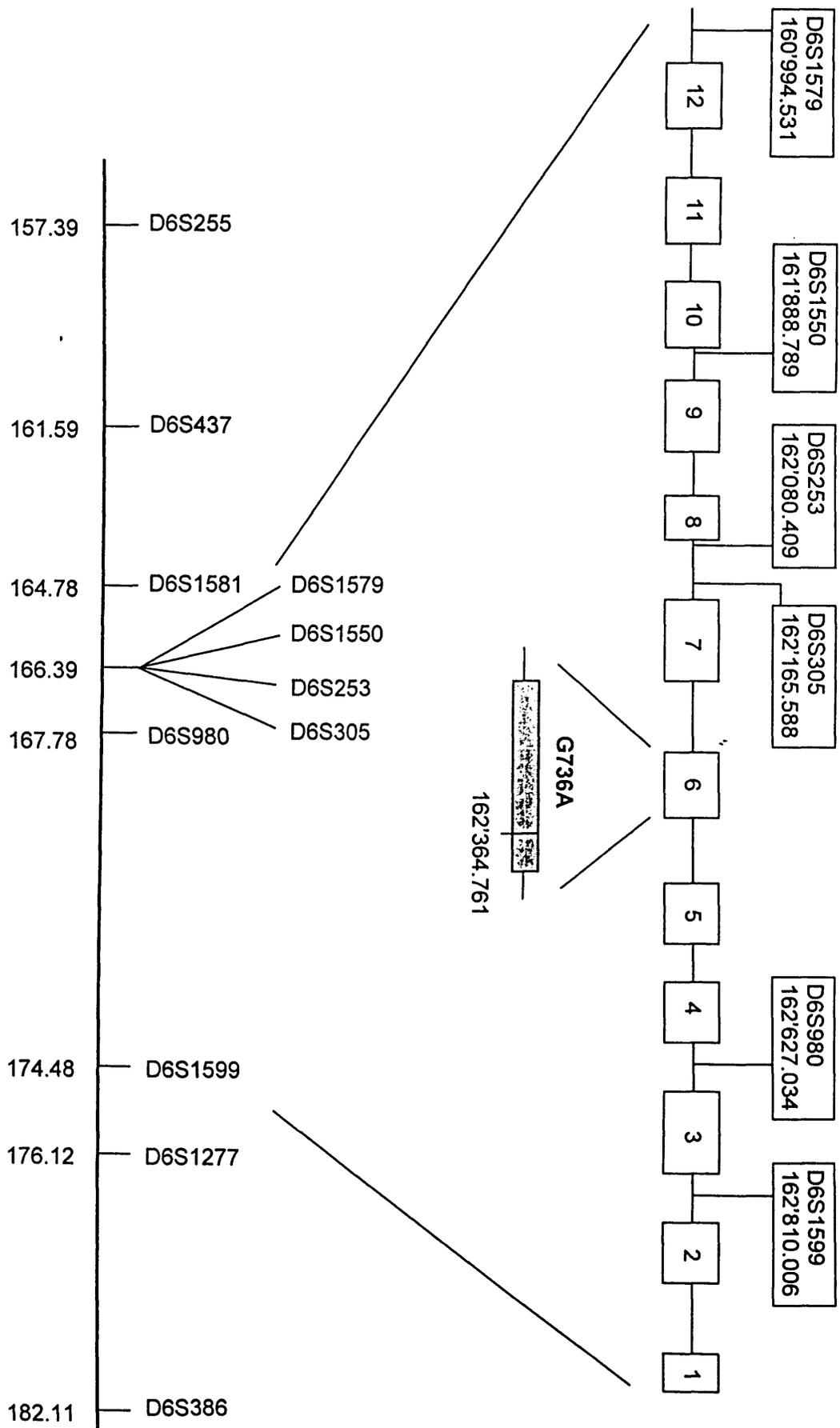
Legends to Figures

Figure 1. Map of the parkin gene region showing the location of the G736A mutation
and the microsatellite markers examined. Nucleotide positions (top) were obtained
from build 35 of the human genome sequence. Genetic distances between markers
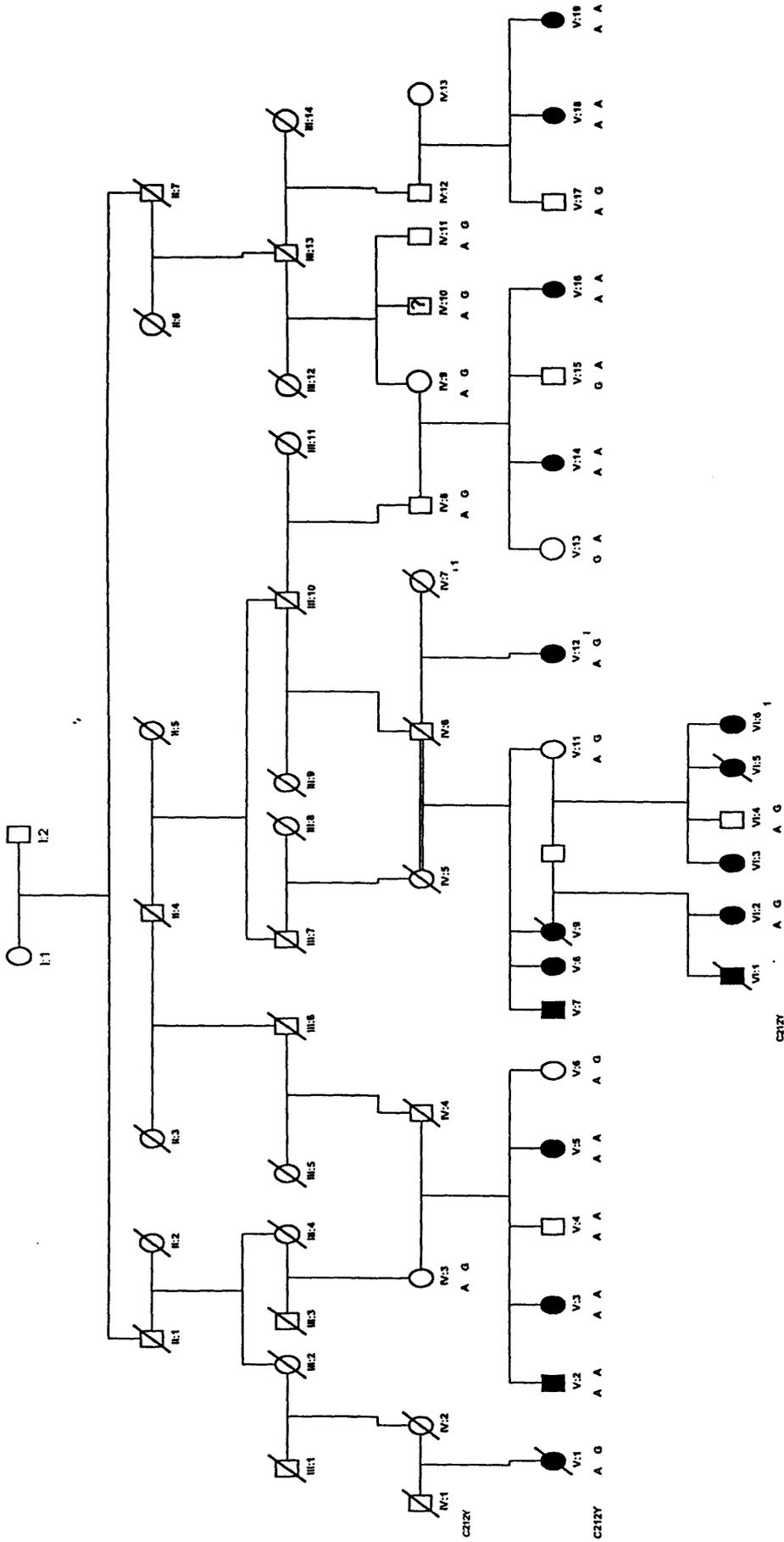are shown at the bottom.

Figure 2. Documented genealogical connections between Antioquian families carrying
the G736A mutation of *parkin*.

Figure 3. Microsatellite haplotypes in the *parkin* gene region for Antioquian individuals
carrying the G736A mutation. The core haplotype shared by all carriers of the
mutation is shaded. Individual labels are as in Figures 2 and 4. PJF1 refers to the
pedigree shown in Figure 4 (A). All other individuals correspond to members of the
pedigree shown in Figure 2.

Figure 4. Microsatellite haplotypes of the *parkin* gene region in (A) a family from
Antioquia (PJF1) with no known relationship with those shown in Figure 2. (B) The
only Spanish family identified so far with the G736A mutation. This family also
segregates a mutation. Haplotype segments shared between family members have
been color-coded. Allele numbers are as in Figure 3.

157.39 — D6S255

161.59 — D6S437

164.78 — D6S1581 — D6S1579

166.39 — D6S1550
— D6S253
— D6S305

167.78 — D6S980

174.48 — D6S1599

176.12 — D6S1277

182.11 — D6S386

G736A
162'364.761

12
11
10 — D6S1550 161'888.789
9
8 — D6S253 162'080.409
— D6S305 162'165.588
7
6
5
4 — D6S980 162'627.034
3
2 — D6S1599 162'810.006
1

D6S1579 160'994.531

marker\ID	V:2*		V:3*		V:4		V:5*		V:6		V:12*		V:13		V:14*		V:15		V:16*		V:17		V:18*		V:19*		VI:2*		VI:4		PJF1 II:2*		II:3*		II:4*	
D6S255	4	3	4	3	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	4	3	3	3	3	3	3	3	3	2	2	2	2	2	2
D6S437	2	1	2	1	2	1	2	1	1	1	1	3	1	1	1	1	8	1	1	1	1	2	1	1	1	1	1	5	1	5	1	1	1	1	1	1
D6S1581	3	4	3	4	3	4	3	4	0	4	3	5	5	3	3	3	2	3	3	3	3	1	3	3	3	3	3	1	2	1	0	0	3	3	3	3
D6S1579	1	5	1	5	1	5	1	5	3	5	5	1	1	5	5	5	1	5	5	5	5	1	5	5	5	5	5	1	5	1	1	5	5	5	5	5
D6S1550	3	3	3	3	3	3	3	3	1	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	3	4	3	3	3	3	3	3
D6S253	3	3	3	3	3	3	3	3	1	3	3	5	2	3	3	3	4	3	3	3	3	1	3	3	3	3	3	5	3	5	3	3	3	3	3	3
D6S305	1	1	1	1	1	1	1	1	8	1	1	2	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	1	1	1	1	1	1
D6S980	10	11	10	11	10	11	10	11	3	11	10	3	7	10	10	10	4	10	10	10	10	10	10	10	10	10	10	3	10	3	10	10	10	10	10	10
D6S1599	1	1	1	1	1	1	1	1	2	1	1	2	2	1	1	1	2	1	1	1	1	2	1	1	1	1	1	2	1	2	1	1	1	1	1	1
D6S1277	4	4	3	4	4	4	4	4	3	4	4	4	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
D6S386	5	2	3	2	5	2	5	2	3	2	1	5	5	1	1	5	5	5	1	2	4	5	1	5	1	5	2	1	2	1	3	3	3	3	3	3

A)

D6S255	0.00		
D6S437	4.20		
D6S1581	7.40		
D6S1579	9.00		
D6S1550	9.90		
D6S253	10.10		
D6S305	10.20		
C212Y	10.30		
D6S980	10.40		
D6S1599	17.09		
D6S1277	18.73		
D6S386	24.72		

B)