# Genetic Analysis

# of

# Human Absence Epilepsy

Robert Anthony Robinson MA MBBS MRCP

A thesis submitted for the degree of Doctor of Philosophy

UCL

September 2009

## Declaration

I, Robert Anthony Robinson, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

The ascertainment of all families was performed by me with the assistance of Pauline Boland, Clinical Research Nurse. I phenotyped all UK patients using available clinical and neurophysiological data. Phenotyping of some patients from outside the UK was performed by local collaborators.

Blood samples and cheek swabs for DNA extraction were collected by me and Pauline Bolland. I performed all DNA extraction and quantitation except where DNA was sent directly from collaborators.

I carried out all genotyping of SSLPs in the laboratory including polyacrylamide gel electrophoresis and allele detection. Genotyping of SNPs was largely carried out at the MRC Geneservice and by Kbiosciences.

I performed some DNA sequencing on the ABI 373 sequencer at UCL. However the majority of candidate gene sequencing was performed by the Advanced Biotechnology Centre at Imperial College London. All sequence analysis was performed by me.

All linkage analysis, association analysis, SNP selection, haplotype construction and bioinformatics analysis of variants was performed by me.

The family with absence epilepsy and episodic ataxia was ascertained and phenotyped by Dr SL Jaffe, Louisiana State University School of Medicine, USA. I performed the DNA extraction, genotyping and linkage analysis. DNA sequencing of *CACNA1A* and functional analysis in this family was performed by Paola Imbrici at the Institute of Neurology, UCL.

Signed Date

Robert Anthony Robinson

## Abstract

Idiopathic Mendelian epilepsies have been typically identified as channelopathies. Evidence suggests that mutations in genes encoding $GABA_A$ receptors, $GABA_B$ receptors or voltage-dependent calcium channels (VDCCs) may underlie childhood absence epilepsy (CAE), an idiopathic generalised epilepsy with complex inheritance.

The aims of this project were:

i) Ascertainment of a patient resource
ii) Investigation of candidate genes by linkage analysis
iii) Mutation analysis by direct sequencing
iv) Construction of single nucleotide polymorphism (SNP) based haplotypes in candidate genes
v) Intra-familial association analysis using SNP based haplotypes

DNA and clinical data were obtained from: 53 nuclear CAE pedigrees; 29 families including individuals with CAE and a broader 'absence' epilepsy phenotype; 217 parent-child trios; a North American family in which absence epilepsy segregates with episodic ataxia type 2 (EA2)

Sixteen calcium channel genes and seven $GABA_A$ and two $GABA_B$ receptor subunit genes were excluded by linkage analysis. Significant linkage was demonstrated for *CACNG3* on chromosome 16p12-p13.1 for both CAE and the broader absence phenotype. Positive linkage was also obtained at the *GABRA5*, *GABRB3*, *GABRG3* cluster on chromosome 15q11-q13. Non-parametric linkage analysis was significant at both the 16p and 15q loci. Two-locus analysis supported a digenic effect from these two loci. Sequencing of *CACNG3* revealed 34 sequence variants, none clearly causal, although bioinformatic analysis provided supportive functional evidence. Association analysis showed significant transmission disequilibrium both for individual single nucleotide polymorphisms (SNPs) and SNP based haplotypes spanning *CACNG3*. This work has provided genetic evidence that *CACNG3* and at least one of the three $GABA_A$ receptor genes are susceptibility loci for absence epilepsy.

Linkage analysis performed in the family with absence epilepsy and EA2 was suggestive that the VDCC *CACNA1A* was the causative gene. This was

subsequently confirmed by sequence analysis in collaboration with the Institute of Neurology, UCL. This is the first reported family in which a *CACNA1A* mutation that impairs calcium channel function cosegregates with typical absence seizures and 3Hz spike-wave discharges on EEG.

**Contents**

**List of Figures**

**List of Tables**

## Acknowledgments

I would especially like to thank the following people:

Dr Barry Chioza, Dr Kate Everett, Dr Michele Rees and Dr Nichole Taske worked with me at various times on the genetic analysis of human absence epilepsy and tolerated my deficiencies in the laboratory with great patience.

Dr Jane Barclay, Dr Louise Bate, Dr Rahul Chodhari, Dr Eddie Chung, Dr Caroline Fertleman, Dr Maggie Meeks, Dr Hannah Mitchison, Dr Sara Mole and Mr Keith Parker worked within the same laboratory on other research projects but were always available for advice and support. Their encouragement, enthusiasm and friendship were indispensable.

I would like to thank Pauline Boland for her help with patient ascertainment

This work would not have been possible without the involvement of the many families with epilepsy and their referring physicians to whom I am very grateful.

My enthusiasm for and knowledge of paediatric epilepsy owes much to my clinical supervisors in paediatric neurology over the years. In particular I would like to thank Professor Brian Neville, Professor Helen Cross, Professor Robert Surtees, Professor Richard Robinson, Dr Sarah Benton and Dr Carlos de Sousa.

I am indebted to the Wellcome Trust, the Epilepsy Research Foundation and Action Medical Research for financially supporting me in this work.

Professor Mark Gardiner’s supervision, support, guidance, encouragement and knowledge have been invaluable from the conception to the completion of this research project.

Finally the support, patience and tolerance of my wife, Yeva, and my children Maya, Ruby and Jacob during the writing of this thesis has been limitless (well, almost). Their amazement at the completion of this thesis is almost as great as mine.

**List of Abbreviations**

| | |
|---|---|
| ABC | Advanced Biotechnology Centre |
| ACD | Acid citrate dextrose |
| AD | Autosomal dominant |
| ADNFLE | Autosomal dominant nocturnal frontal lobe epilepsy |
| ADPEA | Autosomal dominant partial epilepsy with auditory features |
| ADPEAF | Autosomal dominant partial epilepsy with auditory features |
| AE | Absence epilepsy (atypical CAE variants) |
| AEA | Absence epilepsy with ataxia |
| AED | Antiepileptic drug |
| AMPA | α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate |
| AR | Autosomal recessive |
| AS | Absence seizure |
| BAC | Bacterial artificial chromosome |
| BCECTS | Benign childhood epilepsy with centrotemporal spikes |
| BFIC | Benign familial infantile convulsions |
| BFNC | Benign familial neonatal convulsions |
| BFNIS | Benign familial neonatal infantile seizures |
| BPNA | British Paediatric Neurology Association |
| CA | Cerebellar ataxia |
| CAE | Childhood absence epilepsy |
| CAMR | Centre for Applied Microbiology and Research (Porton Down) |
| cDNA | Complementary DNA |
| CEDaR | Canadian Epilepsy Database and Registry |
| CEPH | Centre de'Etude du Polymorphism Humain |
| CNS | Central nervous system |
| DASH | Dynamic allele specific hybridisation |
| ddNTP | Dideoxynucleoside triphosphate |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxynucleoside triphosphate |
| dsDNA | Double stranded DNA |
| EA | Episodic ataxia |
| EA1 | Episodic ataxia type 1 |
| EA2 | Episodic ataxia type 2 |
| EAR | Epilepsy associated repeat |
| EBV | Epstein-Barr virus |
| ECACC | European Collection of Cell Cultures |
| EDTA | Ethylenediaminetetraacetic acid |
| EEG | Electroencephalogram |
| ELISA | Enzyme-linked immunosorbent assay |
| EMA | Epilepsy with myoclonic absences |
| ENCODE | Encyclopaedia of DNA Elements |
| EPSP | Excitatory postsynaptic potential |
| ESE | Exonic splicing enhancers |
| EST | Expressed sequence tag |
| FPEVF | Familial partial epilepsy with variable foci |
| FRET | Fluorescence resonance energy transfer |
| FS | Febrile seizure |
| GABA | Gamma-aminobutyric acid |

| | |
|---|---|
| GAERS | Genetic absence epilepsy rat from Strasbourg |
| GAS | Genetic Analysis System, © Alan Young 1993-1998 |
| GEFS+ | Generalised epilepsy with febrile seizures plus |
| GHB | Gamma-hydroxybutyric acid |
| giSNP | Genetically indistinguishable SNPs |
| GPSWD | Generalised polyspike-wave discharges |
| GSA | Generalised seizure, absence |
| GSAmy | Generalised seizure, absence + myoclonus |
| GSM | Generalised seizure, myoclonic |
| GSWD | Generalised spike-wave discharges |
| GTCS | Generalised tonic-clonic seizures |
| HGMP | Human Genome Mapping Project (UK) |
| HLOD | Heterogeneity LOD |
| htSNP | Haplotype tagging SNP |
| HUGO | Human Genome Organisation |
| HVA | High-voltage activated |
| IBD | Identical by descent |
| IBE | International Bureau for Epilepsy |
| IBS | Identical by state |
| ICCA | Infantile convulsions and paroxysmal choreoathetosis |
| ICEGTC | Intractable childhood epilepsy with generalized tonic clonic seizures |
| IGE | Idiopathic generalized epilepsy |
| ILAE | International League Against Epilepsy |
| JAE | Juvenile absence epilepsy |
| JME | Juvenile myoclonic epilepsy |
| LCP | Linkage control program |
| LD | Linkage disequilibrium |
| LOD | Logarithm of odds |
| LRP | Linkage reporting program |
| LRR | Leucine rich repeat |
| LVA | Low-voltage activated |
| MAF | Minor allele frequency |
| MALDI-TOF | Matrix-associated laser desorption time-of-flight mass spectrometry |
| McSNP | Melting curve SNP analysis |
| MDS | Multiple displacement amplification |
| MRC | Medical Research Council |
| mRNA | Messenger RNA |
| MS | Myoclonic seizure |
| nAChR | Neuronal nicotinic acetylcholine receptor |
| NCBI | National Center for Biotechnology Information |
| NMDA | N-methyl-D-aspartic acid |
| NPL | Non-parametric linkage |
| NRT | Nucleus reticularis thalami |
| OD | Optical density |
| PAC | P1-derived artificial chromosome |
| PCR | Polymerase chain reaction |
| PDT | Pedigree disequilibrium test |
| PET | Positron emission tomography |

| | |
|---|---|
| PIC | Polymorphism information content |
| PS | Photosensitive seizures |
| RCPCH | Royal College of Paediatrics and Child Health |
| RFLP | Restriction fragment length polymorphism |
| RNA | Ribonucleic acid |
| RRM | RNA-recognition motifs |
| RSP | Restriction site polymorphism |
| S/W EEG | Spike-wave EEG |
| SDS | Sodium dodecyl sulphate |
| SMEB | Severe myoclonic epilepsy of infancy – borderland |
| SMEI | Severe myoclonic epilepsy of infancy |
| SNP | Single nucleotide polymorphisms |
| SR | Serine/arginine-rich |
| SSCP | Single strand conformational polymorphism |
| SSLP | Simple sequence length polymorphism |
| TAMRA | Tetramethyl-6-Carboxyrhodamine |
| TBE | Tris/Borate/EDTA |
| TDT | Transmission disequilibrium test |
| TE | Tris/EDTA |
| TEMED | Tetramethylethylenediamine |
| Tm | Melting temperature |
| TSC | The SNP consortium |
| UCL | University College London |
| UCSC | University of California, Santa Cruz |
| US | Unclassified seizure |
| VDCC | Voltage-dependant calcium channel |
| WHO | World Health Organisation |

# PART 1 Introduction

# *Chapter 1 Epilepsy*

## 1.1 Historical review

The term “epilepsy” derives from the Greek επιλαμβανέιν (epilambanein), to be attacked, taken hold of or seized (Engel and Pedley 1997). However the earliest descriptions are found in Mesopotamian, Egyptian, Indian and Chinese literature. The British Museum houses a Babylonian tablet which gives a detailed account of epilepsy, including many different seizure types. The tablet is a chapter from a Babylonian textbook of medicine dating back to 2000BC. Each seizure type is associated with an evil god or spirit. An Assyrian text of 650BC describes a seizure in terms of demonic possession, referring to “eyes moving to the side”, “lip puckering”, “saliva flowing from the mouth”, and “hand, leg and trunk jerking like a slaughtered sheep”, and tells of circumstances in which the demon can or cannot be “driven out” (Porter 1997).

A less supernatural account of epilepsy appears in the Charaka Samhita, dated to 400BC, a description of the Ayurvedic medical system of Ancient India developed during the Vedic period (4500-1500BC). Referring to “apasmara”, meaning “loss of consciousness”, the Charaka Samhita describes the causes, symptoms, diagnosis and treatment of epilepsy.

The Hippocratic text On the Sacred Disease, which dates to 410BC, rejected the idea of a divine origin for epilepsy (Porter 1997). Instead, the symptoms are associated with phlegm blocking the airways, causing the body to convulse as it struggles to free itself. Galen, an influential Greek physician of the second century AD, dissected the brain. He believed that epilepsy was caused by phlegm and bile accumulating within the cerebral ventricles. However, supernatural views of epilepsy persisted in Europe until the 19th Century, when the concept of epilepsy as a brain disorder became more widely accepted. Suspicion, fear and mistrust of people with epilepsy were widespread, and discrimination and punishment was not unusual.

Bromide, the world's first effective anti-epileptic drug, was introduced in 1857. In 1873 Hughlings Jackson, a London neurologist proposed that seizures were the result of sudden brief electro-chemical discharges in the brain, and that the character of the seizures depended on the location and function of the site of the discharges. The first experimentally induced seizures recorded by electroencephalography were described by Pavel Yurevich Kaufmann, a student of Pavlov, in Russia in 1912. Spontaneous brain electrical activity in humans recorded from the scalp was reported in 1929 by Dr Hans Berger, a professor of psychiatry at the University of Jena in Germany. He was also the first person to record human spike-wave activity. Advances over the course of the 20th century have included many diagnostic imaging techniques, the development of a panoply of antiepileptic drugs and the use of neurosurgical therapies.

In the understanding of the basic processes of epileptogenesis, molecular genetics has assumed an increasingly important role, with the discovery of single gene disorders and structural gene defects. This is now one of the most promising and exciting areas in epilepsy research.

## 1.2 Definition of epileptic seizures and epilepsy

An epileptic seizure is a transient episode of abnormal cortical neuronal activity apparent to the patient or an observer. The abnormal cortical activity may be manifest as a motor, sensory, cognitive or psychic disturbance. A seizure is diagnosed on clinical criteria, although electrophysiological data may be used in determining the precise seizure type.

Epilepsy is a disorder of the brain characterised by recurrent, unprovoked epileptic seizures. Therefore single epileptic seizures, provoked seizures (e.g. febrile convulsions, hypoglycaemic seizures or seizures occurring during drug withdrawal) or seizures that occur during an acute illness (e.g. encephalitis) are not classified as epilepsy. An epilepsy syndrome is a complex of signs and symptoms that define a unique epilepsy condition. It is

characterised by a particular pattern of seizure types occurring with certain physical and investigative findings. Epilepsy syndromes are generally defined in terms of aetiology, seizure type, age of onset, precipitating factors, EEG findings, and prognosis. Epilepsy also occurs as a feature of several neurological and multi-system diseases (e.g. neurocutaneous syndromes, neurodegenerative disorders).

## 1.3 Importance

Epilepsy is one of the most common serious brain disorders affecting up to 60 million people worldwide at any one time. It occurs most frequently in childhood, adolescence and old age. The 1998 'European Declaration on Epilepsy', estimated that epilepsy costs the countries of Europe over 20 billion euros annually, and a WHO/World Bank study in 1990, ‘Investing in Health’, showed that epilepsy accounted for nearly 1% of the world’s economic disease burden. In 1997, the World Health Organisation (WHO) together with two non-governmental agencies, the International League Against Epilepsy (ILAE) and International Bureau for Epilepsy (IBE), launched a Global Campaign against Epilepsy. The first three years focussed on increasing awareness and education, and the second phase was launched in 2001 as part of the World Health Day theme, Mental Health and Brain Disorders. In the European White Paper on Epilepsy, presented to the European Parliament in March 2001, the Call to Action focused on three areas: eradication of stigma, discrimination in the workplace and inadequate research. In 2001, the WHO Global Burden of Disease study reported that, for the year 2000, epilepsy accounted for 0.5% of the whole burden of diseases in the world, using the measure ‘years of life lived with disability’ (Leonardi and Ustun 2002).

## 1.4 Classification

Since epilepsy was first recognised there have been descriptions of different seizure types. However, the development of the first structured classification of seizures began in the 1960's, and was introduced by the ILAE in 1981 (ILAE 1981). A classification of different epilepsies followed in 1985 (ILAE 1985), and was revised in 1989 and 2001 (ILAE 1989; Engel 2001). A proposed classification at that time was divided into five parts, or axes, which facilitate a logical approach to diagnosis and management of individual patients.

### 1.4.1 Axis 1: Ictal phenomenology

Axis 1 consists of a description of the ictal semiology, without reference to aetiology, anatomy or mechanisms. A standardised Glossary of Descriptive Terminology should be used to aid communication among clinicians and researchers (Blume, Luders et al. 2001). This descriptive terminology recognises nine possible components of seizure semiology: i) Motor (elementary motor and automatism); ii) Non-motor (aura, sensory and dyscognitive); iii) Autonomic; iv) Somatotopic modifiers (laterality, body part involved and centricity); v) Modifiers and descriptors of seizure timing (incidence, state dependent and catamenial); vi) Duration; vii) Severity; viii) Prodrome; ix) Postictal phenomena.

### 1.4.2 Axis 2: Seizure type

Axis 2 is the epileptic seizure type or types experienced by the patient. An accepted list of seizure types represents diagnostic entities with aetiologic, therapeutic and prognostic implications. These are divided into self-limited seizures and continuous seizures (status epilepticus), and further divided into generalised and focal seizures (see below). Axis 2 also includes precipitating factors for reflex seizures.

### *i. Generalised seizures*

Generalised seizures are those whose initial semiology and/or ictal EEG patterns indicate, or are consistent with, more than minimal involvement of both cerebral hemispheres. Consciousness is not always impaired. Generalised seizures include tonic-clonic, clonic, tonic, absence (typical, atypical and myoclonic), myoclonic and atonic seizures.

### *ii. Focal seizures*

Focal seizures are those whose initial semiology and/or ictal EEG patterns indicate, or are consistent with, initial activation of only part of one cerebral hemisphere. Although it is recommended that the term ‘focal’ replaces ‘partial’ or ‘localization-related’, this does not mean that the epileptogenic region is a small, well-delineated focus of neuronal pathology. Focal seizures usually result from diffuse, sometimes widespread areas of cerebral dysfunction. The term ‘simple’ is used to describe focal seizures during which consciousness is unimpaired, and ‘complex’ signifies some degree of altered consciousness.

### *iii. Generalised or focal?*

In practice there are a range of conditions between focal and generalised epileptogenic dysfunctions including diffuse hemispheric abnormalities, multifocal abnormalities, and bilaterally symmetrical localised abnormalities. Any focal seizure can become secondarily generalised. In fact, apparently generalised seizures may be shown to have a focal origin when the EEG is examined in detail, indicating that a dichotomous classification of ‘focal’ versus ‘generalised’ is not always appropriate.

## 1.4.3 Axis 3: Epilepsy syndrome

Axis 3 is the epilepsy syndrome diagnosis, which is not always possible to make. The recommended list of syndromes is constantly evolving and includes syndromes in development. There is a distinction between epilepsy syndromes and conditions with epileptic seizures that do not require a

diagnosis of epilepsy (e.g. febrile seizures). The terms ‘focal’ and ‘generalised’ are also applied to epilepsy syndromes, reflecting the seizure type. The main epilepsy syndromes include idiopathic focal epilepsies of infancy and childhood, familial focal epilepsies, symptomatic focal epilepsies, idiopathic generalised epilepsies, reflex epilepsies, epileptic encephalopathies and progressive myoclonus epilepsies.

### 1.4.4 Axis 4: Aetiology

Axis 4 specifies the aetiology when it is known. Epilepsy syndromes are classified as symptomatic, probably symptomatic or idiopathic, according to whether a cause is apparent or not.

#### *i. Idiopathic epilepsy*

An idiopathic epilepsy is defined as a syndrome that is epilepsy alone, with no underlying structural brain lesion or other neurological signs or symptoms. These are known or presumed to have a genetic basis and are usually age dependent. It is estimated that 40- 60% of all epilepsies may be classified as idiopathic (Annegers, Rocca et al. 1996). The majority of idiopathic epilepsies show a complex pattern of inheritance.

#### *ii. Symptomatic epilepsy*

A symptomatic epilepsy is defined as a syndrome in which the epileptic seizures are the result of one or more identifiable structural lesion of the brain. The aetiology may be a disease frequently associated with epilepsy, a genetic disorder, a specific pathological substrate, a cerebral malformation or tumour, or an environmental insult.

#### *iii. Probably symptomatic epilepsy*

The term ‘cryptogenic’ has been replaced by ‘probably symptomatic’, used to define a syndrome that is believed to be symptomatic, but no aetiology has been identified.

### 1.4.5 Axis 5: Impairment

Axis 5 is an optional designation of the degree of impairment caused by the epilepsy. A World Health Organisation International Classification of Functioning and Disability is in preparation, and may be modified for application to epileptic disorders.

### 1.4.6 Current proposals on classification

A draft recommendation regarding the terminology and concepts used in the classification of the epilepsies and epileptic seizures was due for discussion by the ILAE Commission on Classification and Terminology at the International Epilepsy Congress in Budapest, on June 30th 2009 (ILAE 2009). The emphasis of the draft report is that the epilepsies should be flexibly organized according to whichever dimensions are most relevant to a specific purpose. The key recommendations related to the following three areas (Table 1-1):

#### *i. Mode of seizure onset and classification of seizures*

Generalised epileptic seizures originate within and involve bilaterally distributed networks but may be asymmetric and appear localised (not consistently). Generalised seizure types can be distinguished although there may be some overlap.

Focal seizures originate from unihemispheric networks, with consistent ictal onset for each seizure type. They are not classifiable as distinct entities as current understanding of mechanisms does not define a set of natural classes.

#### *ii. Syndrome versus epilepsy*

The term “syndrome” should be restricted to those conditions reliably identified by a cluster of electro-clinical characteristics. Other epilepsies can be grouped and organised in terms of other factors, such as cause (where known), seizure type, age of onset etc. The description of syndromes as being ‘localization-related’ versus ‘generalised’, and ‘idiopathic’,

'symptomatic' or 'cryptogenic' should be abandoned for a more flexible, multidimensional system of organisation.

### *iii. Aetiology*

The terms, idiopathic, symptomatic, and cryptogenic used to classify epilepsies should be replaced with the terms genetic (or presumed genetic), structural/metabolic, and unknown respectively. 'Genetic', unlike 'idiopathic', should no longer imply 'benign'. 'Structural/metabolic' includes those genetic conditions where a distinct disorder is interposed between the genetic defect and the epilepsy. 'Epilepsy of unknown cause' no longer presumes a structural/metabolic cause, and may include some of the traditional "idiopathic" developmental syndromes.

For the purposes of this thesis, the terminology of the 1989 classification will be used.

**Table 1-1** *Comparison of 1989/2001 ILAE classification and new proposals*

| | 1989/2001 Classification | 2009 Proposed Classification |
|---|---|---|
| **Classification of seizures:** | | |
| Generalised seizures: | More than minimal involvement of both cerebral hemispheres | Occur in and rapidly engage bilaterally distributed networks |
| | Multiple seizure types | Simplified classification of seizure types |
| Focal seizures: | Initial activation of only part of one cerebral hemisphere | Occur in networks limited to one hemisphere |
| | Simple or complex. | No classification. Described according to manifestation |
| **Classification of syndromes:** | Focal or generalised | No classification |
| **Aetiology:** | Idiopathic, symptomatic or cryptogenic | Genetic, structural-metabolic or unknown |

## 1.5 Basic Mechanisms of epileptogenesis

Epileptic seizures occur as a result of excessive discharge in hyperexcitable neurons. These originate most commonly in cortical and hippocampal structures, although subcortical structures may be involved. The changes underlying neuronal hyperexcitability may involve increased excitatory synaptic transmission, decreased inhibitory neurotransmission, altered ion-channel function or altered intra- or extra-cellular ion concentrations leading to membrane depolarisation. Genetically determined alterations of both ligand-gated and voltage-gated ion channels are of particular importance in epileptogenesis. Changes in ionic currents mediated by sodium, potassium and calcium channels, as well as alteration in function of acetylcholine, glutamate and $\gamma$-aminobutyric acid receptors can lead to different types of epilepsy. Abnormal activity in single neurons leads to recruitment of a critical mass of hyperexcitable cells in highly synchronised activities (Avanzini and Franceschetti 2003). The neurons in an epileptic neuronal aggregate consistently discharge in the form of protracted bursts of action potentials – called paroxysmal depolarisation shifts. Propagation through normal or pathological pathways then occurs resulting in the clinical manifestation of a seizure. Thus both altered neuronal excitability and neuronal circuitry are important in the generation and spread of epileptic discharges.

# *Chapter 2 Molecular genetics of epilepsy*

## 2.1 Introduction

The heterogeneous nature of epilepsy is reflected in the diversity of its many causes. However, a genetic contribution to aetiology is estimated to be present in 40 % of patients, and this proportion is higher in childhood epilepsies. The familial clustering of human epilepsy has long been recognised, and the concept of a genetic predisposition to epilepsy was proposed over 40 years ago (Lennox and Lennox 1960). More recent twin studies have shown that genetic factors are particularly important in the generalised epilepsies but also play a role in the partial epilepsies (Berkovic, Howell et al. 1998). The high frequency of concordant monozygotic twins with the same major syndrome suggests the existence of syndrome specific genetic determinants rather than a single broad genetic predisposition to seizures.

## 2.2 Classification of genetic epilepsies

The human genetic epilepsies can be categorised in several ways, including mechanism of inheritance, whether they are idiopathic (primary) or symptomatic and, where known, which class of gene is involved.

### 2.2.1 Mechanism of inheritance

The mechanism of inheritance identifies three major groups:

- Mendelian epilepsies, in which mutations in a single gene can account for segregation of the disease trait.

- Non-mendelian or ‘complex’ epilepsies, in which several loci interact with environmental factors to produce the pattern of familial clustering. This group includes those epilepsies which exhibit a maternal inheritance pattern due to mutations in mitochondrial DNA.

- Chromosomal disorders, in which the epilepsy results from a gross cytogenetic abnormality.

### 2.2.2 Idiopathic and symptomatic epilepsy

The most commonly observed mendelian epilepsies are ‘symptomatic’, in which recurrent seizures occur as a result of one or more identifiable structural lesions of the brain, and are often one component of a diverse neurological phenotype. Over 200 mendelian diseases include epilepsy as part of the phenotype. The genetic mutations in these disorders may cause abnormal brain development, progressive neurodegeneration or disturbed energy metabolism, so the mechanism of seizure generation is often indirect. The idiopathic epilepsies, in which seizures occur in individuals who are otherwise neurologically and cognitively intact, rarely display a mendelian inheritance pattern. The common familial idiopathic epilepsies, such as juvenile myoclonic epilepsy (JME), childhood absence epilepsy (CAE) and benign childhood epilepsy with centrotemporal spike (BCECTS), tend to display ‘complex’ inheritance. These epilepsies often show age-dependant penetrance with peak onset in childhood. A number of gross chromosomal abnormalities are associated with epilepsy (Singh, Gardner et al. 2002).

### 2.2.3 Class of gene

Although rare, the idiopathic mendelian epilepsies have provided the major recent advances in the molecular basis of the epilepsies. Mutations have been identified in families segregating benign familial neonatal convulsions (BFNC), autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE), generalised epilepsy with febrile seizures plus (GEFS+), childhood absence epilepsy with febrile seizures, autosomal dominant partial epilepsy with

auditory features (ADPEAF) (Table 2-1). Other than *LGI1* and *MASS1*, all mutations occur in genes encoding ion channels, identifying some idiopathic mendelian epilepsies as channelopathies.

Success in determining the molecular genetic basis of the common familial epilepsies has been relatively slow. No gene identified in a mendelian epilepsy acts as a major locus in any non-mendelian epilepsy, and the extent of heterogeneity is likely to be far greater in the complex epilepsies. However, the completion of the human genome project and the development of high-throughout molecular genetic techniques will significantly increase the prospect of identifying the common epilepsy susceptibility genes.

**Table 2-1** *Genes identified in mendelian epilepsies*

| Gene class | Gene | Gene location | Epilepsy Syndrome | Inheritance | Key References |
|---|---|---|---|---|---|
| *Voltage-dependent ion channels:* | | | | | |
| Sodium channels | *SCN1A* | 2q24 | GEFS+/SMEI | AD | Escayg et al. 2000; Claes et al. 2001 |
| | *SCN2A* | 2q23-q24 | GEFS+/BFNIS | AD | Sugawara et al. 2001; Heron et al. 2002 |
| | *SCN1B* | 19q13 | GEFS+ | AD | Wallace et al. 1998 |
| Potassium channels | *KCNQ2* | 20q | BFNC | AD | Leppert et al. 1989; Singh et al. 1998 |
| | *KCNQ3* | 8q24 | | | Lewis et al. 1993; Charlier et al. 1998 |
| | *KCNA1* | 12p13 | EA1 with partial epilepsy | AD | Spauschus et al. 1999 |
| Calcium channels | *CACNA1A* | 19p13 | Generalised epilepsy with ataxia | Sporadic | Jouvenceau et al. 2001 |
| | *CACNB4* | 2q22-23 | IGE/JME/ataxia | AD | Escayg et al. 2000 |
| *Ligand-gated ion channels:* | | | | | |
| Nicotinic acetylcholine receptors | *CHRNA4* | 20q13.2 | ADNFLE | AD | Phillips et al. 1995; Steinlein et al. 1995; Steinlein et al. 1997 |
| | *CHRNB2* | 1p21 | | | Phillips et al. 2001 |
| | *CHRNA2* | 8p21 | | | Aridon et al. 2006 |
| $GABA_A$ receptor | *GABRG2* | 5q34 | GEFS+/SMEI<br>FS with CAE | AD<br>AD | Baulac et al. 2001; Wallace et al. 2001 |
| | *GABRA1* | 5q34 | JME | AD | Cossette et al. 2002 |
| | *GABRD* | 1p36.3 | GEFS+/JME | AD | Dibbens et al. 2004 |
| *Epilepsy associated repeat genes:* | | | | | |
| Leucine-rich, glioma inactivated protein | *LGI1* | 10q24 | ADPEAF | AD | Kalachikov et al. 2002; Scheel et al. 2002 |
| G-protein-coupled receptor | *MASS1* | 5q14 | Febrile and afebrile seizures | AD | Nakayama et al. 2002 |

GEFS+: Generalised epilepsy with febrile seizures plus; SMEI: Severe myoclonic epilepsy of infancy; BFNIS: Benign familial neonatal-infantile seizures; BFNC: Benign familial neonatal convulsions; EA1: Episodic ataxia type1; IGE: Idiopathic generalised epilepsy; ADNFLE: Autosomal dominant nocturnal frontal lobe epilepsy; FS: Febrile seizures; CAE: Childhood absence epilepsy; JME: Juvenile myoclonic epilepsy; ADPEAF: Autosomal dominant partial epilepsy with auditory features

## 2.3 Genetic heterogeneity

The idiopathic mendelian epilepsies demonstrate both locus heterogeneity (mutations in more than one gene causing the same clinical phenotype) and phenotypic heterogeneity (mutations in the same gene causing different clinical phenotypes). This is not surprising given the complexity of the pathways involved in neuronal excitability and the diversity of genes likely to be involved.

## 2.4 Mendelian epilepsy genes

Over 200 mendelian diseases include epilepsy as part of the phenotype, including tuberous sclerosis, fragile X syndrome, neurofibromatosis and large number of metabolic disorders. The small number of primary epilepsies which are inherited in a mendelian fashion are described here. Although they are rare and account for only a small number of epilepsy cases, recognition of the characteristic features and presence of a family history enable a correct diagnosis to be made. Identification of genes responsible for some of these disorders has provided valuable insights into the molecular mechanisms underlying epilepsy.

### 2.4.1 Benign familial neonatal convulsions (BFNC)

Benign familial neonatal convulsions is a rare autosomal dominant idiopathic epilepsy, first described in 1963, and was the first epileptic syndrome to be localised by linkage (Rett and Teubel 1964). Seizures occur in otherwise well neonates from the second or third day of life and remit by week two to three. The characteristic seizures comprise tonic posturing with ocular and autonomic features, followed by a clonic phases with motor automatisms. The prognosis for neurological and intellectual development is favourable, although seizures recur later in life in about 10% of individuals.

BFNC is a good illustration of both clinical and genetic heterogeneity, the latter of which can be explained by the underlying molecular genetics. The first locus (*EBN1*) identified in 1989 was localised to chromosome 20q by linkage analysis in a four generation family with 19 affected individuals (Leppert, Anderson et al. 1989). Six French pedigrees confirmed this linkage (Malafosse, Leboyer et al. 1992). However, in a study of two North American families, one family which showed linkage to *EBN1* included family members with seizures persisting up to 2 years of age, and in one individual, into adolescence (Ryan, Wiznitzer et al. 1991). The other family, none of whose members had seizures after two months of age, could be excluded from linkage to *EBN1*, and was subsequently linked to a second locus (*EBN2*) on chromosome 8q (Lewis, Leach et al. 1993).

The gene for *EBN1*, subsequently named *KCNQ2*, was identified by characterisation of a sub-microscopic deletion on chromosome 20q13.3 in affected individuals and demonstrated significant homology with a voltage-dependent delayed rectifying potassium channel gene, *KCNQ1*(Singh, Charlier et al. 1998). Members of the *KCNQ* potassium channel family comprise six transmembrane-spanning segments (S1-S6), a pore forming loop linking S5 and S6, and intracellular N and C termini. These channels open on membrane depolarisation and are involved in the repolarisation of the action potential and thus in the electrical excitability of nerve and muscle. Mutations in *KCNQ1* can cause the paroxysmal cardiac dysrhythmias long QT syndrome and Jervell-Lange-Nielson cardioauditory syndrome (Wang, Curran et al. 1996; Neyroud, Tesson et al. 1997). Six allelic variants of *KCNQ2* were subsequently identified to segregate with the disease in families with BFNC, including one family whose affected members subsequently developed myokymia (spontaneous contractions of skeletal muscle fibres) (Dedek, Kunath et al. 2001). All mutations involved regions of the gene important for ion conduction, including the S6 domain, the channel pore and the C terminus. Exonic deletions and duplications have now also been identified in families with BFNC (Heron, Cox et al. 2007).

Following identification of *KCNQ2*, a BLAST search was made of the human expressed sequence tag (EST) database, to find cDNA sequences showing significant homology to *KCNQ2*. Rather fortuitously, a novel gene, *KCNQ3*, was identified with 69% similarity to *KCNQ2*, which mapped to the *EBN2* critical region on chromosome 8q24, and was mutated in affected members of the BFNC/*EBN2* family (Charlier, Singh et al. 1998). The missense mutation identified altered a conserved amino acid in the critical pore forming region (the same amino acid was mutated in *KCNQ1* in a patient with long QT syndrome (Wang, Curran et al. 1996)).

*KCNQ2* and *KCNQ3* are co-expressed in most areas of the brain, especially the hippocampus, neocortex and cerebellum. They coassemble and form a heteromeric channel with essentially identical biophysical properties and pharmacologic sensitivities to the native neuronal M-channel (Wang, Pan et al. 1998). The M-channel is a slowly activating and deactivating potassium conductance that plays a critical role in determining the subthreshold electroexcitability of neurons. Therefore mutations in either *KCNQ2* or *KCNQ3* disrupt the native M-current and result in an identical disease phenotype (Figure 2-1(a)).

A possible third locus for AD BFNC was identified in a 3 generation family (Concolino, Iembo et al. 2002). Linkage to *KCNQ2* and *KCNQ3* was excluded, but a pericentric inversion on chromosome 5 was found to segregate with the disease phenotype.

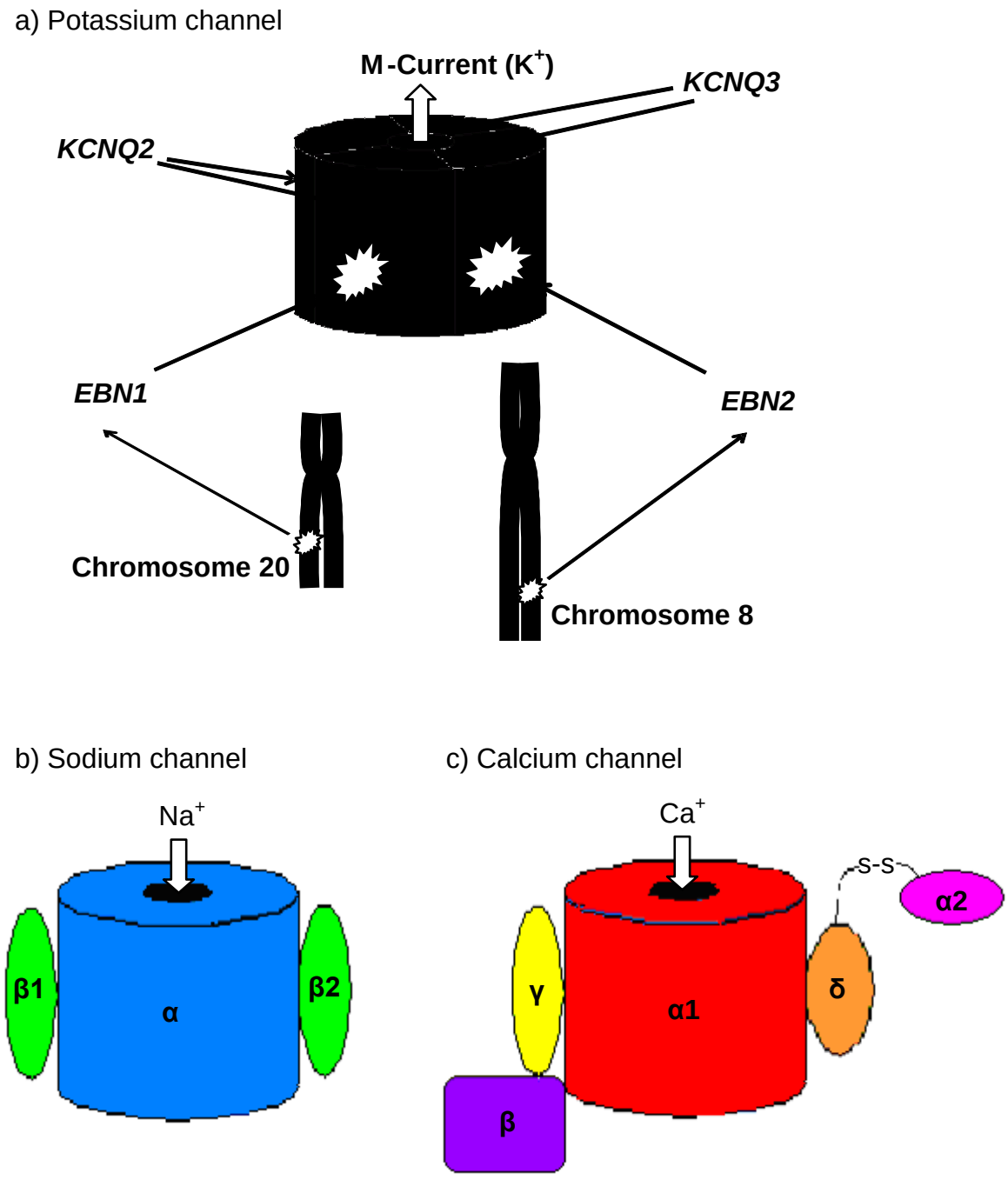



**Figure 2-1** Voltage-dependent ion channel structure. (a) Mutations at either benign familial neonatal convulsions (BFNC) locus, EBN1 (KCNQ2) or EBN2 (KCNQ3), disrupt the heteromeric M-channel (potassium conductance) and result in an identical phenotype. (b) Sodium channels contain a pore forming α subunit with two β subunits. (c) Calcium channels contain a pore forming α subunit with regulatory β, γ, and α2δ subunits.

### 2.4.2 Benign familial infantile convulsions (BFIC)

This mendelian idiopathic epilepsy was first described as an autosomal dominant disorder in families of Italian origin (Vigevano, Fusco et al. 1992), and further families were described in France and Singapore (Lee, Low et al. 1993; Echenne, Humbertclaude et al. 1994). Affected individuals manifested both partial and generalised seizures commencing at age 3-12 months. Patients responded well to conventional anti-epilepsy drugs and prognosis was good with resolution of seizures and no psychomotor retardation.

Gene mapping studies have demonstrated locus heterogeneity. A locus, *BFIC1*, was mapped to chromosome 19q by linkage analysis in Italian families (Guipponi, Rivier et al. 1997). A common haplotype was evident suggesting a founder effect. A second locus, *BFIC2*, was mapped to 16p12-q12 (Caraballo, Pavek et al. 2001). The disease genes have yet to be identified and it is possible that *BFIC2* is an allelic variant of the gene causing infantile convulsions with choreoathetosis (ICCA) which maps to the same genomic region (section 2.4.4).

### 2.4.3 Benign familial neonatal-infantile seizures (BFNIS)

As well as BFNC and BFIC, an intermediate phenotype BFNIS was first described in 1983, with seizure onset from 2 days to 3½ months (Kaplan and Lacey 1983). Missense mutations in the sodium channel subunit gene *SCN2A* were identified in two families with BFNIS. The different mutations identified in the two families both disrupted a conserved leucine, and were predicted to reduce the rate of inactivation of the sodium channel. The number of *SCN2A* mutations identified was subsequently increased to six, in eight families with BFNIS (Berkovic, Heron et al. 2004). Affected individuals showed varying seizure frequencies but all seizures remitted by the age of 12 months. *SCN2A* mutations also cause febrile seizures associated with afebrile seizures, part of the GEFS+ spectrum (section 2.4.6)

### 2.4.4 Familial infantile convulsions and paroxysmal choreoathetosis (ICCA)

This condition was first described in four families from north western France in which infantile convulsions (onset 3-12 months, afebrile, partial with secondary generalisation, remittance by 12 months) were inherited as an autosomal dominant trait in combination with paroxysmal choreoathetosis of variable expression (Szepetowski, Rochette et al. 1997). A genome-wide screen provided evidence of linkage to the pericentromeric region of chromosome 16 encompassing a 10cM interval 16p12-q12. Confirmation of linkage to this region has subsequently been reported in a Chinese family with nine affected members (Lee, Tay et al. 1998). It is also noteworthy that another phenotype, autosomal recessive rolandic epilepsy with paroxysmal exercise-induced dystonia and writer's cramp maps to a region encompassed by the ICCA critical region (Guerrini, Bonanni et al. 1999).

### 2.4.5 Autosomal dominant nocturnal frontal lobe epilepsy (ADNFLE)

First described in six families from Australia, Canada and the UK, autosomal dominant nocturnal frontal lobe epilepsy is characterised by the occurrence of partial seizures almost exclusively during sleep, which may be misdiagnosed as nightmares or other sleep disorders (Scheffer, Bhatia et al. 1994). The familial nature is easily missed, as there is a pronounced variation in severity among family members, and the penetrance is incomplete (approximately 70%). Seizures begin predominantly in childhood, persist into adulthood, and occur in clusters soon after falling asleep or before waking. An aura is often reported. Seizures are characterised by brief tonic or hyperkinetic motor activity with retention of consciousness, although secondary generalisation often occurs. Affected individuals are neurologically and intellectually normal. Ictal-EEG may show bifrontal sharp and slow wave activity. The inter-ictal EEG is usually normal.

Linkage analysis in a single large Australian pedigree with 27 individuals assigned the gene to chromosome 20q13.2 (Phillips, Scheffer et al. 1995). The gene for the α4 subunit of the neuronal nicotinic acetylcholine receptor (nAChR), *CHRNA4*, was known to map to the same chromosomal region, and also to be expressed in the frontal cortex. The nAChRs are presynaptic ligand-gated ion channels with a role in modulating neurotransmitter release during fast synaptic transmission. Mutational analysis of *CHRNA4* identified a missense mutation that cosegregated with the disease in the chromosome 20-linked family (Steinlein, Mulley et al. 1995). The mutation converts a serine to phenylanaline in the M2 transmembrane domain, a crucial structure mediating ionic permeability, and is likely to be disease causing. The mutation was not found in 333 healthy controls. Further site-specific mutations in *CHRNA4* affecting pore forming amino acids have been associated with ADNFLE (Steinlein, Magnusson et al. 1997; Hirose, Iwata et al. 1999; Saenz, Galan et al. 1999; Phillips, Marini et al. 2000; Steinlein, Stoodt et al. 2000).

Other families with ADNFLE were not linked to chromosome 20q. A second locus was mapped to chromosome 15q24 in one family (Phillips, Scheffer et al. 1998). A cluster of nAChR genes (*CHRNA3/CHRNA5/CHRNB4*) in this region seemed good candidates, but no mutations were identified in the pore-forming regions of these genes. However, mutations have been identified in the gene for the β2 subunit of the nAChR, *CHRNB2*, on chromosome 1p21 (De Fusco, Becchetti et al. 2000; Phillips, Favre et al. 2001). The main neuronal nAChR has a heteropentameric structure comprised of α4 and β2 subunits. Thus mutations in the pore forming M2 domains of both *CHRNA4* and *CHRNB2* can produce similar functional effects causing the ADNFLE phenotype.

Mutations in a third nAChR subunit gene, *CHRNA2* on chromosome 8p21, have recently been identified in a Sardinian family with nocturnal frontal lobe seizures inherited in an autosomal dominant pattern (Aridon, Marini et al. 2006).

### 2.4.6 Generalised epilepsy with febrile seizures plus (GEFS+)

This syndrome was first described in 1997 in an Australian family originating from Yorkshire in the UK (Scheffer and Berkovic 1997). Genealogical information was obtained on 2000 family members dating back to the mid-1700s, and clinical information obtained on 289 individuals, of which 28 had seizures. The commonest phenotype, denoted as 'febrile seizures plus' (FS+) comprised a childhood onset of multiple febrile seizures persisting beyond the age of 6, as well as a spectrum of afebrile seizures including absences, myoclonic seizures, atonic seizures and rarely myoclonic-astatic epilepsy. The pattern of inheritance was autosomal dominant.

A second Australian family with generalised epilepsy and febrile seizures plus (GEFS+) was linked to chromosome 19q13.1, and a point mutation identified in *SCN1B*, which encodes the β1 subunit of the voltage gated sodium channel (Wallace, Wang et al. 1998).

Neuronal voltage-gated sodium channels are essential for the generation and propagation of the action potential. They contain a large α subunit associated with two smaller β subunits (**Figure 2-1**(b)). The pore-forming α subunit contains 4 homologous domains each containing six membrane-spanning units. The β subunits contain a single transmembrane region and modulate the gating properties of the channel and are required for normal inactivation kinetics. Mutations in α subunit genes cause several paroxysmal disorders of muscle, including hyperkalaemic periodic paralysis, paramyotonia congenita (*SCN4A*), long QT syndrome (*SCN5A*) and Brugada syndrome (SCN5A) (Bulman 1997). Brugada syndrome is associated with sudden death due to ventricular fibrillation, and fever may be a trigger. Sodium channels are also modulated by antiepileptic drugs such as phenytoin and carbamazepine (Macdonald and Kelly 1994). They are thus good candidate genes for epilepsy.

The *SCN1B* mutation segregated with disease status in the 19q13.1-linked GEFS+ family. It changes a conserved cysteine residue that disrupts a

disulfide bridge normally maintaining an extracellular immunoglobulin-like fold in the β-subunit. This alters the secondary structure of the extracellular domain, which modulates channel gating. The predicted effects of the disruption are to reduce sodium channel expression, slow inactivation, and slow recovery from inactivation. Coexpression of α and mutant β1 subunits exerts significant modulatory effects on the channel gating kinetics and expression levels. The net effect is to cause persistent inward neuronal sodium currents, increased membrane depolarisation, and neuronal hyperexcitability. This may also exaggerate the normal effects of temperature on both conductance and gating of neuronal sodium channels, explaining the apparent temperature dependence of the GEFS+ phenotype.

Two further families with GEFS+ showed linkage to chromosome 2q24, and mutations were identified in *SCN1A*, the gene encoding the sodium channel α1 subunit (Escayg, MacDonald et al. 2000). De novo mutations in this gene have also been identified in patients with severe myoclonic epilepsy of infancy (section 2.4.7). A mutation in the gene encoding the α2 sodium channel subunit, *SCN2A*, has now been identified in a patient with febrile seizures associated with afebrile seizures, consistent with GEFS+. This mutation also slows channel inactivation, suggesting involvement in the epilepsy phenotype (Sugawara, Tsurubuchi et al. 2001).

The GEFS+ phenotype is not only caused by mutations in voltage-gated sodium channels. In one large GEFS+ family, mutations have been identified in the $GABA_A$ receptor α-subunit gene, *GABRG2* (Baulac, Huberfeld et al. 2001). $GABA_A$ receptors are anion-selective ligand-gated channels which mediate fast synaptic inhibition. Binding of GABA opens an integral chloride channel, producing an increase in membrane conductance that results in inhibition of neuronal activity. The GEFS+ mutation substitutes a serine for a methionine in the extracellular loop between transmembrane segments M2 and M3, which was shown to decrease the amplitude of GABA-activated currents when expressed in *xenopus* oocytes. A mutation in *GABRG2* also causes severe myoclonic epilepsy of infancy (section 2.4.7) and a phenotype

of childhood absence epilepsy and febrile seizures (Wallace, Marini et al. 2001).

A mutation in the $GABA_A$ receptor δ-subunit gene, *GABRD*, has now been identified in a small family with GEFS+ (Dibbens, Feng et al. 2004).

### 2.4.7 Severe myoclonic epilepsy of infancy (SMEI)

SMEI was first reported by Dravet, who described a syndrome with onset of prolonged febrile seizures (generalised or unilateral) in the first year of life, followed by afebrile seizures including tonic-clonic, myoclonic, absence seizures and recurrent status epilepticus (Dravet 1978). Neurological development appears normal before the onset of seizures but becomes progressively delayed as seizure frequency increases.

Heterozygous de-novo mutations were identified in *SCN1A* (Claes, Del-Favero et al. 2001; Ohmori, Ouchida et al. 2002; Sugawara, Mazaki-Miyazaki et al. 2002). Mutations included insertions, deletions and missense mutations, several resulting in premature stop codons and truncated channels. As *SCN1A* mutations have been identified in both SMEI and in families with GEFS+, and some patients with SMEI have a family history of seizures consistent with the spectrum of GEFS+, it has been suggested that SMEI is the most severe phenotype in the GEFS+ spectrum (Singh, Andermann et al. 2001). A recent study of 333 patients with SMEI identified *SCN1A* mutations in 242 (73%), with no obvious genotype-phenotype correlations (Depienne, Trouillard et al. 2009).

Mutations in *GABRG2* have also been identified in a patient with SMEI, who belonged to a family with several members with GEFS+ (Harkin, Bowser et al. 2002).

Two related syndromes have been described. SMEI-borderland (SMEB), referring to children who lack several of the key features of SMEI such as myoclonic seizures or generalized spike-wave activity, has been attributed to *SCN1A* mutations in 69% of cases (Harkin, McMahon et al. 2007). Intractable

childhood epilepsy with generalized tonic clonic seizures (ICEGTC), also used to describe children with a phenotype similar to SMEI but without minor seizures, particularly myoclonic seizures, was identified to have SCN1A mutations in 70% of cases (Fujiwara, Sugawara et al. 2003).

### 2.4.8 Autosomal dominant partial epilepsy with auditory features (ADPEA)

Autosomal dominant partial epilepsy with auditory features was first described in a three generation family with 11 members diagnosed with an idiopathic/cryptogenic epilepsy (Ottman, Risch et al. 1995). The epilepsy was clearly localized in all but one of these cases, with both simple and partial complex seizures progressing to secondarily generalised tonic-clonic seizures. Six of those with idiopathic epilepsy reported auditory disturbances as a simple partial component of their seizures. The seizures were infrequent with an age of onset between 8 and 19 years.

A genome screen identified linkage over a 10cM region on chromosome 10q23.3-24 (Ottman, Risch et al. 1995). This region was narrowed to 3cM by a genome screen in a large Basque family segregating lateral temporal lobe epilepsy with auditory and visual features (Poza, Saenz et al. 1999). Further families confirmed the linkage but failed to narrow the region. Construction of a physical map identified 28 putative genes of which 21 were sequenced in an affected individual from three families, and mutations subsequently checked in a further two families (Kalachikov, Evgrafov et al. 2002). Mutations were identified in the leucine rich, glioma-inactivated 1 gene (*LGI1*) in all affected individuals and obligate carriers, as well as six unaffected members, consistent with a 71% disease penetrance. The five mutations identified were not present in 123 unrelated controls. Analysis of ten families with ADPEAF identified *LGI1* mutations in 50% of families, with penetrance of 54% (Ottman, Winawer et al. 2004). LGI1 mutations have not been identified in individuals with other forms of temporal lobe epilepsy (Berkovic, Izzillo et al. 2004).

*LGI1* is a member of the leucine-rich repeat (LRR) superfamily, in particular the adhesive proteins and receptors. It was first described following the observation that it was disrupted by translocation in a glioblastoma multiforme cell line and in over one quarter of primary tumours. The *LGI1* protein consists of an extracellular domain with LRR repeat motifs, a transmembrane segment and an intracellular segment of unknown function (Somerville, Chernova et al. 2000). The extracellular portion aligns most closely with a group of proteins, including slit, toll and tartan, involved in CNS development and in which the LRRs bind nerve growth factor and other neurotrophins. Interestingly, a C-terminal repeat motif, now referred to as the EAR (epilepsy-associated repeat) domain, has been identified in both *LGI1* and the *MASS1* gene. *MASS1* is mutated in the Frings mouse model of audiogenic epilepsy (Scheel, Tomiuk et al. 2002) and in one family with febrile and afebrile seizures (Nakayama, Fu et al. 2002). This EAR domain is likely to play a role in the pathogenesis of epilepsy. *LGI1* is expressed predominantly in brain (cerebellum, cortex, medulla, occipital pole, frontal lobe, temporal lobe and putamen), muscle and spinal cord. *LGI1* enhances AMPA receptor-mediated synaptic transmission in hippocampal slices and binds to *ADAM22* (Fukata, Adesnik et al. 2006). Of the five mutations identified in ADPEA, three were missense mutations with predicted premature truncation of the *LGI1* protein, one was a non-synonymous point mutation in the highly conserved extracellular and C-terminal region, and one was an intronic mutation predicted to alter a splice site. *LGI1* is therefore the first non ion-channel gene identified as causing an idiopathic epilepsy in humans.

### 2.4.9 Familial partial epilepsy with variable foci (FPEVF)

This idiopathic epilepsy displays autosomal dominant inheritance with reduced penetrance and locus heterogeneity. It was first described in an Australian pedigree with suggestive linkage to chromosome 2 (Scheffer, Phillips et al. 1998) and subsequently in two large French-Canadian families

with linkage to chromosome 22q11-q12 (Xiong, Labuda et al. 1999). The phenotype is characterised by recurring partial seizures originating from different cortical areas, usually in the frontal or temporal lobes. The epileptic focus varies between family members. In the Australian family seizures were diurnal and the age of onset varied from one year to 44 years. In the French-Canadian pedigrees age of onset peaked at 5 and 25 years and seizures were predominantly or exclusively nocturnal.

### 2.4.10 Infantile spasms (West syndrome)

Infantile spasms are divided into those that are symptomatic and those that are cryptogenic or idiopathic. The majority (70-80%) are symptomatic and may be attributed to a prenatal, perinatal or postnatal cause, of which prenatal aetiologies are the most common (50%). Many of theses are genetically determined, including disorders of brain development, neurocutaneous syndromes, metabolic disorders and chromosomal abnormalities.

Most cases of idiopathic infantile spasms are sporadic, and the recurrence risk is less than 1% (Dulac, Feingold et al. 1993). However several familial cases of early infantile epileptic encephalopathy progressing to infantile spasms have been identified consistent with X-linked inheritance. Feinberg and Leahy first reported five affected males in four sibships of a three-generation family (Feinberg and Leahy 1977). Subsequently, five further families have been identified, some of which also include individuals with X-linked mental retardation without infantile spasms (Rugtveit 1986; Claes, Devriendt et al. 1997; Bruyere, Lewis et al. 1999; Stromme, Sundet et al. 1999; Stromme, Mangelsdorf et al. 2002). Linkage analysis in these families mapped the disease gene to chromosome Xp21.3-Xp22.1 (Claes, Devriendt et al. 1997; Bruyere, Lewis et al. 1999; Stromme, Sundet et al. 1999). The aristaless-related, homeobox gene, *ARX*, was considered a candidate on the basis of its expression pattern in fetal, infant and adult brain. Screening of this gene identified mutations in four of the five families with infantile spasms (Stromme, Mangelsdorf et al. 2002). Mutations have also been identified in

families with various combinations of mental retardation together with myoclonic seizures, spasticity, ataxia, chorea and dystonia, with and without infantile spasms (Scheffer, Wallace et al. 2002; Guerrini, Moro et al. 2007). Two recurrent mutations identified result in expansion of polyalanine tracts of the *ARX* protein. These are likely to cause protein aggregation, as has been demonstrated in other human diseases caused by alanine expansions (Calado, Tome et al. 2000). Homeobox-containing genes are important in the regulation of key stages of development. *ARX* encodes one of a class of proteins incorporating a C-terminal aristaless domain presumed particularly important in the differentiation and maintenance of specific neuronal subtypes in the cerebral cortex (Bienvenu, Poirier et al. 2002).

A family with three siblings (twin girls and a boy) with severe mental retardation, seizures including infantile spasms and features of Rett syndrome displayed mutations in *CDKL5* (Weaving, Christodoulou et al. 2004). A recent review of girls with epilepsy associated with *CDKL5* mutations, identified three clinical stages: early epilepsy (onset between one and ten weeks), epileptic encephalopathy with hypsarrhythmia and developmental stasis, and late multifocal and myoclonic epilepsy (Bahi-Buisson, Kaminska et al. 2008). *CDKL5* and *MECP2* (the gene responsible for Rett syndrome) interact in vivo and in vitro and they demonstrate overlapping expression during neural maturation and synaptogenesis (Mari, Azimonti et al. 2005).

## 2.5 Non-mendelian epilepsy genes

Juvenile myoclonic epilepsy, benign childhood epilepsy with centrotemporal spikes and febrile seizures are discussed in this section. The genetics of absence epilepsy will be reviewed in chapter 3.

### 2.5.1 Juvenile Myoclonic Epilepsy (JME)

Juvenile myoclonic epilepsy (JME) is one of the idiopathic generalized epilepsies, and accounts for 5 to 10% of all epilepsy. It was first described as a distinct electroclinical syndrome by Janz and Christian in 1957 (Janz and Christian 1957). The onset of seizures occurs usually between the ages of 8 and 26, with a peak in adolescence. It is characterized by the presence of generalized myoclonic seizures involving the upper limbs predominantly. Myoclonic seizures occur predominantly in the morning, soon after awakening, and are exacerbated by sleep deprivation and alcohol. In addition, nearly all affected individuals have generalized tonic-clonic seizures, often following a series of myoclonic jerks, and about 20-40% have absence seizures. The EEG shows spike and wave or polyspike and wave at a frequency of 4 to 6 Hz. Photosensitivity is quite common. The susceptibility to seizures is life long and patients require long-term maintenance on antiepileptic medication. The drug of choice is sodium valproate.

A genetic contribution to the aetiology of JME is well established (Janz 1969; Tsuboi and Christian 1973; Sundqvist 1990), but the mode of inheritance is uncertain. Autosomal dominant (Delgado-Escueta, Greenberg et al. 1990), autosomal recessive (Panayiotopoulos and Obeid 1989), and two locus models have all been proposed (Greenberg, Delgado-Escueta et al. 1989).

Evidence for linkage of the JME trait to the serologic markers HLA and properdin factor B on chromosome 6p was first found in 1988 (Greenberg, Delgado-Escueta et al. 1988), and the locus designated *EJM1* (subsequently

renamed *EJM3*). These findings were subsequently confirmed in a separately ascertained group of 23 families using HLA serologic markers (Weissbecker, Durner et al. 1991). Analysis of a subset of these families together with one new family, using HLA-DQ restriction fragment length polymorphisms (RFLPs), gave similar results (Durner, Sander et al. 1991). Further work by Greenberg and Delgado-Escueta (1993) on a larger group of 35 families once again confirmed linkage to the serologic markers HLA and properdin factor B, with a maximum lod score of 4.2 (Greenberg and Delgado-Escueta 1993). A study in a single large pedigree of Belize origin using microsatellite markers on chromosome 6p gave a maximum lod score of 3.67 ($\theta$=0) between the marker *D6S257* and a trait defined as the presence of clinical JME or an EEG showing diffuse 3.5-6 Hz multispike and slow wave complexes (Liu, Delgado-Escueta et al. 1995). In addition, linkage analysis in 28 families ascertained through a JME patient in which family members with IGE (idiopathic generalised epilepsy) were classified as affected gave a lod score for the DQB1 locus of 4.2. The linkage pattern observed suggested heterogeneity and an excess of transmission from mothers (Greenberg, Durner et al. 2000). Association analysis found significant linkage disequilibrium between JME and a haplotype of five SNPs and microsatellites between HLA-DQ and HLA-DP on chromosome 6p21 (Pal, Evgrafov et al. 2003). DNA sequencing identified two JME-associated SNP variants in the promoter region of *BRD2*, a putative nuclear transcriptional factor, but no potentially causative coding mutations.

Two studies from a single group have failed to find evidence for the existence of a locus on chromosome 6p (Whitehouse, Diebold et al. 1993; Elmslie, Williamson et al. 1996). No significant evidence in favour of linkage was obtained under any of the models tested. These results suggest that genetic heterogeneity exists within this epilepsy phenotype.

Chromosomal regions harbouring genes for subunits of the neuronal nicotinic acetylcholine receptor were tested for linkage to the JME trait in 34 European pedigrees. Two-point lod scores were negative for all loci except *D15S128* and *D15S118* on chromosome 15q14. Seven additional marker loci

encompassing a 20.1-cM region were selected in order to investigate this region further. A maximum multipoint lod score of 4.18 was obtained under the assumption of heterogeneity at $\alpha$ = 0.64 (where $\alpha$ is proportion of linked families). Analysis of recombinant events defined the 10-cM interval between *D15S144* and *D15S1012* as being the region in which the gene lies. The $\alpha$-7 subunit of the neuronal nicotinic acetylcholine receptor (*CHRNA7*) maps within this interval and therefore represents an excellent candidate gene. These results indicate that a major susceptibility locus for JME, termed *EJM2*, may map to this region of chromosome 15q (Elmslie, Rees et al. 1997). A further study was undertaken to test this linkage finding and to explore whether this susceptibility locus contributes to the pathogenesis of a broader spectrum of IGE syndromes. Linkage analysis with 7 microsatellite markers encompassing the *CHRNA7* region failed to replicate evidence of linkage in 11 families with at least two JME members. No evidence in favour of linkage to 15q14 was found under a broadened diagnostic scheme in 27 families of JME probands or in 30 families of probands with idiopathic absence epilepsy (Sander, Schulz et al. 1999). A subsequent study has clarified the linkage data in relation to the current map of the region study (Taske, Williamson et al. 2002). The *CHRNA7* gene and its partial duplication *CHRFAM7A* were screened for mutations, but no causative sequence variants could be identified.

Linkage analysis in families from Belize and Mexico, comprising members with JME and asymptomatic individuals with polyspike wave complexes on EEG, identified a disease locus at chromosome 6p21.2-p11, termed *EJM1* (Liu, Delgado-Escueta et al. 1995; Liu, Delgado-Escueta et al. 1996; Bai, Alonso et al. 2002). Heterozygous mutations were identified in the *EFHC1* gene, in both affected individuals and unaffected family members from six families, but not in 382 healthy unrelated controls (Suzuki, Delgado-Escueta et al. 2004). *EFHC1* appeared to increase *CACNA1E* mediated R-type $Ca^{2+}$ currents in patch clamp studies, and this was reversed by the mutations associated with JME. Overexpression of *EFHC1* in mouse hippocampal primary culture neurons induced apoptosis, and this was significantly lowered by the mutations. *EFHC1* mutations have also been identified in sporadic

individuals with juvenile absence epilepsy and ‘cryptogenic temporal lobe epilepsy’ (Stogmann, Lichtner et al. 2006).

Linkage analysis in an extended French Canadian autosomal dominant JME pedigree mapped the locus to 5q34, and a missense mutation (Ala 322 Asp) was identified in this family in the *GABRA1* gene (Cossette, Liu et al. 2002). Mutations have also been identified in individuals or single families with JME in *CACNB4* (Escayg, De Waard et al. 2000) and *GABRD* (Dibbens, Feng et al. 2004).

### 2.5.2 Benign Childhood Epilepsy With Centrotemporal Spikes (BCECTS)

The syndrome of benign childhood epilepsy with centrotemporal spikes (BCECTS), or benign rolandic epilepsy, was first described by Nayrac and Beaussart (1958) (Nayrac and Beaussart 1958). The onset of seizures occurs during childhood between the ages of 3 and 13 in a child who is neurologically intact. The typical seizures often occur at night and are preceded by a somatosensory aura around the mouth. This is followed by excessive salivation, a gurgling noise and speech arrest with retention of consciousness. Unilateral motor seizures of the face follow and can progress to a secondary generalized seizure (Panayiotopoulos 1993). Following the seizure, the child may feel numbness or pins and needles in the side of the face involved in the seizure. The pattern of the seizures varies diurnally, with nocturnal seizures more likely to generalize secondarily (Holmes 1993). The prognosis is good. Seizures usually resolve within 1 to 3 years of onset, and rarely persist beyond 16 years.

The EEG is characteristic. High-amplitude spikes or sharp waves are followed by slow waves, occurring unilaterally or independently bilateral in the central (Rolandic) and mid-temporal areas. About 20% of children who have Rolandic discharges on EEG will not have seizures (Beaussart 1972).

A family history of epilepsy is known to be common among probands with BCECTS, although the proportion with a family history varies from study to

study. Figures of 9-59% have been obtained (Blom and Heijbel 1975; Lerman and Kivity 1975). A high proportion of relatives also have abnormal EEGs.

Bray and Wiser analyzed the EEGs from the families of 40 patients with seizures and centrotemporal spikes and found that 36% of the siblings and 19% of parents had focal epileptiform activity on the EEG (Bray and Wiser 1965). They postulated that an autosomal dominant gene with age-dependent penetrance was responsible for producing centrotemporal spikes. Heijbel et al (1975) studied 19 probands with BCECTS (Heijbel, Blom et al. 1975). Of 34 siblings 15% had Rolandic discharges and seizures, and a further 19% had Rolandic discharges in isolation. Their findings tended to support the suggestion of an autosomal dominant gene with age-dependent penetrance. However a study of 8 pairs of twins (6 monozygotic, 2 dizygotic ), where one twin had a diagnosis of BCECTS, found no concordance for the phenotype in any of the twin pairs (Vadlamudi, Harvey et al. 2004). This provided evidence against a strong genetic component.

Generalized epileptiform activity has also been found in the EEGs of relatives of patients with BCECTS. Degen and Degen found epileptiform activity in the EEGs of 26 out of 69 (38%) siblings of 43 probands with BCECTS (Degen and Degen 1990). Among those siblings aged 5 to 12 years, 54% had an abnormal EEG, and the proportion declined in the younger and older age groups. This pattern of age-dependent penetrance and the finding of generalized EEG abnormalities in the siblings of patients has led to the suggestion that BCECTS and absence epilepsy may be linked (Bray and Wiser 1969).

Twenty-two nuclear families segregating BCECTS were examined for linkage to chromosomal regions known to harbour neuronal nicotinic acetylcholine receptor (nAChR) subunit genes. Evidence was found for linkage with heterogeneity to a region on chromosome 15q14 in the vicinity of the alpha 7 nAChR subunit gene, *CHRNA7* (Neubauer, Fiedler et al. 1998).

### 2.5.3 Febrile seizures

Susceptibility to febrile seizures is known to have a strong genetic basis, and a significant proportion of patients have a family history of febrile convulsions or other epilepsies. The proportion of probands with an affected first degree relative has been estimated as between 8% and 49% (Rich, Annegers et al. 1987; Wallace 1988). The mode of inheritance seems to depend on the frequency of febrile convulsions in the proband. Complex segregation analysis performed on 467 nuclear families ascertained through probands with febrile convulsions showed clear evidence for polygenic inheritance in those families in which the proband had a single febrile convulsion (Rich, Annegers et al. 1987). The heritability of liability was estimated at 68% in these families. However, in the families of probands with more than three febrile convulsions there appeared to be a single major locus contributing to seizure susceptibility. Another study of the families of 52 probands with febrile seizures found that 40 families (77%) had at least one further affected member, and these were consistent with an autosomal dominant mode of inheritance with reduced penetrance (64%) (Johnson, Kugler et al. 1996).

The investigation of large pedigrees has led to the identification of at least ten putative loci for febrile seizures, *FEB1-10*, including three genes in which mutations have been found:

- A heterozygous loss-of-function *SCN1A* mutation was identified in an Italian family comprising 12 individuals with simple febrile seizures resolving by the age of six (Mantegazza, Gambardella et al. 2005). Three family members developed afebrile seizures of mesial temporal lobe origin, and two of these had evidence of hippocampal sclerosis on MRI.

- A nonsense mutation in *MASS1(GPR98)* was identified in a Japanese family with febrile and afebrile seizures (Nakayama, Fu et al. 2002). The mutation was found in the proband who had febrile and afebrile seizures before the age of 4, her brother who had two febrile seizures age 6 and 7 years, and their unaffected father. *MASS1* encodes a

fragment of a large G protein-coupled receptor containing an epilepsy associated repeat (EAR) domain also found in *LGI1* (section 2.4.8). *MASS1* is mutated in the Frings mouse model of audiogenic epilepsy (Scheel, Tomiuk et al. 2002)

- A heterozygous missense mutation in *GABRG2* was identified in a family whose individuals manifested febrile seizures with or without childhood absence epilepsy (Wallace, Marini et al. 2001). A splice site mutation in *GABRG2* was identified in a similar family with febrile seizures and absence epilepsy (Kananura, Haug et al. 2002). A third family with isolated febrile seizures was also found to have a missense mutation in *GABRG2* (Audenaert, Schwartz et al. 2006). Febrile seizures occur as a part of the syndrome of Generalised Epilepsy with Febrile Seizures Plus (GEFS+), for which several genes have been identified, including *GABRG2* (section 2.4.6).

# *Chapter 3 Absence epilepsy*

Childhood absence epilepsy (CAE), juvenile absence epilepsy (JAE) and epilepsy with myoclonic absences comprise a group of idiopathic absence epilepsies which vary in their age of onset, seizure frequency and pattern of evolution. Eyelid myoclonia with absences, perioral myoclonia with absences and episodic ataxia with absences seem to represent further distinct absence epilepsy syndromes.

## 3.1 Childhood absence epilepsy (CAE)

CAE is an idiopathic generalised epilepsy of childhood which is well characterised in terms of age of onset, clinical features and EEG findings. There is strong evidence for a genetic aetiology.

### 3.1.1 Historical background

Childhood absence epilepsy was probably first described in 1705 by Poupart. However a more widely known description is that of Tissot in 1770, who gives an account of 'petits accés' (minor attacks) occurring in a fourteen year old girl from the age of seven years: '…the young female patient had frequently had very short minor attacks which were recognised only by an instantaneous loss of consciousness which stopped the patient's speech, accompanied by a very slight movement in the eyes' (Loiseau 1992). The term 'petit-mal' was first used in 1815 by Esquirol for non-convulsive epileptic seizures, and in 1824 Calmeil used the term 'absence d'esprit' to describe the 'fleeing of the spirit from the eyes of the afflicted patient'. The term pyknolepsy was introduced by Saur in 1916, and described by Adie in 1924 as 'a disease with an explosive onset between the age of 4 and 14 years, of frequent, short, very slight, monotonous minor epileptiform seizures of uniform severity which recur almost daily for weeks, months or years, are

uninfluenced by antiepileptic remedies, do not impede normal mental and psychological development, and ultimately cease spontaneously never to return'. There has subsequently been some confusion around the use of petit mal, as at various times it has been applied to (i) brief interruptions of consciousness associated with a rhythmic 3 cycle/second discharge of regular spike and wave complexes on the EEG (Gibbs, Davis et al. 1935), (ii) a petit mal type of dysrhythmia in a routine EEG record in the absence of clinically obvious seizures, even if the particular patient has no history of epilepsy (Loiseau 1992) and (iii) a petit mal triad consisting of absences, myoclonic seizures and akinetic seizures (Lennox 1945). The International Classification of Epileptic Seizures (Gastaut 1970) distinguished typical absences from atypical absences, and the term petit mal was discarded.

### 3.1.2 Epidemiology

CAE represents approximately 8% of epilepsy in school age children (Cavazzuti 1980) and is the most frequent absence epilepsy syndrome. The estimated annual incidence of CAE is 6.3-8.0 per 100,000 children aged 1-15 years (Loiseau, Loiseau et al. 1990). The prevalence of absence seizures, including CAE, is 0.1-0.7 per 1000 in children up to the age of 16 years. CAE is commoner in girls than boys (Loiseau 1992).

### 3.1.3 Clinical features

A diagnosis of CAE should be restricted to an epilepsy with the following characteristics:

- Onset between 3 and 12 years of age (peak at 6-8 years)
- Absence seizures occurring in children who are otherwise developmentally and neurologically normal
- Frequent absence seizures (of any type except myoclonic absences) as the initial and predominant seizure type

- Absence seizures associated with characteristic EEG changes (see below)

- Generalised tonic-clonic seizures may also occur

The absence seizures of CAE are short (typically less than 10 seconds), have an abrupt onset and termination and occur frequently throughout the day (up to 200 per day). The child shows a loss of awareness and responsiveness, with vacant eyes which may drift upwards, and stops all ongoing activities. However the International Classification of Epileptic Seizures (ILAE 1981) distinguishes six types of absence seizures, according to their associated clinical features, several of which may occur in a given patient or during a single absence seizure:

- Simple absences, with impairment of consciousness only

- Absences with mild clonic components (often involving eyelids only, although jerking of the head and limbs may occur). True eyelid myoclonia with absences represents a distinct epilepsy syndrome.

- Absences with atonic components. A gradual postural loss occurs, with drooping of the head and shoulders. A sudden fall to the ground usually signifies atypical absences.

- Absences with tonic components. Increased postural tone of extensor muscles causes the eyes to rotate upwards and the head to draw backwards.

- Absences with automatisms. These may be perseverative (continuing an activity such as eating or walking) or de novo (lip licking, swallowing, fumbling, mumbling).

- Absences with autonomic components, including flushing, tachycardia and urinary incontinence.

Absence seizures occur spontaneously but a number of precipitating factors may be identified. Hyperventilation is an extremely common trigger. Other precipitating factors include strong emotions, boredom and exercise.

### 3.1.4 EEG characteristics

The ictal EEG shows bilateral, synchronous, symmetrical discharges of 2.5-4Hz spike-wave complexes (Figure 3-1). Onset is abrupt and synchronous, with the anterior leads usually involved first. The amplitude is highest under the fronto-central leads. The initial frequency of the spike-wave complexes is 3 Hz, which then slows to 2-2.5 Hz as the amplitude decreases towards the end of the seizure. Irregular spike-wave discharges and polyspike-waves are compatible with a diagnosis of CAE, but are more likely to be associated with abnormal interictal EEGs, prolonged absence seizures, multiple seizure types and developmental delay (Holmes, McKeever et al. 1987).

The interictal EEG is usually normal. Interictal high-voltage 3-4 Hz occipital delta rhythms may be present. This posterior rhythm is blocked by eye-opening and enhanced by hyperventilation. Some photosensitivity may be present, but marked photosensitivity should be considered as exclusion criteria for CAE.

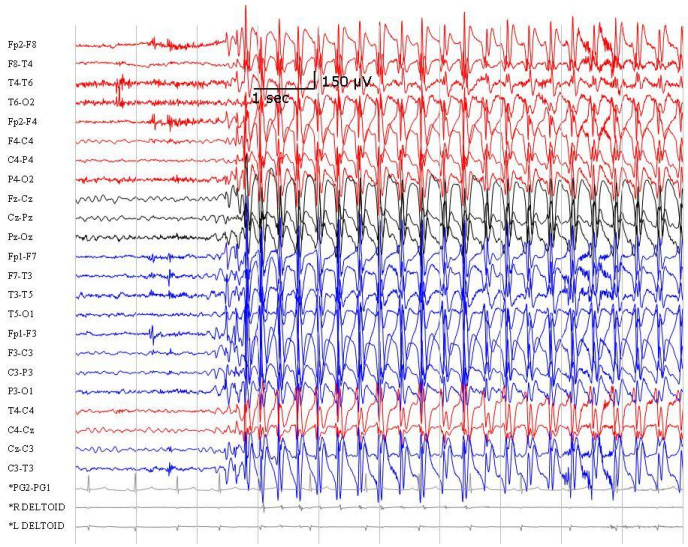



**Figure 3-1** *EEG in childhood absence epilepsy, showing bilateral, synchronous, symmetrical discharges of 3Hz spike-wave complexes*

### 3.1.5 Prognosis

70-80% of patients with CAE have their seizures controlled by a single antiepileptic drug (usually ethosuximide, sodium valproate or lamotrigine). There are three long-term outcomes for patients:

- Patients become seizure free. Remission rates of between 33% and 79.3% have been reported (Loiseau 1992). The longer the follow up, the higher the occurrence of generalised tonic-clonic seizures. In one series, the average duration of absence seizures was 6.6 years (Currier, Kooi et al. 1963). However, absence seizures continue to remit beyond puberty and well into adulthood.

- Absence seizures persist. This occurs in only about 6% of patients (Currier, Kooi et al. 1963). Absence seizures become less frequent and tend to be very short.

- Tonic-clonic seizures develop. These occur in about 16-44% of patients, usually between 8 and 15 years of age, but may appear beyond 30 years (Currier, Kooi et al. 1963; Livingston, Torres et al. 1965; Loiseau, Duche et al. 1995). Onset of tonic-clonic seizures is most frequently 5-10 years after the onset of absence seizures. Several predisposing factors have been identified:

    1. Absence seizures occurring after 8 years of age
    2. Female sex
    3. Poor response to initial treatment
    4. Incorrect choice of initial drug
    5. Abnormal background EEG activity (other than posterior delta rhythms) or photosensitivity

## 3.2 Juvenile absence epilepsy

Juvenile absence epilepsy was defined as a syndrome distinct from CAE following studies of patients with 'spike-wave absences' by Doose et al in 1965 (Doose, Volzke et al. 1965) and 'age-related minor seizures' by Janz in 1969 (Janz 1969). The Revised International classification of Epilepsies and Epileptic Syndromes (ILAE 1989) describes JAE as an age-related idiopathic generalised epilepsy with onset of absences around puberty (after 10 years of age), less frequent absences than in CAE, a more frequent association with generalised tonic clonic seizures and myoclonic seizures occurring not infrequently. JAE is less common than CAE and affects both sexes equally.

Additional clinical and EEG features have been described by Duncan and Panyiotopoulos (Duncan and Panayiotopoulos 1992) These include, in comparison to CAE, absences of longer duration and with less severe impairment of cognition, and a poorer prognosis with seizures persisting throughout life but decreasing in severity. Myoclonic jerks occur in 15-25% of patients. However eyelid or perioral myoclonus, rhythmic limb jerking and single or arrhythmic myoclonic jerks of the head, trunk or limbs during the absence seizure are not compatible with JAE. The EEG shows generalised spike or polyspike and slow wave complexes, starting at 3 Hz (>2.5 Hz) and gradually reducing in frequency. The inter-ictal EEG is generally normal. Photosensitivity or other sensory precipitants are unusual.

## 3.3 Other absence epilepsies

CAE and JAE are the two IGEs recognised by the International Classification of Epilepsies and Epileptic Syndromes (ILAE 1989) in which absence seizures are the main seizure type. A third syndrome, epilepsy with myoclonic absences, was classified as 'cryptogenic or symptomatic'. Eyelid myoclonia with absences and perioral myoclonia with absences seem to be two distinct IGE syndromes but have not been recognised as such by the ILAE. A syndrome of episodic ataxia with absences also now appears to

represent a distinct genetic entity, and this is described in **Error! Reference source not found.**. Epilepsy with myoclonic absences, eyelid myoclonia with absences, perioral myoclonia with absences and episodic ataxia with absences will be described briefly. Absence seizures also occur in JME and in epilepsy with generalised tonic-clonic seizures on awakening, but are not the predominant seizure type.

### 3.3.1 Epilepsy with myoclonic absences

An epilepsy syndrome in which myoclonic absences constitute the only or predominant seizure type was first described by Tassinari et al in 1969 (Tassinari, Lyagoubi et al. 1969). Epilepsy with myoclonic absences was included among the cryptogenic or symptomatic generalised epilepsies in the International Classification of Epilepsies and Epileptic Syndromes (ILAE 1989). It is characterised by absences accompanied by severe bilateral (but sometimes asymmetrical) rhythmical myoclonic jerks of shoulders, arms and legs, often associated with a tonic contraction leading to progressive elevation of the arms. When facial myoclonias occur they tend to be perioral, with eyelid involvement rare. Autonomic manifestations include changes in respiration and urinary incontinence. Impairment of consciousness is variable, and patients may be aware of the jerks. The ictal EEG shows bilateral, synchronous, and symmetrical spike-wave discharges at 3 Hz, similar to CAE. Age of onset is 7 years (11 months-12 years), there is a male preponderance and a family history of epilepsy in about 20% of cases (Tassinari, Michelucci et al. 1995). Myoclonic absences resolve in about half of patients, after a mean period of 5.5 years from onset. A poor prognosis is associated with a high incidence of associated seizures, mainly generalised tonic clonic seizures and atonic seizures, and inappropriate or inadequate medical therapy. Refractory myoclonic absences persist for 20 years or more, and these are associated with a decline in intellectual function.

### 3.3.2 Eyelid myoclonia with absences

Jeavons first described eyelid myoclonia with absences as a rare but specific entity in 1977 (Jeavons 1977). It is characterised by eyelid myoclonia, typical absences and photo-sensitivity (Appleton 1995). Onset occurs in the first ten years of life, usually before the age of six. The initial seizure type may be absences or eyelid myoclonia, which differs from the eyelid flutter seen in CAE. The fast eyelid myoclonic jerks last between 1and 3 seconds and may be accompanied by upward deviation of eyes, head or shoulders. A brief absence usually occurs simultaneously. Limb myoclonus is rare, but automatisms and chewing movements do occur. Photosensitivity is common and present from an early age. Cognitive impairment does not generally occur. The ictal EEG shows brief (1-4s) spike-wave or polyspike-wave discharges at a frequency of 3-6 Hz, occurring on eye-closure and on photic stimulation. 80% of patients are female (Appleton 1995). Remission of seizures is rare, anti-epileptic medication may be required for life and complete seizure control should not be expected.

### 3.3.3 Perioral myoclonia with absences

Panayiotopoulos reported two patients in 1992 who could not be satisfactorily classified, in whom perioral myoclonus was a feature of absences (Panayiotopoulos 1993). At least six further patients have been identified and described (Panayiotopoulos, Ferrie et al. 1995). The majority (75%) are female. Onset is often with a generalised tonic-clonic seizure in childhood or early adolescence, followed by frequent typical absence seizures with rhythmical myoclonus of the perioral or masticatory facial muscles. GTCS occur invariably and may follow absence status. Ictal EEG shows high amplitude generalised irregular spike or polyspike and slow wave discharges at 3-4Hz. There is no photosensitivity. Seizures may persist throughout life and be resistant to medication.

.

## 3.4 Genetic basis of absence epilepsy

A genetic component to aetiology is well established but the mechanism of inheritance and the key genes involved are unknown. Approximately 1.6% of siblings of probands with CAE will have absence epilepsy, giving a $\lambda_s$ (the risk to a sibling of an affected proband compared with the population risk) of at least 27 (Beck-Mannagetta and Janz 1991). Studies of familial clustering indicate that CAE has a 'complex' non-mendelian mode of inheritance (Gardiner 1995). A study of concordance of seizure type and seizure syndrome within 31 families with IGE found that 65% of families were concordant for seizure type (absence, myoclonic, or both) with an expected concordance by chance of 24% ($p<0.0001$) (Winawer, Rabinowitz et al. 2003). Concordance of epilepsy syndrome (JME versus CAE plus JAE) in 25 families was 68% with an expected concordance by chance of 48% ($p=0.012$). When restricted to families with only JME, CAE and JAE, concordance of epilepsy syndrome (JME versus CAE plus JAE) in 19 families was 79% with an expected concordance by chance of 58% ($p=0.002$). These results provide evidence for distinct genetic effects on both seizure type and epilepsy syndrome. One segregation analysis of CAE was consistent with an autosomal dominant inheritance with reduced penetrance (Buoni, Grosso et al. 1998).

## 3.5 Candidate genes for CAE

A number of candidate genes and chromosomal regions can be identified for CAE, from studies on the mechanism by which spike-wave seizures are generated, the targets for anti-absence drugs, isolation of genes causing spike-wave seizures in rodents, mendelian epilepsies which include absence seizures as part of the phenotype, and initial linkage and association studies in man.

### 3.5.1 Mechanisms of spike-wave seizures generation

Traditionally spike-wave discharges are considered to be generalised from the onset. However, both frontal and centrencephalic origins have been demonstrated in patients and animal models, with up to a 40ms time lag in the involvement of other cortical regions (Pavone and Niedermeyer 2000).

Evidence for a key role of the thalamus and cortex in the pathogenesis of absence seizures was provided by the demonstration of a focal increase in thalamic blood flow on positron emission tomography (PET) (Prevett, Duncan et al. 1995), and hypermetabolism in cortical and thalamic areas (Iannetti, Spalice et al. 2001) during an absence seizure. Neuronal elements of the thalamocortical loop appear to be crucial for the full expression of spike wave discharges (Crunelli and Leresche 2002) (Figure 3-2). The hypothesis that spike-wave discharges may develop by the same thalamocortical circuits which normally create sleep spindles under conditions of cortical hyperexcitability was proposed by Gloor in 1978 (Gloor 1978). Subsequent electrophysiological studies in vitro and in vivo revealed the role of nucleus reticularis thalami (NRT) neurons and thalamocortical-corticothalamic circuit oscillations in the generation of both spindle rhythms and spike-wave discharges (Kostopoulos 2000).

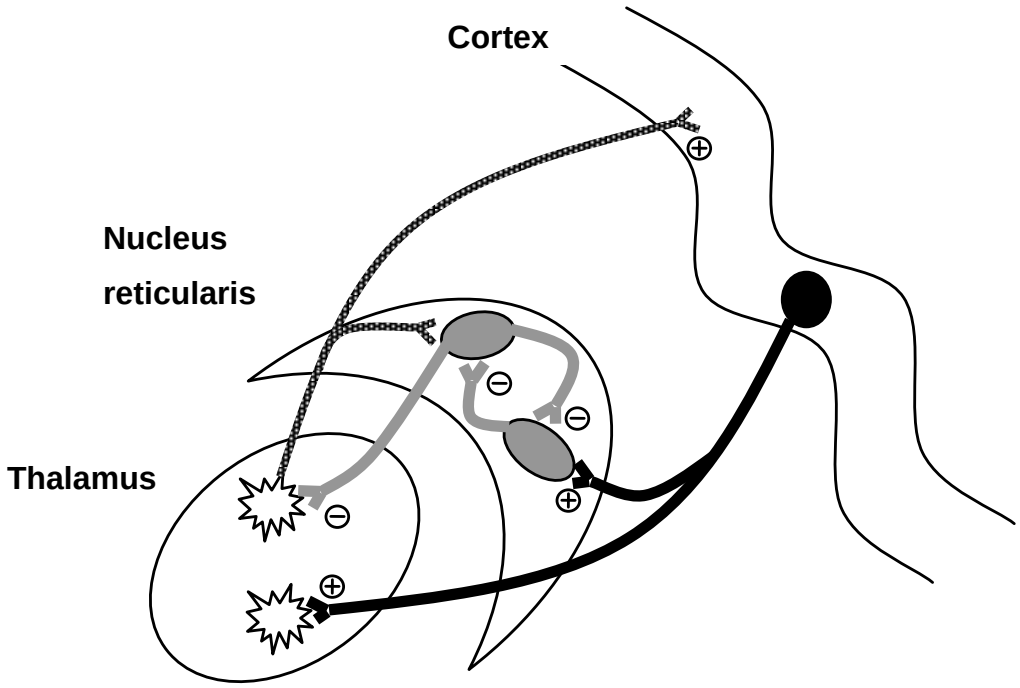



**Figure 3-2** *Thalamocortical excitatory (+) and inhibitory (-) connections involved in the generation of spike-wave discharges*

Animal models demonstrate that during spike-wave discharges, rhythmic depolarisations of cortical neurons occur in phase with the EEG spike. Synchronous cortical volleys of discharges and converging corticothalamic inputs result in bursts of excitatory postsynaptic potentials (EPSPs) in the NRT in vivo (Giaretta, Avoli et al. 1987). Hyperpolarization followed by the EPSP bursts generates a low threshold $Ca^{2+}$ potential and a high frequency burst of action potentials corresponding to each EEG spike. NRT abnormalities may be required for the clinical manifestation of absence seizures, but may not be essential for the generation of spike-wave discharges as they still occur in the ipsilateral cortex of cats after hemithalamectomy, but not in the thalamus after decortication (Steriade and Contreras 1998).

Possible mechanisms for the generation of abnormal spike-wave activities in cortical and thalamic neurons include the loss of $GABA_A$ receptor-mediated inhibition between thalamic reticular cells, strong activation of thalamic GABAergic neurons by corticothalamic or thalamocortical afferents, and increased NMDA (N-methyl-D-aspartic acid)-mediated excitation in deep layers of the cortex or the enhancement of the low threshold $Ca^{2+}$ current (Pumain, Louvel et al. 1992; Leresche, Parri et al. 1998; Pinault, Leresche et al. 1998; Huguenard 1999; Kim, Song et al. 2001; Schuler, Luscher et al. 2001). The role of the low threshold $Ca^{2+}$ current in thalamocortical activity during pharmacologically induced spike-wave discharges is demonstrated by *CACNA1G* (a T-type calcium channel) knockout mice (Kim, Song et al. 2001). These mice are resistant to spike-wave discharges induced by γ-butyrolactone. However the authors considered that this could also be related to reduced low threshold $Ca^{2+}$ potentials in cortical neurons or changes in the steady state low threshold $Ca^{2+}$ current. Investigation of the contribution of $GABA_A$ and $GABA_B$ receptors to thalamic neuronal activity during spontaneous spike-wave discharges in WAG/Rij rats by microiontophoretic application of $GABA_A$ and $GABA_B$ receptor antagonists indicated that (1) $GABA_A$ receptor-mediated events are critical, (2) a contribution from $GABA_B$ receptors is not necessary and (3) $GABA_A$ receptor blockade potentiates spike-wave discharge related activity (Staak and Pape 2001).

### 3.5.2 Animal models of human absence epilepsy

The molecular basis of epilepsy has been explored in the mouse both by identification of genes underlying naturally occurring epilepsy mutants and by exploring the phenotypic consequences of targeting genes which may cause epilepsy on biological grounds (Meisler, Kearney et al. 2001). Six spontaneous AR mouse mutants have spike-wave seizures similar to those occurring in human absence epilepsy. Four of these arise from mutations in voltage-dependant calcium channel (VDCC) subunit genes: *Cacna1a* (tottering) (Fletcher, Lutz et al. 1996), *Cacnb4* (lethargic) (Burgess, Jones et

al. 1997), *Cacng2* (stargazer) (Letts, Felix et al. 1998) and *Cacna2d2* (ducky) (Barclay, Balaguero et al. 2001) (Figure 3-3). VDCCs contain a pore forming α subunit with regulatory β, γ, and α2δ subunits (**Figure 2-1** (c)). In contrast, the slow-wave epilepsy mouse gene, *Slc9a1*, encodes a Na+/H+ exchanger, a ubiquitous membrane-bound enzyme involved in pH regulation of cells (Cox, Lutz et al. 1997). The mocha mouse mutant is hyperactive and has 6-7 Hz spike-wave discharges on EEG. It is a model for Hermansky-Pudlack syndrome in humans, showing coat and eye colour dilution, reduced levels of renal lysosomal enzymes in urine, and prolonged bleeding due to storage pool deficiency in the dense granules of platelets. Mocha is a null allele of the *Ap3d1*, a subunit of the adaptor-protein complex gene. *Ap3d1* is implicated in intracellular biogenesis and trafficking of pigment granules and possibly platelet dense granules and neurotransmitter vesicles (Kantheti, Qiao et al. 1998).

Gene targeting has produced mice lacking the $GABA_A$ receptor β3 subunit implicated in the epilepsy of Angelman syndrome and in CAE by linkage and association, and is illuminating the mechanisms underlying spike-wave seizures. Mice lacking αIG T-type $Ca^{2+}$ channels lack burst firing of thalamocortical relay neurons and are resistant to generation of spike-wave discharges in response to $GABA_B$ receptor activation. Spontaneous absence seizures with generalised spike-wave discharges at 5Hz are exhibited by mice with targeted disruption of *HCN2*, which encodes a hyperpolarisation-activated cation channel involved in the generation of pacemaker depolarisations in the heart and the control of neuronal excitability.

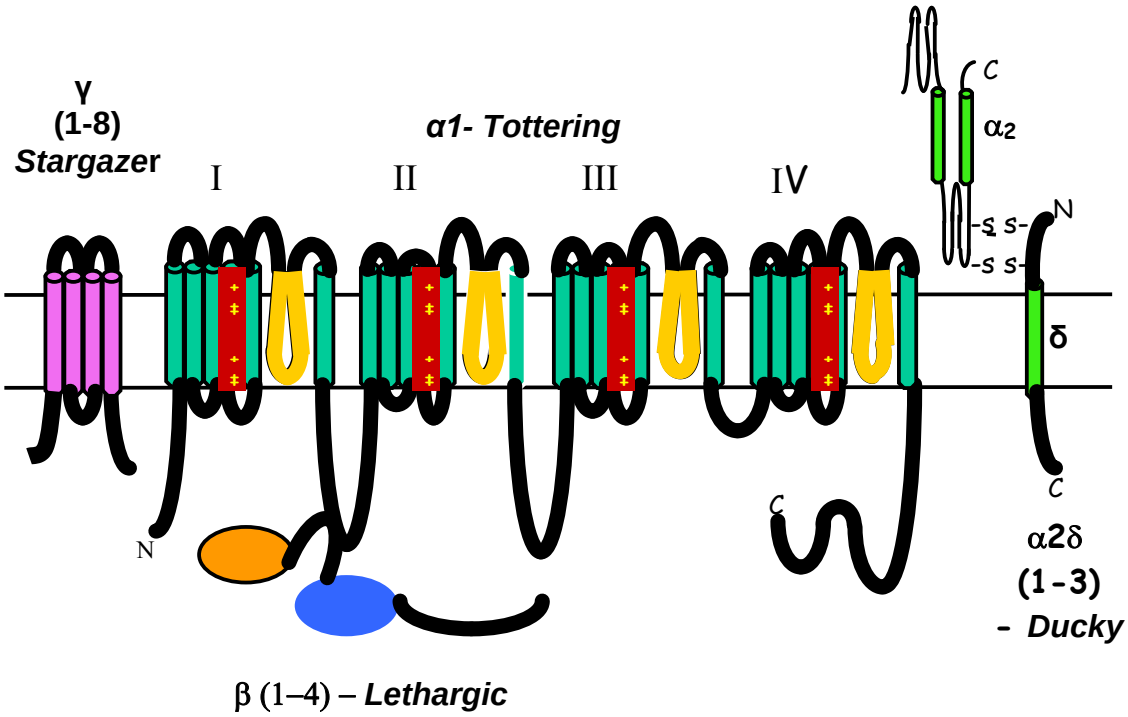



**Figure 3-3** *Voltage-dependent calcium channel subunits and associated mouse models of human spike-wave epilepsy*

### 3.5.3 Targets of anti-absence drugs

Ethosuximide is a drug specifically effective in the control of absence seizures. Application of ethosuximide to thalamocortical and nucleus reticularis thalami neurones results in a 40% reduction in the amplitude of the T-type $Ca^{2+}$ current (Coulter, Huguenard et al. 1989; Coulter, Huguenard et al. 1990). This has implicated an action of ethosuximide on thalamic T-type calcium channels. However other studies have failed to demonstrate an effect of ethosuximide on native T-type $Ca^{2+}$ currents in thalamic and nonthalamic neurones (Sayer, Brown et al. 1993; Leresche, Parri et al. 1998). Ethosuximide has been observed to decrease the persistent $Na^{+}$ and sustained $K^{+}$ currents in thalamocortical and cortical neurones (Leresche,

Parri et al. 1998). Recent evidence also suggests that targeting of the thalamus alone may be insufficient, and that actions in the cortex may be required for a full anti-absence action for ethosuximide (Richards, Manning et al. 2003).

### 3.5.4 Mendelian epilepsies

Mutations in *GABRG2* cause both GEFS+, a phenotype which includes absence seizures, and a phenotype of childhood absence epilepsy and febrile seizures (section 2.4.6). A novel *CACNA1A* mutations have been identified in a boy with progressive and episodic ataxia, learning difficulties and absence epilepsy (Hanna, Jouvenceau et al. 2001). A heterozygous mutation in the VDCC gene *CACNA1A* has previously been reported in a child with early onset absence epilepsy and cerebellar ataxia (Jouvenceau, Eunson et al. 2001). The mutation gave rise to a premature stop codon, with complete loss of function of the heterologously expressed mutated CaV2.1$\alpha$1 subunit. A further mutation in *CACNA1A*, predicted to result in a frameshift and premature stop codon, has been reported in association with EA2 and unspecified epilepsy (Jen, Kim et al. 2004).

### 3.5.5 Linkage and association studies of human absence epilepsy

Linkage analysis of a five generation family from Bombay in which affected patients had a persisting form of CAE provided evidence for a locus on chromosome 8q24 (Fong, Shah et al. 1998). The candidate region for this locus, designated *ECA1*, has been refined, but a gene remains to be identified (Sugimoto, Morita et al. 2000). Study of another extended pedigree in which affected individuals manifested both CAE and febrile seizures revealed a linked marker on chromosome 5 close to a cluster of genes

encoding $GABA_A$ receptor subunits (Wallace, Marini et al. 2001). A mutation was found in *GABRG2* which changes a conserved amino acid and appeared to contribute to the CAE phenotype. Mutations in this gene have also been identified in a family segregating generalised epilepsy with febrile seizures plus (GEFS+) (Baulac, Huberfeld et al. 2001).

A possible association was documented between a polymorphism in *GABRB3* and patients with CAE (Feucht, Fuchs et al. 1999), and suggestive linkage to this gene found in 8 families (Tanaka, Castroviejo et al. 2000). Mutation screening of 45 CAE patients identified no functionally relevant mutations in *GABRB3* (Urak, Feucht et al. 2006). However a SNP haplotype in the region between the promoter and intron 3 showed significant association with the CAE phenotype, and reporter gene assays in the human neuronal-like cell line NT2 suggested that the disease-associated haplotype causes reduced transcriptional activity. Heterozygous mutations in *GABRB3* have been identified in four out of 48 families with CAE from Mexico and Honduras, and all mutations demonstrated hyperglycosylation and reduced whole cell GABA-evoked current density when expressed in vitro (Tanaka, Olsen et al. 2008). *GABRB3* maps to 15q11-q13, the region deleted in Angelman syndrome (Wagstaff, Knoll et al. 1991), and mice with targeted disruption of the *Gabrb3* gene have the epilepsy phenotype and behavioural characteristics of Angelman syndrome (DeLorey, Handforth et al. 1998).

An association has also been documented between polymorphisms in *CACNA1A* and IGE including CAE (Chioza, Wilkie et al. 2001).

Direct sequencing of *CACNA1H*, a T-type $Ca^{2+}$ channel gene, in a Chinese population identified 12 missense mutations that were only found in patients with CAE (14 of 118 patients) and not in any of 230 unrelated controls (Chen, Lu et al. 2003). The mutations occurred in highly conserved residues of the calcium channel gene, and are likely to be functionally significant. However each mutation was also identified in an apparently unaffected parent of each of the 14 affected children, suggesting that *CACNA1H* is one of perhaps several susceptibility genes. Evaluation of Caucasian European patients with CAE found no evidence of linkage to *CACNA1H* in 44 pedigrees and no

*CACNA1H* variants were identified in 220 unrelated patients (Chioza, Everett et al. 2006). *CACNA1H* was subsequently screened in 240 epilepsy patients (Heron, Khosravani et al. 2007). Over 100 variants were detected, including 19 missense mutations in subjects with childhood absence, juvenile absence, juvenile myoclonic and myoclonic astatic epilepsies, as well as febrile seizures and temporal lobe epilepsies. Nine of eleven variants tested altered channel properties. The authors proposed that the variants identified contribute to an individual's susceptibility to epilepsy but are not sufficient to cause epilepsy on their own.

Linkage analysis in 65 CAE pedigrees found suggestive evidence for linkage to *CLCN2* with heterogeneity (Everett, Chioza et al. 2007). Mutation analysis identified 45 sequence variants in *CLCN2* and intrafamilial association analysis found suggestive evidence for transmission disequilibrium with the CAE phenotype.

# *Chapter 4 Identification of Human Disease Genes*

## 4.1 Introduction

Since the early 1980's, advances in molecular biological techniques have provided a range of approaches to the investigation of the molecular basis of human inherited disease. Prior to this, very few human disease genes had been identified, mainly for conditions in which the underlying biochemical mechanisms had been elucidated. Modern techniques involving recombinant DNA technology and the polymerase chain reaction (PCR) have been applied with most success to mendelian diseases, and strategies for tackling these disorders are now well established. These include positional cloning, candidate gene identification and mutational analysis, which have allowed the identification of several mendelian epilepsy genes. The complex genetic diseases, however, have generally proved resistant to these strategies. The completion of the human genome mapping project and progress in the construction of a haplotype map (HapMap) of the human genome has allowed new approaches to the investigation of complex diseases.

## 4.2 Evolution of Molecular Genetics

### 4.2.1 Mendelian and non-mendelian inheritance

Gregor Mendel (1822-1884) proposed a theory of particulate inheritance, in which characteristics were determined by discrete units of inheritance (later known as genes) that were passed from one generation to the next. Mendel's experiments involving the cultivation of pea plants allowed him to formulate his laws of heredity.

### 4.2.2 Mendel's law of segregation

This introduces the concept of genes and alleles (alternative versions of genes). These terms were not actually used by Mendel, but the essence of what Mendel discovered was:

1. Alternative 'alleles' account for variations in inherited characters.

2. For each character, an organism inherits two 'genes', one from each parent.

3. If two alleles differ (a heterozygote), then one, the 'dominant allele', is fully expressed in the organism, whilst the other, the 'recessive allele', has no noticeable effect.

4. The two genes for each character segregate during gamete production (meiosis).

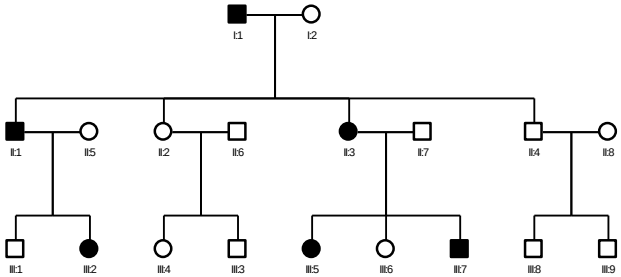
### 4.2.3 Mendel's law of independent assortment

The most important principle of this law is that the emergence of one trait will not affect the emergence of another. This results in the observed 3:1 ratio between dominant and recessive phenotypes for a single trait, whilst phenotypes comprising two traits show a 9:3:3:1 ratio.

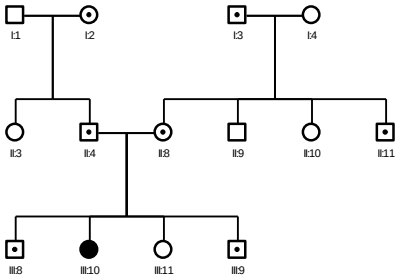
### 4.2.4 Mendelian patterns of inheritance

Autosomal or X-linked traits can be dominant or recessive. This does not apply to Y-linked traits as there is at most one Y chromosome. Males are hemizygous for loci on the X and Y chromosomes. There are therefore five basic mendelian pedigree patterns: i) autosomal dominant, ii) autosomal recessive, iii) X-linked recessive, iv) X-linked dominant, v) Y-linked.

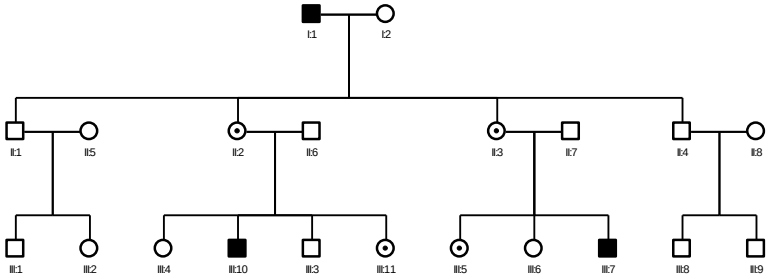
i) Autosomal dominant



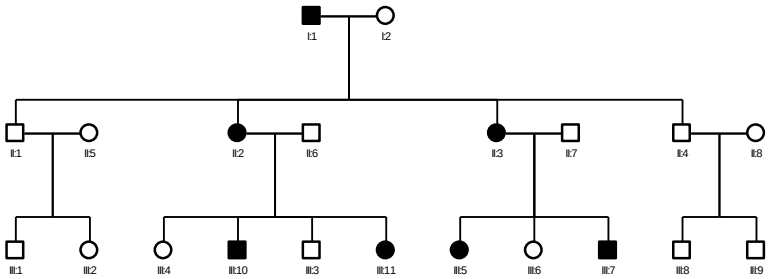

ii) Autosomal recessive




iii) X-linked recessive




iv) X-linked dominant




v) Y-linked

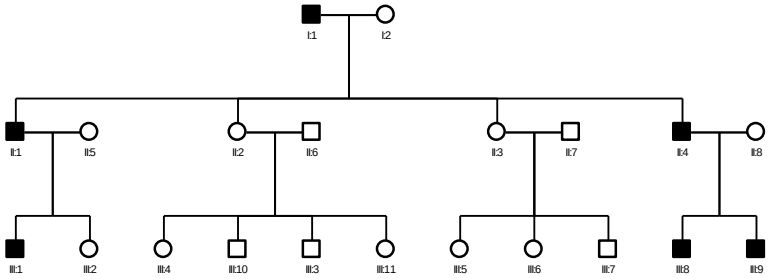



**Figure 4-1** *Basic mendelian pedigree patterns*

### 4.2.5 Non-mendelian ‘complex’ inheritance

By 1900, Mendel’s work had been rediscovered and formed the basis of the Mendelian school of genetics. However, a rival school, based on the tradition of biometrics developed by Francis Galton, argued that continuously variable quantitative variable characters (such as height, strength, intelligence) could not be explained by mendelian genetics and were not inherited in a mendelian fashion. In 1918, RA Fisher demonstrated that traits with quantitative variation and complex inheritance patterns could be explained by the interaction of many independent mendelian factors (polygenic inheritance). This applies to most clinical syndromes which are determined by a complex interaction of factors involved in physiological or developmental processes. The genetic architecture of complex diseases may comprise a few interacting loci (oligogenic), many loci each with small effect (polygenic), or the interaction of genetic and environmental factors (multifactorial). Non-penetrance or variable expression of an apparently dominant disease, where different ‘affected’ family members show different clinical features (or no features at all), is likely to be a reflection of a single major causative locus interacting with modifying loci or environmental factors.

### 4.2.6 DNA/RNA and polypeptides

In 1943, Oswald Avery discovered evidence that molecules of DNA (deoxyribonucleic acid) store the genetic information which determines the characteristics of living organisms. James Watson and Francis Crick at the Cavendish Laboratories in Cambridge, using the x-ray crystallography data generated by Maurice Wilkins and Rosalind Franklin at King’s College London, determined the double helix structure of DNA in 1953. They demonstrated that the four bases, adenine, guanine, cytosine and thymine, which form the genetic code, are arranged in hydrogen-bonded pairs of adenine-thymine and guanine-cytosine which hold the two strands together. This structure explained how a gene could replicate, as the double helix could split and each strand on its own could build its counterpart from free bases, thus forming two new but identical DNA molecules.

By 1960, it was still not clear how the basic genetic information encoded by DNA was expressed. It was known that the polypeptide proteins that determine a cell's physiology, morphology and development are encoded by individual genes in chromosomes. In 1960, Francois Jacob and Matthew Meselson determined that proteins are assembled on ribosomes in the cytoplasm of the cell. The following year, Jacob and Jacques Monod hypothesised that a special form of RNA synthesised from the DNA template of genes is transported to the ribosome where this messenger RNA (mRNA), complementary to the genetic DNA sequence, provides the information for polypeptide synthesis. This mRNA hypothesis was confirmed by Sydney Brenner, Jacob, and Meselson in 1964. It had previously been demonstrated in 1961, by Francis Crick and co-workers, that the actual instructions for a polypeptide exist as a series of overlapping, three-base code words, each "triplet" specifying one of the 20 amino acids.

### 4.2.7 Mapping of genes

Genetic mapping relies on the principle that the segregation of two loci during meiotic recombination depends on the distance between them. If two loci are on different chromosomes they will segregate independently. If they lie on the same chromosome, they will segregate together unless a cross-over occurs between the two loci during meiosis. The further apart the two loci are on a chromosome, the more likely it is that a crossover will separate them. The recombination fraction is a measure of the genetic distance between the two loci, and this genetic distance can be used to estimate the physical distance using a mapping function. Genetic linkage studies measure the genetic distance between markers of known chromosomal location and a putative disease locus.

In 1911, Thomas Huntley Morgan first used the principles of linkage analysis to propose that the genes for white eyes, yellow body, and miniature wings in Drosophila are linked together on the X chromosome. Two years later, A. H. Sturtevant, an undergraduate working with Morgan at Columbia, provided the experimental basis for the linkage concept in Drosophila and produced the

first genetic map. Early linkage studies in humans used blood group and serum protein (with variant electrophoretic mobility) markers to determine the chromosomal location of disease genes. Mapping disease genes more accurately became a realistic possibility in the 1980s with the discovery that variations in human DNA could be used as genetic markers in linkage studies. The first genetic linkage map in humans, using restriction-fragment length polymorphism (RFLP) markers was published in 1980 (Botstein, White et al. 1980). These were typed by hybridizing southern blots of restriction digests with radiolabelled probes. Subsequently, abundant highly polymorphic short tandem repeat DNA sequences were discovered which enabled the mapping of a larger number of mendelian disease loci (Weber and May 1989).Disease genes could then be identified by positional cloning (section 4.5.2)

Linkage analysis has proved extremely powerful when applied to mendelian phenotypes. It has also led to the identification of genes responsible for rare subsets of more common familial disorders such as diabetes, alzheimer's disease and breast cancer. However, the majority of genes for the common 'complex' diseases have proven resistant to linkage analysis, due to the probable existence of multiple genes with small effect and weak genotype-phenotype correlations. Approaches using candidate genes or high-density genome scans dependent on linkage disequilibrium may prove more successful.

## 4.3 Molecular genetic techniques

### 4.3.1 Polymerase chain reaction(PCR)

PCR is an extremely powerful method for synthesizing and amplifying a specific DNA sequence within a source of DNA in vitro. It requires the presence of a heat stable DNA polymerase for synthesis of new DNA strands. The theory of PCR was first described in 1971 by Kleppe (Kleppe, Ohtsuka et al. 1971). However, the first heat-stable DNA polymerase was not discovered until 1975, when taq polymerase was isolated from the thermus aquaticus microorganism in the hot springs of Yellowstone National Park. Large scale automation of DNA amplification was then developed in 1983, when Kary Mullis, working at the Perkin-Elmer Cetus Corporation, realised its immense potential (Saiki, Scharf et al. 1985). To enable selective amplification, the DNA sequence flanking the target region must be known. This is used to construct two oligonucleotide primer sequences, which will bind to complementary DNA sequences flanking the target region, when added to denatured genomic DNA. To allow the synthesis of new DNA, the four deoxynucleoside triphosphates (dATP, dCTP, dGTP and dTTP) must be present in the reaction.

The PCR reaction involves three steps repeated sequentially: denaturation, reannealing and DNA synthesis (Figure 4-2). Each cycle results in doubling of the target DNA present.

PCR has a wide range of applications, including typing genetic markers, mutation screening, DNA cloning and DNA sequencing.

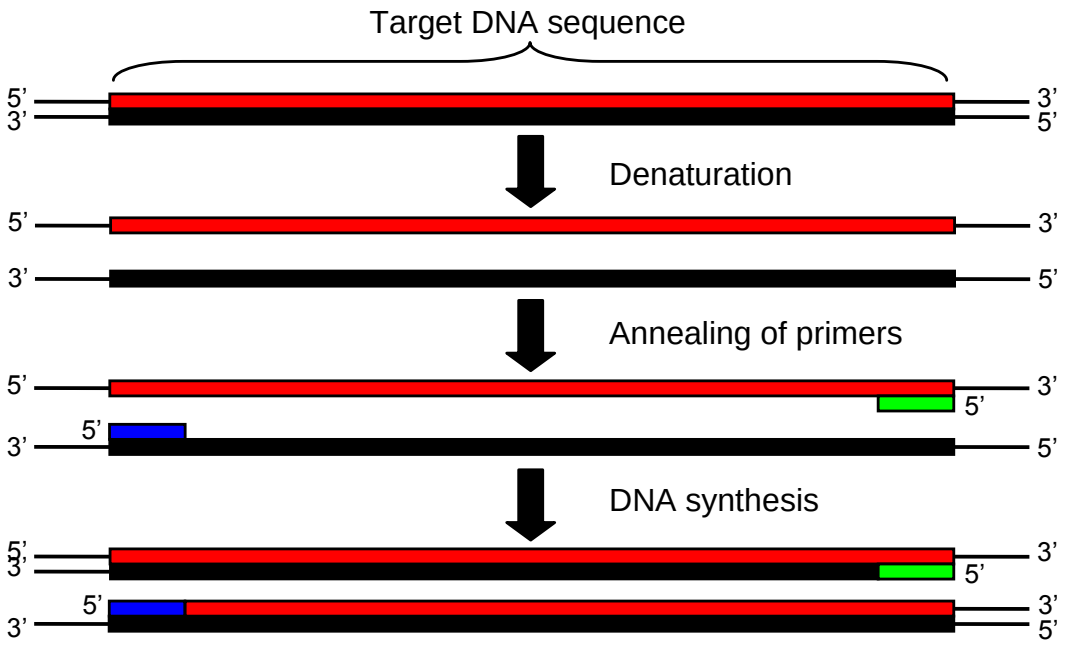



**Figure 4-2** *Polymerase chain reaction*

### 4.3.2 Genetic marker typing

Restriction site polymorphisms (RSPs) are biallelic markers, either possessing or lacking a specific restriction site. Conventionally, these are assayed by digesting genomic DNA with the restriction endonuclease specific for that restriction site. This results in fragments whose length depends on the presence or absence of the restriction site. These fragments are then size-fractionated by agarose gel electrophoresis, with smaller fragments travelling more rapidly and therefore further through the gel. The DNA fragments are detected by Southern blot hybridisation. This involves denaturing the DNA fragments in alkali, transfer by blotting onto a nitrocellulose membrane, hybridising with a radiolabelled single stranded DNA probe and detection of the radiolabelled fragments using a x-ray film. As a an easier alternative, RSPs can be typed by using PCR to amplify the region containing the restriction site and then cutting the PCR product with

the restriction enzyme. PCR amplifies enough of the target DNA sequence to allow visualization of the fragments following agarose gel electrophoresis.

Simple sequence length polymorphisms (SSLPs) are much more powerful than RSPs when used as genetic markers, as they tend to be highly polymorphic. An SSLP consists of a short nucleotide sequence that is tandemly repeated a variable number of times. Each allele of a particular marker is determined by the number of repeat sequences. The most commonly used SSLPs are dinucleotide (e.g. (CA)n), trinucleotide or tetranucleotide repeats. Typing of SSLPs involves PCR amplification using primers flanking the repeat region. The size of the PCR product indicates the particular allele of the SSLP. By fluorescent labelling of one of the PCR primers, the PCR products can be separated by polyacrylamide gel electrophoresis and detected by laser excitation.

### 4.3.3 DNA sequencing

The basis for virtually all DNA sequencing is the dideoxy-chain terminating reaction, developed by Sanger in 1975. A sequencing reaction consists of a single stranded DNA template (multiple copies) and a complementary sequencing primer which binds to the template and flanks the region to be sequenced. DNA polymerase and the four deoxynucleoside triphosphate (dNTP) DNA precursors are added for synthesis of complementary DNA strands. In addition, a low concentration of one of four dideoxynucleoside triphosphates (ddNTPs) is included which correspond to the four dNTPs. These ddNTPs act as terminators of DNA chain synthesis. These are incorporated normally into the growing DNA chain, but as they lack a hydroxyl group at the 3’ carbon position, they cannot form a phosphodiester bond with the 5’ carbon of the next dNTP in the sequence, and DNA synthesis terminates. As the DNA sequencing reaction proceeds, chain termination occurs randomly whenever a ddNTPs is incorporated instead of the corresponding dNTP. This results in a population of DNA strands of varying lengths, each terminating at the 3’ end with the particular ddNTP in the reaction. Traditionally, four parallel reactions are run, each with one of the

four ddNTPs. The products are separated on a denaturing polyacrylamide gel. By incorporating radioisotope labelled nucleotides, the band positions can be detected by exposure to x-ray film.

Modern sequencing reactions use fluorescent labelling of ddNTPs. By labelling each of the four ddNTPs with a different coloured label, they can be combined in a single sequencing reaction and the products separated in a single lane. As the laser causes the dyes to fluoresce, each fluoresces at a different wavelength and can be detected automatically and converted into sequence information (Figure 4-3)

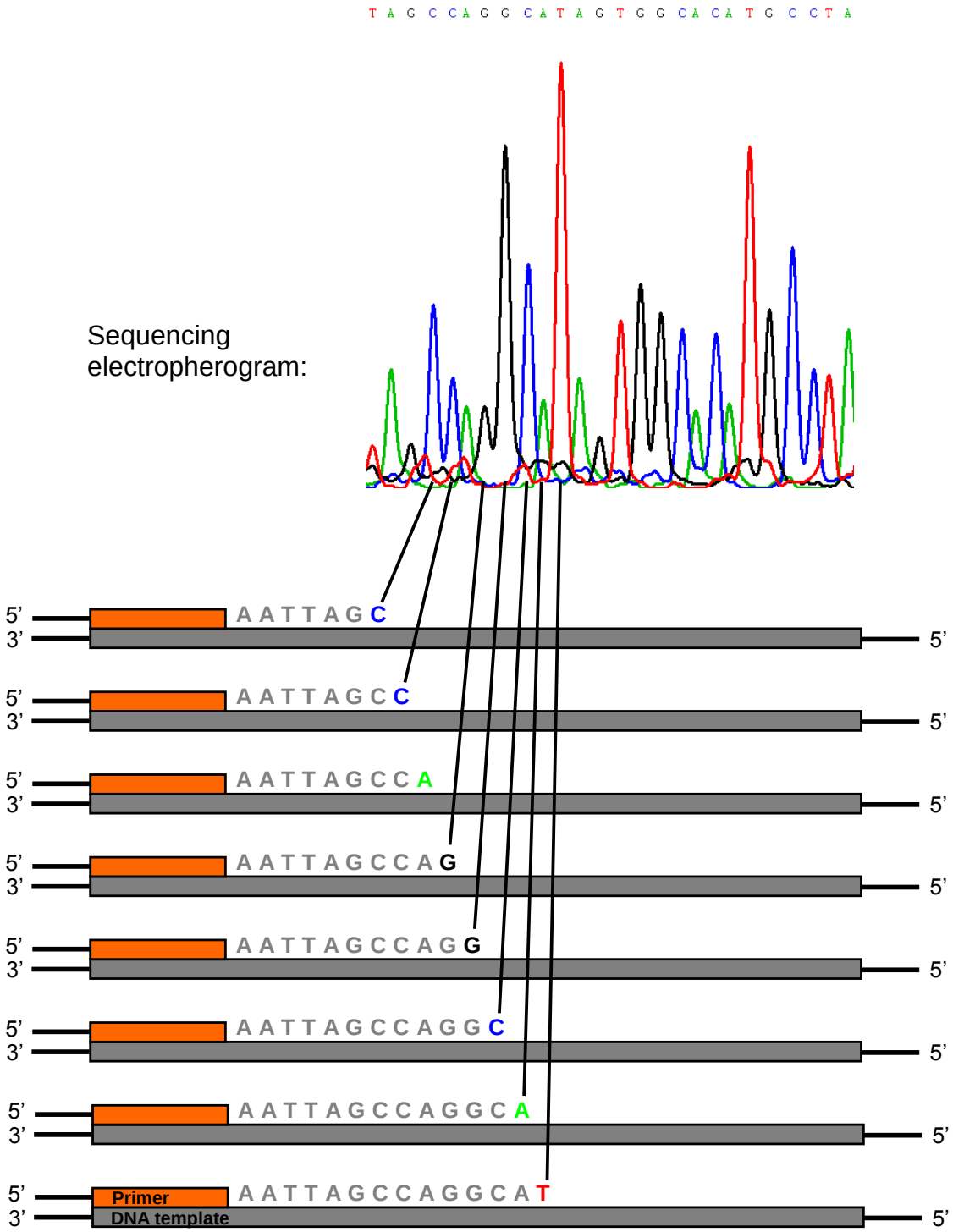



**Figure 4-3** *Automated DNA sequencing using fluorescent primers. Random incorporation of the labelled ddNTPs produces a series of DNA fragments in which chain growth has been terminated at each successive position.*

### 4.3.4 Single nucleotide polymorphism (SNP) genotyping and microarrays

SNPs are DNA sequence variations that occur when a single nucleotide is altered. A variation can (and probably does) occur at any of the approximately 3,000,000,000 nucleotides in the human genome, but to be considered a SNP, it must occur in at least 1% of the population. SNPs occur every 100 to 300 bases along the human genome, accounting for about 90% of all human genetic variation. SNPs and SNP haplotypes can be used as genetic markers for genome-wide and candidate-gene association studies. The availability of high density SNP maps with knowledge of the haplotype structure of the human genome will provide an extremely valuable tool for the investigation of complex genetic diseases. In 1999 the SNP consortium (TSC) was established between pharmaceutical companies and the Wellcome Trust, with the aim of mapping 300,000 common SNPs. By February 2003, 3.7 million SNPs had been mapped. In 2002 the international HapMap project was established to develop a haplotype map of the human genome and to identify which 200,000 to 1 million tagging SNPs provide almost as much mapping information as the 10 million SNPs in total. The phase II HapMap published in 2007 characterised over 3.1 million human SNPs genotyped in 270 individuals from four geographically diverse populations and include 25–35% of common SNP variation in the populations surveyed.

A number of SNP genotyping methods have been developed. An ideal method should be easy, quick, robust, flexible and automated for high throughput. Important aspects of methodology include allelic discrimination, assay format, and detection technique (Kwok 2001). Some illustrative methodologies will be briefly described. All depend on PCR amplification of the genomic DNA prior to genotyping. Melting curve SNP analysis (McSNP) is described in section 6.7.1

### 4.3.5 Allele specific hybridisation

Techniques such as that developed by Affymetrix and the Whitehead institute involve hybridisation on a GeneChip microarray. Allele specific hybridisation uses allele specific probes which only anneal to the target DNA when there is a perfect match. The allele specific probes are immobilized on a solid support, and after PCR amplification of the target DNA and hybridization to the microarray, the unbound targets are washed away and the hybridised DNA visualised by fluorescent labelling. The PCR assays allow a degree of multiplexing, but the detectable SNPs are limited to those on the pre-designed microarray.

### 4.3.6 Dynamic allele specific hybridisation (DASH)

This method investigates allele-specific hybridisation by using the difference in melting temperature between annealing of the allele-specific probe to matched and mis-matched DNA. Hybridisation is monitored over a range of temperatures and formation of double stranded DNA detected using a DNA intercalating dye, such as Syber Green I, which fluoresces in the presence of double-stranded DNA. Assays can be designed for almost all SNPs. However, the PCR amplified target DNA must be rendered single-stranded by used of biotinylated PCR primers which allow the product to be captured on a solid support, denatured, and the unlabelled strand washed away.

### 4.3.7 Homogeneous hybridisation with fluorescence resonance energy transfer (FRET)

Techniques such as Molecular Beacon and TaqMan (section 6.7.1) use the principle of fluorescence resonance energy transfer. A fluorescent reporter is held in close association with a universal quencher such that fluorescence is only observed when the reporter and quencher are separated. The reporter and quencher can be held by a molecular beacon which is a stem-loop structure in which binding to target DNA exactly complementary to the sequence of the loop portion forces open the stem and allows fluorescence.

A single base mismatch reduces the strength of hybridisation so that the stem-loop structure remains intact. In both the Molecular Beacon and TaqMan assays, the increase in fluorescence can be monitored in real time, or two differently labelled allele specific probes can be used, and the signal ratio of the two probes measured at completion of the PCR.

### 4.3.8 DNA polymerase-assisted genotyping

Minisequencing, or single nucleotide primer extension, uses a primer that anneals to the target DNA immediately adjacent to the SNP. Primer extension by DNA polymerase only occurs with the nucleotide complementary to the nucleotide at the site of the SNP, and this is detected by differential fluorescent labelling of the four ddNTPs. Modifications include the use of HAPTEN-labelled nucleotide analogues that are detected colometrically in an ELISA (enzyme-linked immunosorbent assay), and the Amplifluor® SNPs Genotyping System is based on competitive allele-specific PCR, using two fluorescently labelled Amplifluor® SNP Primers in a one-step, single-tube reaction with standard PCR. Pyrosequencing detects the pyrophosphate which is released whenever a dNTP is incorporated during primer extension (see section 6.7.3). Mass spectrometry or MALDI-TOF (matrix-associated laser desorption time-of-flight mass spectrometry) can be used to detect the primer extension products. This can be an extremely sensitive method, and allows multiplexing, but is expensive.

As with allele specific hybridisation, single nucleotide primer extension can be used with microarray technology for parallel genotyping of SNPs. The power of discrimination between genotypes appears to be significantly greater than for allele specific hybridisation. Minisequencing on microarrays uses one primer for each SNP immobilised on a solid support. After primer extension with labelled ddNTPs, the genotypes are detected by fluorescence scanning. An alternative microarray method is allele-specific primer extension, where two allele-specific primers per SNP are immobilised and extended with fluorescent oligonucleotides and templates of target DNA (or RNA created by RNA polymerase and then used as templates with reverse

transcriptase). Microarrays can also be used in conjunction with cyclic nucleotide primer extension reactions in solution using tagged primers and complementary oligonucleotide tags on the microarray. The microarray is used to capture and sort the minisequencing products by hybridisation. Tagged primers can also be used with arrays of microspheres, where different classes of microspheres, each with a specific tag, are detected by fluorescence in a flow cytometer, and the genotype of the SNP is determined by the captured fluorescent product.

### 4.3.9 Ligation methods

Methods such as the oligonucleotide ligation assay utilise DNA ligases which only ligate hybridised probes and target DNA when there is a perfect match. This can be combined with colimetric detection methods, fluorescently labelled ligation probes and microarray formats.

## 4.4 Genome Mapping Projects

### 4.4.1 Human genome project

The human genome project was a collaboration involving 20 groups from the United States, United Kingdom, Japan, France, Germany and China. In 1988 the US national Research Council published a report entitled ‘Mapping and Sequencing the Human Genome’. This recommended a programme to include: the creation of genetic, physical and sequence maps of the human genome; similar projects in model organisms (bacteria, yeast, worms, flies and mice); the development of supportive technology; research into related ethical, legal and social issues. Genomic research programmes were launched by the Department of Energy and the National Institute of Health in the US, the Medical Research Council and the Wellcome Trust in the UK, as well as other bodies in Europe and Japan. By 1990 the Human Genome Project had been launched as a 15 year project and the Human Genome Organisation (HUGO) to coordinate international genome research. A draft genome sequence was published in February 2001, covering about 94% of the human genome. The sequencing of the human genome was declared complete in April 2003 (99% of gene-containing part of human sequence finished to 99.99% accuracy). However the completed sequences of individual chromosome continue to be released until 2006. By March 2003, 15,000 full-length human cDNAs had been identified and these could be compared with those of other organisms with sequenced genomes. The total number of human genes was less than predicted at around 25,000, with alternative splicing resulting in around 100,000 proteins.

### 4.4.2 Model organisms

Comparison of genome sequences from evolutionarily diverse species is a powerful tool for identifying functionally important conserved genomic elements, defining the functional nature of some sequences, establishing

relationships between genotype and biological function, providing insights into the distinct anatomical, physiological and developmental features of different organisms, defining the genetic basis for speciation and facilitating the characterization of mutational processes.

The first multicellular eukaryotic genome to be completed was the Caenorhabditis elegans genome, carried out by the Sanger Centre and the Genome Sequencing Centre at Washington University (C. Elegans Sequencing Consortium 1998). The sequence of the Drosophila melanogaster genome, originally determined in a collaboration between Celera and the Berkeley Drosophila Genome Project, was published in 2000 (Adams, Celniker et al. 2000). The results of an international collaboration to produce a high-quality draft sequence of the mouse genome were published in 2002, as well as an initial comparative analysis of the mouse and human genomes (Waterston, Lindblad-Toh et al. 2002). The complete genome sequences of Plasmodium falciparum (Gardner, Hall et al. 2002), the main cause of human malaria, Plasmodium yoelii (Carlton, Angiuoli et al. 2002), the infectious agent in rodent malaria and the malaria mosquito Anopheles gambiae (Holt, Subramanian et al. 2002), were published simultaneously in October 2002. As of June 2003, the Rat Genome Sequencing Consortium, led by the Baylor College of Medicine Human Genome Sequencing Centre, had released a draft genome assembly of the Rattus norvegicus genome covering more than 90% of the estimated 2.8 Gb genome. In February 2001 the Sanger Institute started sequencing the genome of the zebrafish (Danio rerio) following two strategies: clone mapping and sequencing from BAC (bacterial artificial chromosome) and PAC (P1-derived artificial chromosome) libraries and whole genome shotgun sequencing with subsequent assembly. It is expected that 90% of the zebrafish genome will be available by the end of 2009.

### 4.4.3 The future of genome research

A number of future projects for genome research were proposed following the completion of the human genome project (Collins, Green et al. 2003).

These challenges were related to three major themes - genomics to biology, genomics to health and genomics to society, and to six elements – resources; technology development; computational biology; training; ethical, legal and social implications; and education. Within the genomics to biology theme, the key targets included: i) identifying all structural and functional components, including protein-coding and non protein-coding regulatory sequences encoded in the human genome (the aim of the Encyclopaedia of DNA Elements (ENCODE) Project launched by the National Human Genome Research Institute); ii) elucidating how genes and gene products interact and how the contribute to functioning of cells and organisms; iii) developing a complete description of the genetic variation in the human genome and its relationship to disease and response to drugs (for example the SNP Consortium and the International HapMap project); iv) understanding evolutionary mechanisms; and v) disseminating genome information as effectively as possible. Much of this work is ongoing.

## 4.5 Identifying mendelian disease genes

Positional cloning, including positional candidate gene approaches, together with mutational analysis has been particularly successful in the identification of mendelian disease genes. Diseases with ‘complex’ inheritance have generally required alternative approaches, such as association analysis.

### 4.5.1 Functional Cloning

A disease gene is identified on the basis of information about its function, such as the biochemical abnormality underlying the disease. Historically, isolation of a purified gene product has allowed amino acid sequencing and identification of regions with minimal codon degeneracy (ie relatively few codon permutations can encode the amino acid sequence). This enabled the synthesis of labelled oligonucleotides corresponding to all possible codon permutations, which were then used to screen cDNA libraries, identify a cDNA clone, and screen a genomic DNA library to fully characterize the gene. This approach was used to identify the factor VIII gene responsible for Haemophilia A. An alternative approach involves raising antibodies against the normal protein product, and then using the antibodies to screen expression cDNA libraries.

### 4.5.2 Positional Cloning

This approach uses linkage analysis to map a gene locus to a small chromosomal region. Linkage analysis tests polymorphic genetic markers distributed across the genome (a ‘genome-wide’ scan) for co-segregation (linkage) with the disease phenotype. A single large pedigree or several small pedigrees are required. A ‘linked’ marker has undergone few recombinations with the causative gene during meiosis and therefore lies on the same chromosomal segment (Figure 4-4). A physical map of a linked

chromosomal region can then be constructed, and coding DNA sequences identified for mutation screening. Linkage analysis as a method for mapping disease genes in man was revolutionised by the development of methods for detecting polymorphism at the DNA level as well as a comprehensive genetic marker map of the human genome consisting of 5,264 simple sequence length polymorphisms (Dib, Faure et al. 1996).

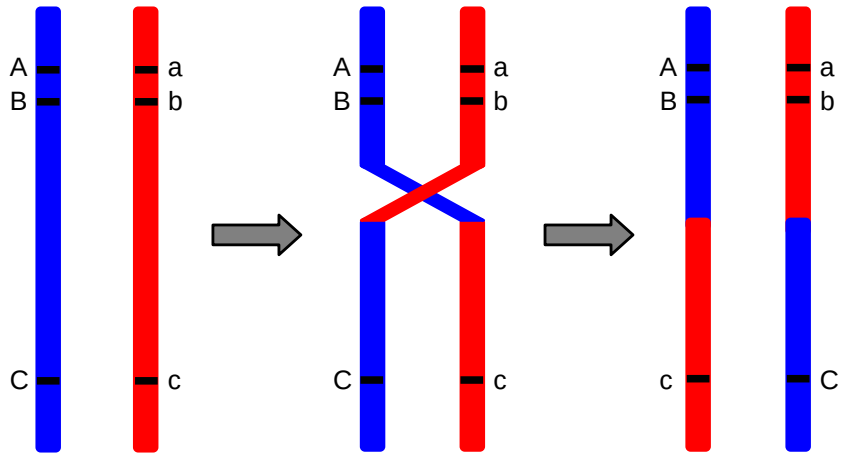



**Figure 4-4** *Linkage analysis. During prophase of meiosis 1, pairs of homologous chromosomes synapse and exchange segments. Alleles at 3 loci are shown - A, B, C on the blue (paternal) chromosome and a, b, c on the red (maternal) chromosome. Recombination between closely spaced loci (A and B) is likely to occur much less frequently than between well separated loci (A and C). If A is the disease gene and B and C are polymorphic genetic markers, linkage analysis allows the disease gene to be mapped relative to the markers B and C.*

Positional cloning has led to the identification of disease genes for a large number of mendelian disorders including several epilepsies. There has been some limited success in non-mendelian disorders such as Crohn disease (Hugot, Chamaillard et al. 2001; Ogura, Bonen et al. 2001). However, the statistical power of this approach is greatly reduced by genetic heterogeneity, uncertainty concerning the mode of inheritance and penetrance of disease alleles, and several other confounding factors.

### 4.5.3 Genetic Mapping

The genetic distance between loci is defined by the recombination fraction. A recombination fraction of 1% signifies a crossover occurring in 1% of meioses. Genetic distance is measured in centimorgans (cM), ie two loci that show 1% recombination are 1 cM apart on a genetic map. However the relationship between recombination fraction and genetic map distance is not linear, as recombination fraction cannot exceed 50% (two loci which segregate independently will be recombinant on average 50% of the time). Recombination fraction ($\theta$) and genetic distance in centimorgans (D) are related by a mapping function. A commonly used mapping function is the Kosambi function (Kosambi 1944), which takes into account interference (inhibition of chiasma formation by the presence of a nearby previous chiasma):

$$D = 1/4\ln\{(1+2\theta)/(1-2\theta)\}$$

### 4.5.4 Genetic Markers

The development of human genetic markers is described in section 4.2.7 and section 4.3.2. For a particular marker, the proportion of informative meioses (those meioses for which the segregation of alleles can be determined) for linkage analysis depends on the informativeness of the marker. A meiosis

can only be informative if the individual is heterozygous for a particular marker. The heterozygosity of a marker indicates the proportion of individuals within a population heterozygous at the marker locus. The polymorphism information content (PIC) refers to the proportion of informative meioses. A PIC of 0 indicates a marker that is never informative and a PIC of 1 indicates a marker that is always informative.

### 4.5.5 Linkage analysis

The likelihood that two loci, or a marker and disease are linked is tested using two-point LOD score analysis (Morton 1955). The LOD score, Z, is the logarithm of odds that the loci are linked with a particular recombination fraction, $\theta$, rather than unlinked ($\theta$=0.5). Each meiosis is considered in turn to calculate the overall LOD. The LOD score is calculated over a range of $\theta$ values to find the maximum values. The traditional criteria for positive linkage for a mendelian trait is Z=3.0 (corresponding to odds of 1000:1), with a 5% chance of error. The criteria for exclusion of linkage is Z<-2.0.

Multipoint linkage analysis allows multiple loci to be analysed simultaneously. This can compensate for uninformative markers at particular meioses and for missing marker data by extracting data simultaneously from neighbouring markers and allowing construction of marker haplotypes.

Standard LOD score analysis requires a precise genetic model to be specified, including the mode of inheritance, gene frequency and penetrance of each genotype. This data is often an approximation or estimate. Model-free, or non-parametric analysis tests allele sharing amongst affected individuals in pedigrees, without any assumptions about model of inheritance. This bypasses the inherent problems of parametric analysis such as misspecification of parameters. Model-free analysis must distinguish between allele sharing identical by state (IBS) from alleles that are identical by descent (IBD). Identity by descent signifies shared alleles that have been inherited from a common ancestral chromosome, whereas identity by state refers to alleles that are shared coincidentally but do not have a common

ancestral origin. Exclusion of IBS can be increased by construction of multilocus haplotypes.

### 4.5.6 Positional candidate cloning

The availability of detailed genetic and physical maps, including the human genome project, has allowed for much easier identification of a potential causative gene. The positional candidate approach involves (i) linkage analysis to localize a disease gene to a chromosomal subregion (ii) searching databases for a candidate gene within that subregion, based on known gene function and an understanding of the disease pathophysiology and (iii) testing the candidate gene for disease-causing mutations.

### 4.5.7 Mutational analysis

Several methodologies are available for the identification of potential disease causing mutations in candidate genes. These include single strand conformational polymorphism (SSCP), heteroduplex analysis and oligonucleotide arrays. Direct sequencing is ultimately required to confirm a sequence change. Functionally important regions of the gene are initially screened for mutations, and any identified sequence variant is then evaluated to determine its significance. The association of a mutation with the disease trait can be investigated both in family-based linkage studies and population-based association studies (section 4.6.2). The functional consequences of a sequence variation depend in part on the location relative to the gene, pathogenic mutations often occurring within the coding sequence of the gene but also occurring in intragenic noncoding sequences or regulatory sequences outside exons. A functional variant is also likely to alter a highly conserved amino acid or affect a splice site or stop codon. Ultimately, the effect of a mutation on gene expression and protein function may be investigated using a variety of in-vivo and in-vitro techniques.

## 4.6 Identifying genes for diseases with 'complex' inheritance

### 4.6.1 Genetic architecture of 'complex' disease

Genetic architecture has a major influence on the optimal strategy for detecting susceptibility loci responsible for 'complex' disease traits. Alleles at loci determining Mendelian traits are rare (low frequency) and of major effect (highly penetrant). The same trait may be caused by rare, high penetrance alleles at distinct loci (locus heterogeneity). At a given locus many different disease alleles may occur (allelic heterogeneity).

Alleles at susceptibility loci are of small effect and likely to be of relatively high frequency (>1%). The number of loci, the magnitude of their individual effect on risk, their mode of interaction and the number and population frequency of susceptibility alleles underlying any particular trait is unknown. Oligogenic traits with a few loci of significant effect represent one end of the spectrum, with truly polygenic traits caused by numerous loci of small effect at the other end. The 'common-disease-common variant' hypothesis assumes that there is allelic homogeneity at each locus with the susceptibility allele present at high frequency (>5-10%) (Wright and Hastie 2001).

The methods of linkage analysis and positional cloning which are so powerful when applied to Mendelian disorders are much less useful for those traits which display 'complex' inheritance. Although there is increased familial clustering, segregation in families cannot be explained by the effect of a locus with dominant or recessive disease alleles. Interaction of several loci, each exerting a small effect, together with environmental factors is assumed. The power of linkage analysis is low if individual loci exert a small effect on the phenotype or if there is extensive locus heterogeneity. Thus, adequate power to detect linkage can only be attained with numbers of pedigrees above the number it may be practicable to ascertain. Numerous linkage analyses using a 'genome-wide' approach in complex traits have failed to identify linked loci which have been highly significant (statistically), replicable

or have led on to positional cloning of a susceptibility gene. In 'complex' traits, linkage cannot provide high-resolution localisation and positive results usually cover very large chromosomal regions. Therefore an alternative approach such as association analysis is often required.

### 4.6.2 Association analysis

Association studies detect non-random associations between a trait and either an allele or group of alleles in linkage disequilibrium (LD). LD is the non-random occurrence of specific alleles at adjacent loci (a haplotype). Association studies test for marker–disease correlations by investigating LD in a population. LD is measured between the disease and marker alleles, or preferably a haplotype of several markers, in unrelated individuals.

LD can be measured in several ways. A commonly used parameter of LD is D'. D' is related to the basic measure of LD, D, which is a measure of the deviation of haplotype frequencies from the equilibrium state, ie:

$$D=p(AB)-p(A)p(B)$$

where A and B are alleles at two linked SNPs and p(X) represents the frequency of the X allele or haplotype. D' is the ratio of D compared to its maximum value $D_{max}$ (or compared to its minimum value, $D_{min}$, when $D<0$). Complete LD is denoted by D'=1, and repeated recombination results in D' approaching zero.

SNPs are particularly suited as markers for association analysis. SNPs constitute the majority of the variation between human genomes, which on average differ by 0.1%. Most of these SNPs are 'neutral' but a subset with functional consequences is likely to include the allelic variation that accounts for common disease traits.

The control sample of chromosomes can be taken from an unaffected population. An alternative strategy is intra-familial association analysis, such as the transmission disequilibrium test (TDT) which uses internal controls. The TDT compares the transmission of alleles from heterozygous parents to

affected offspring (Spielman, McGinnis et al. 1993). The affection status of the parents is not relevant, and a single parent can be used. Each marker allele is tested by selecting all parents that are heterozygous for that allele and counting how many times, n, the allele is transmitted to the affected offspring. This is compared with the number of times the alternative allele is transmitted.

The TDT test statistic, $(n_{transmitted}n_{untransmitted})^2/(n_{transmitted}+n_{untransmitted})$, has a $\chi^2$ distribution with 1 degree of freedom. The pedigree disequilibrium test, PDT, is an extended test for intra-familial association that analyses LD in extended pedigrees using all potentially informative data (Martin, Monks et al. 2000; Martin, Bass et al. 2001).

Allelic association has greater power to detect susceptibility alleles of smaller effect, but is critically dependent on the 'common-disease-common variant' hypothesis being true for a particular trait (Reich and Lander 2001). If a wide diversity of low-frequency alleles causes susceptibility, association would be difficult or impossible to detect.

Association studies can be used in combination with linkage analysis to narrow a candidate region identified by linkage, or can be applied directly to test candidate genes. Chromosomal or genome-wide association studies have become a realistic possibility with the development of whole-genome association maps.

### 4.6.3 Association analysis used in combination with linkage analysis

Linkage operates over long distances, typically of the order of greater than 1Mb. Association analysis operates over much shorter distances (less than 100kb) and therefore can be used in certain situations to narrow a candidate region identified by linkage. However it is important to note that this is not always the case, and LD may not be significant in a region identified by linkage analysis. For example, if there is a high degree of allelic heterogeneity then no individual susceptibility allele will be detectable by LD

with neighbouring markers. Similarly, for very old susceptibility alleles, LD in the region may no longer be present due to successive recombination.

### 4.6.4 Candidate gene association analysis

Association analysis can be used to test candidate genes identified on the basis of known gene function and an understanding of the disease pathophysiology. Intragenic SNPs are identified from SNP databases or by direct sequencing in a control population. The number of SNPs required to be tested can be significantly reduced by the identification of the LD block structure of the region encompassing the candidate gene. Extensive blocks of LD are present in the human genome, at least in the Northern European population, which create haplotypes between 25 and 100 kb long (Daly, Rioux et al. 2001). LD blocks are defined by an intermarker LD of |D'| greater than a predetermined value, commonly 0.7. An entire gene may lie within an LD block, or may span multiple blocks. Within an LD block, it is unnecessary to genotype every SNP. A subset of SNPs will represent the majority of the genetic variation within an LD block and can be designated as haplotype tagging SNPs (htSNPs). Typing the htSNPs alone will be sufficient to identify the common haplotypes, and will significantly reduce the amount of genotyping required.

### 4.6.5 Whole genome association analysis

A genome wide association study searches the whole genome for susceptibility alleles with no previous assumptions about the likely location. This has become a practical possibility with the development of rapid, high-throughput, cheap SNP genotyping technology and increasingly comprehensive SNP databases incorporating the LD structure of the human genome (e.g. dbSNP, HapMap). Fortunately it is not necessary to genotype all of the approximately 11 million SNPs with allele frequency greater than 1%. As discussed in section 4.6.4, the pattern of LD allows haplotype tagging SNPs to be selected that capture most of the common genetic variation in a region. The stronger the LD in a particular region, the less htSNPs needed to

be typed to capture the variation. On the basis of the HapMap data, it is estimated that a few hundred thousand htSNPs will be required to cover the whole genome (Hirschhorn and Daly 2005). HtSNPs are usually selected to detect relatively common haplotypes (frequency>5%). Therefore rarer susceptibility alleles may be missed by this approach. However rare SNPs with minor allele frequency <5% still tend to lie on a single commoner haplotype, as rare SNPs are likely to have arisen recently so that they have not had time to either disappear or become common, and the LD block has not been eroded by recombination (Hirschhorn and Daly 2005). Thus rarer alleles that are highly penetrant and still account for a significant proportion of the genetic trait in question should still be detectable by the htSNP approach (Kamatani, Sekine et al. 2004).

An alternative approach which has been suggested to reduce the number of SNPs typed is a sequence-based approach. This targets SNPs in DNA sequences that are likely to be of functional importance, including coding regions of genes (particularly missense SNPs), splice sites, promoter regions and sequences of DNA that are highly conserved between species (Botstein and Risch 2003). This approach may identify susceptibility alleles of lower frequency than the haplotype-based approach, but would miss functionally important changes in non-coding or non-conserved regions.

Several other factor need to be considered in the design of a genome wide association study. These include:

### *i. Sample size*

Large sample sizes are needed to detect variants of modest effect. Power to detect a significant association is further decreased by the problem of multiple hypothesis testing (due to the large number of SNPs tested in a genome-wide association study) and p-values must be corrected for this (see point ii). One possible strategy for reducing sample size is a multistage approach where an initial population is typed for all SNPs, and a second independent population is used to retest only those SNPs with putative associations. Another strategy to increase the power of a population to demonstrate a positive association for rare susceptibility alleles is to study a

founder population, one that has expanded from a bottleneck of relatively few individuals in the recent past (less than 100 generations). Rare alleles are likely to be found on a single large shared haplotype in the founder population, so that fewer markers need to be genotyped.

### *ii. Significance thresholds*

The multiple testing performed during a genome-wide association study requires a correction in *p*-value. The standard is a Bonferroni correction which assumes that all tests are independent:

$$p_{corrected} = 1 - (1-p_{uncorrected})^n$$

However this may be an over correction as it does not take into account LD between markers and marker redundancy. Other strategies for estimating statistical significance include permutation testing (where the data sets for the cases and controls are repeatedly mixed up, resorted randomly and retested to produce a null distribution of best test statistics for comparison with the best actual test result) and Bayesian approaches.

### *iii. False positives*

False positive associations can arise due to population stratification, where the study population consists of more than one genetically distinct group (admixture), and these subgroups are not equally represented in the cases and controls. Thus a difference in marker allele frequencies between cases and controls can simply represent the different genetic subgroups and have no relationship to the trait under study. Population stratification can be detected and corrected for by evaluating large numbers of unlinked markers, of which there is an abundance in a genome-wide association study (Pritchard, Stephens et al. 2000). Intra-familial association studies also avoid the problem of population stratification. Other sources of false-positive associations include technical artefacts and statistical fluctuations.

# PART 2 Materials and Methods

# *Chapter 5 Patients and Families*

## 5.1 Summary

This section describes the ascertainment and characterisation of the total family resource, in collaboration with a network of paediatricians and paediatric neurologists in the UK and abroad.

Criteria for affectedness status were a clinical diagnosis of CAE or JAE according to ILAE criteria. The referring paediatrician usually made the initial diagnosis but confirmatory documentation of the clinical phenotype and EEG data were obtained in all cases. The clinical phenotype of affected individuals was recorded in an electronic database. This data was subjected to security control.

## 5.2 Family Resource

A resource of patients and families with IGE was previously ascertained during a BIOMED 2 funded Concerted Action on Genetic Analysis of Epilepsy (1993-1996). This resource, which includes 78 pedigrees in which at least one individual has absence epilepsy, has been employed predominantly for a genome wide IGE screen.

Multi-centre ethical approval for the current study was obtained through the North Thames Multi-centre Research Ethics Committee (reference number MREC/99/2/61)

Collaboration with clinicians in both the UK and Europe has increased the patient resource to 299 families including 405 affected individuals (CAE, JAE or AE as discussed in section 5.2). 82 families contain more than one affected individual. Detailed clinical information has been documented and DNA has been obtained from all affected individuals and, where possible, parents and unaffected siblings. DNA was also obtained from a family in which absence epilepsy segregates with episodic ataxia type 2 (EA2) in an autosomal dominant fashion through three generations. The patient resource was ascertained in the following ways.

### 5.2.1 Recruitment in the United Kingdom

1. A mailshot was sent to 322 consultant paediatricians in the North and South Thames Regions, identified from the Royal College of Paediatrics and Child Health (RCPCH) handbook (Appendix 1a). The mailshot was also sent to 37 consultant paediatric neurologists throughout the UK, identified from the British Paediatric Neurology Association (BPNA) membership list. 124 paediatricians ultimately expressed interest in collaborating and provided patients for the study.

2. Collaborations were established with neurophysiology departments performing paediatric EEGs throughout the UK (Appendix 2). EEG records were searched for reports consistent with a diagnosis of CAE or JAE. A letter was then sent to the referring consultant in order to confirm the diagnosis and to request permission to pass on information about the study to the family (Appendix 1b).

3. Details of the study including a call for patients were published in the BPNA newsletter in January 2002 (PART 4Chapter 17Appendix 3). A project summary was also circulated at the BPNA Annual Meeting (PART 4Chapter 17Appendix 4).

4. A collaboration was established with Dr AJ Makoff and Dr L Nashef in the epilepsy genetics research group at the Institute of Psychiatry in London. Clinical information and DNA samples from 35 absence patients were contributed to this study.

### 5.2.2 Recruitment outside the United Kingdom

1. Collaborations were established with paediatric neurologists and epilepsy genetics researchers throughout Europe (PART 4Chapter 17Appendix 5). Several of these originated from the Concerted Action on Genetic Analysis of Epilepsy (1993-1996) and the subsequent European Consortium on the Genetics of Idiopathic Generalised Epilepsy.

2. Members of the Canadian Epilepsy Database and Registry (CEDaR) agreed to collaborate following a meeting at the American Epilepsy Society Annual Meeting in Seattle in 2002 (Appendix 5)

## 5.3 Criteria for affectedness status

Criteria for affectedness status were a clinical diagnosis of CAE or JAE according to ILAE criteria (ILAE 1989). An additional category of absence epilepsy, AE, was use to include individuals with a childhood onset of absence seizures but who did not meet the strict criteria for CAE.

### 5.3.1 Childhood absence epilepsy

**Inclusion criteria:**

Brief (4 to 20 seconds, exceptionally longer) and usually frequent (tens per day) absence seizures (of any type except myoclonic absences) with abrupt and severe impairment of consciousness as predominant seizure type. Automatisms may occur.

Onset between 2 and 12 years

Generalised tonic-clonic seizures or unclassified seizures may also occur

Seizures may persist into adulthood

Normal neurological examination and development

Ictal EEG showing bilateral, synchronous, symmetrical high amplitude discharges of 2.5-4 Hz spike and double spike and slow wave complexes on a normal background. 3-4 Hz occipital rhythms may be present.

**Exclusion criteria:**

Prominent myoclonus (brief jerks at onset of absence seizure do not exclude patient)

Eyelid myoclonus (fluttering of eyelids during absence seizure does not exclude patient)

Significant developmental delay

Persistent or focal neurological deficit

Polyspike-wave complexes in the ictal EEG, or clear evidence of photosensitivity

Clear abnormalities on neuroimaging (although most patients will not have had neuroimaging performed)

### 5.3.2 Juvenile absence epilepsy

For the purposes of this study, the following criteria were used to distinguish JAE from CAE:

Onset of absences after 12 years of age

Less frequent absences than in CAE

A frequent association with generalised tonic clonic seizures

Myoclonic seizures occurring not infrequently

Ictal EEG may show polyspike-wave complexes at a frequency of 3.5Hz to 4.5 Hz. Photosensitivity may be present.

### 5.3.3 Absence epilepsy

Individuals were identified with many of the features of CAE and an onset of absence seizures below the age of 12 years, but who could not be classified as CAE due to certain atypical features. For the purpose of this study, these individuals were classified as absence epilepsy (AE). The atypical features included:

- Onset of absences below the age of 2 years
- Prominent myoclonus (other than eyelid myoclonus) as a feature of absence seizures
- Prolonged absence seizures
- Atypical EEG features including polyspike-wave complexes in the ictal EEG, or clear evidence of photosensitivity

## 5.4 Recruitment procedure

Potential participants were identified by the methods above. Whenever possible, initial contact with a family regarding the study was made by their own consultant paediatrician. A letter of invitation to participate was then sent together with information sheets in three versions – for adults, teenagers and younger children (Appendix 6, Appendix 7). All positive responses were followed up by a phone call and, when indicated, a home visit to confirm the clinical details and obtain informed consent for all participating family members. Consent forms were signed by participating subjects or, when appropriate, parents or guardians (Appendix 8).

## 5.5 Clinical data

Clinical data was obtained in the following ways:

1. Direct questioning of affected individuals and family members
2. Direct questioning of referring clinician
3. Clinical documentation and correspondence from referring clinician
4. Description of ictal events from EEG reports including video telemetry where available

Data was initially recorded in Family Clinical Data Sets (Appendix 9). It was then stored electronically in a secured Microsoft® Access Database. The Access database was organised into three tables - Family Investigator, Family Overview and Clinical Data (Table 5-1). Data could be entered using data entry forms (Figure 5-1, Figure 5-2). Blank versions of the Access database were emailed to collaborators to enable data on their local patients to be entered directly on a local computer. This secured data could then be emailed back and added to the master database. Instructions on data entry were provided for collaborators (Appendix 10). The database could be queried to identify subgroups of subjects with particular characteristics.

## 5.6 DNA collection

DNA was obtained from all participating family members, including unaffected parents and siblings. Samples were collected by one of two methods.

### 5.6.1 Venous blood sampling

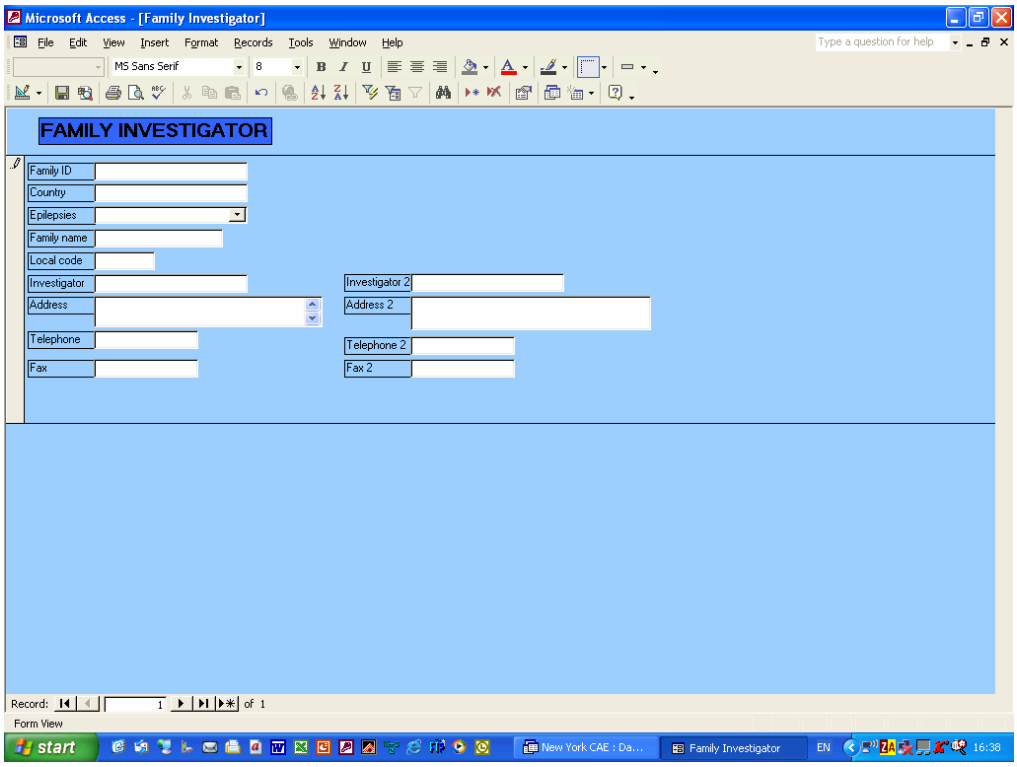
Venous EDTA blood samples were collected (5ml from children, up to 20ml from adults) whenever possible for the high DNA yield. A topical anaesthetic cream (Emla or Ametop) was used for children. For some local families, home visits were organised to perform venepuncture. In other cases, local GPs or hospital phlebotomy departments agreed to perform venepuncture and the blood samples were sent to University College London for DNA extraction. Whenever possible, blood was taken at the same time as routine specimens during clinic appointments. Instructions for blood sample collection and transport were provided (Appendix 11a). When blood samples were taken outside the UK, the blood was frozen and transported on dry ice or DNA was extracted locally.

### 5.6.2 Buccal cell sampling

Buccal cells were obtained using cheek swabs when venepuncture was difficult to perform or family members were unwilling to give blood. Five sterile swabs were provided for each individual and cheek swab kits (including instructions, preservative solution and stamped addressed padded envelopes for return of samples) were posted directly to the families (Appendix 11b).

**Table 5-1** *Clinical database organisation*

| Table | Notes |
|---|---|
| **i) Family Investigator** | |
| Family ID | 3 digit code, one record per family |
| Country | |
| Epilepsies | This represents all the epilepsies/seizures in the family, each represented by a 2-letter code:<br>CA: CAE, JA: JAE, JM: JME, IG: IGE, EM: EMA, FC: Febrile convulsions,<br>PS: Photosensitive seizures, OE: Other epilepsy |
| Individual C/ACode | Concerted action code if relevant |
| Family name | |
| Local code | |
| Investigator name | |
| Address | |
| Telephone | |
| Fax | |
| | |
| **ii) Family Overview** | |
| Family ID | 3 digit code |
| Individual ID | 3 digit code. The usual individual IDs are: father 101, mother 102, ch |
| Full ID | 6 digit code - a combination of the family ID and the individual ID |
| Country | |
| Epilepsies | Epilepsies – this represents all epilepsies/seizures for that individual. |
| Subject Name | |
| Date of Birth | |
| Sex | |
| Relationship | Relationship: P – proband, PP – parent of proband, SP – sib of proba |
| Date Last Observation | |
| Date of Entry | |
| EEG Available | |
| DNA | |
| Ethincity | |
| | |
| **iii) Clinical Data** | |
| Full ID | |
| Country | |
| Epilepsies | |
| Seizures | GSA: generalised seizure, absence, GSM: generalised seizure, myoclonic<br>GTCS: generalised tonic-clonic seizure, FC: febrile convulsion<br>GSAmy: generalised seizure, absence + myoclonus, US: unclassified seizure |
| Past Medical History | |
| Examination | |
| Neurodevelopment | |
| Imaging | |
| Clinical Comments | |
| Date of EEG | Up to three EEGs can be recorded |
| Normal? | |
| Patient age at EEG | |
| Background | |
| Patient state | |
| EEG Ictal | |
| Ictal Phenomena | |
| Duration of EEG | |
| Video-Telemetry | |
| Medication at EEG | Up to three medications can be recorded |
| Abnormality spontaneous | |
| Hyperventilation | |
| Photo Sensitive | |
| Generalized Abnormality | |
| Frequency in HZ | |
| Focal abnormalities | |



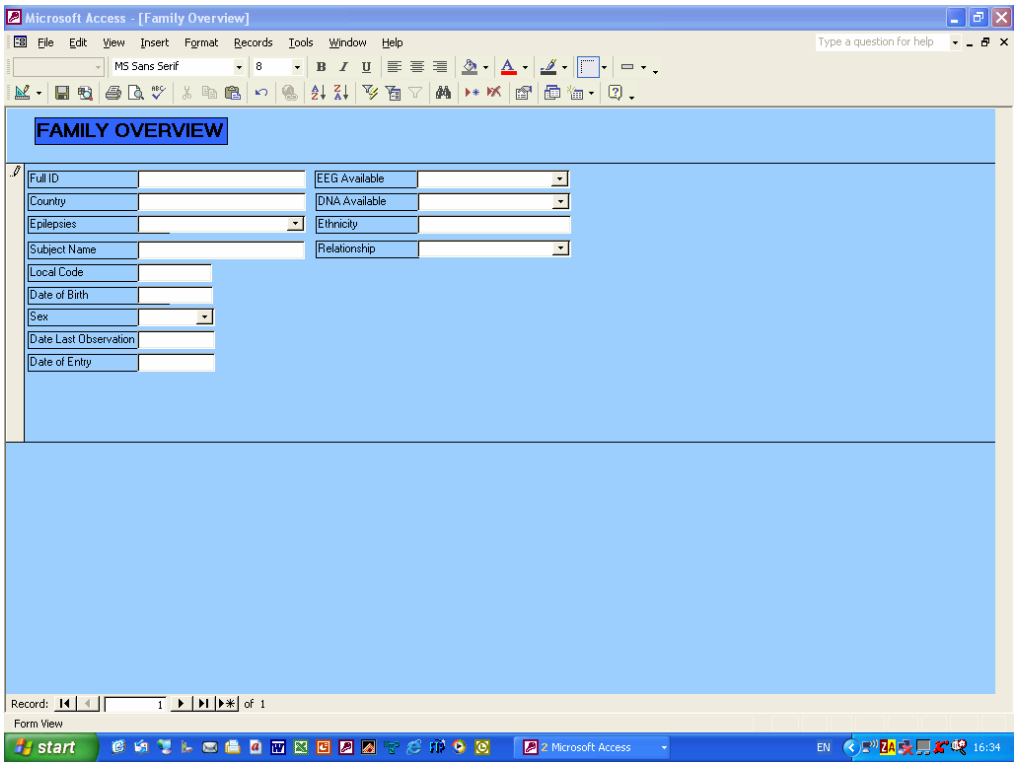
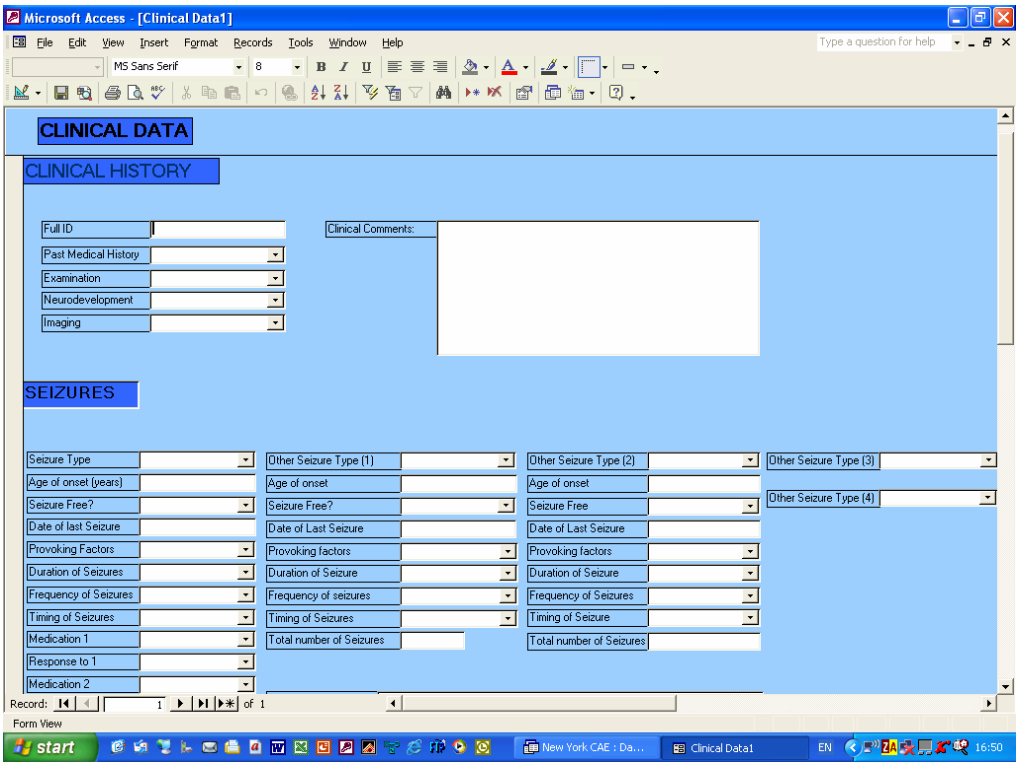




Figure 5-1 *Clinical database - screen shots showing family data entry forms*



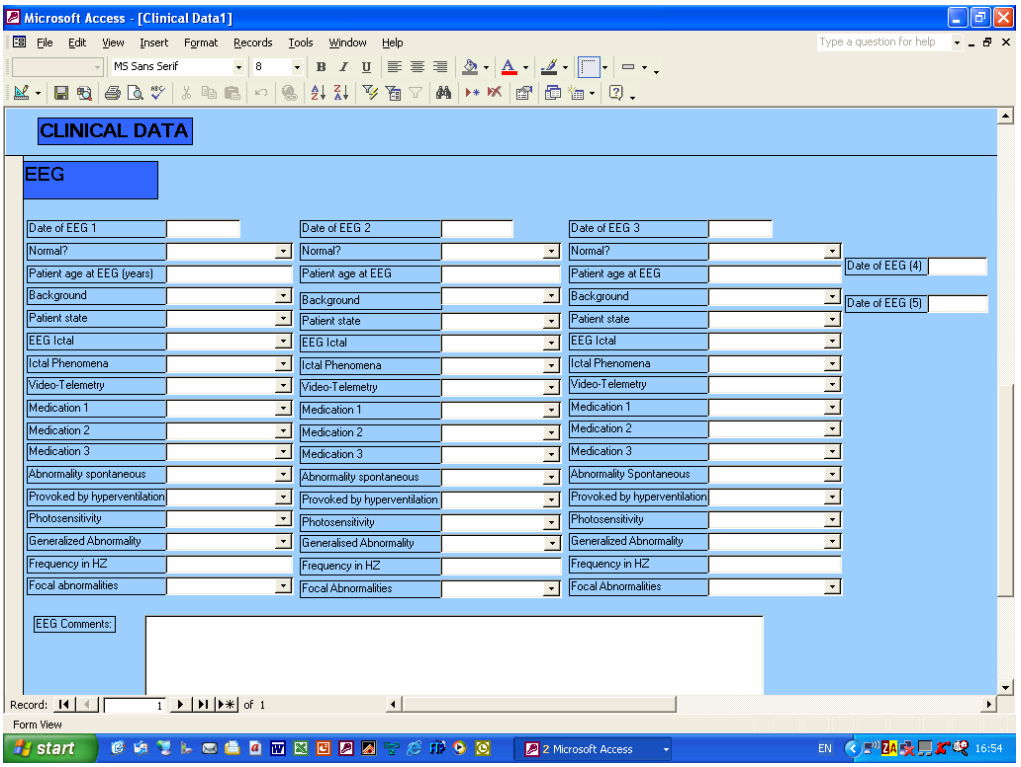




Figure 5-2 *Clinical database - screen shots showing clinical data entry forms*

## 5.7 Nuclear pedigrees

### 5.7.1 Pedigrees used for initial linkage analysis

An initial cohort of 33 nuclear families was used to test 30 candidate gene loci (seventeen VDCC subunit genes, ten $GABA_AR$ subunit genes, two $GABA_B$ receptor genes and the *ECA1* locus on 8q24) by linkage analysis (Figure 5-3,Table 7-1). The families were ascertained from European populations including the UK, France, Germany, Denmark, Finland and Greece. All families included two or more individuals with a provisional diagnosis of CAE or JAE. The 33 families included 172 individuals of which 83 were classified as affected. DNA was not available on 6 unaffected and 2 affected individuals. Of the affected individuals, 11 did not have EEG data available, but otherwise had typical CAE or JAE. For this preliminary linkage analysis, polyspike-wave complexes in the ictal EEG (1 individual), or clear evidence of photosensitivity (7 individuals) were not used as exclusion criteria for CAE. Five individuals were classified as affection status unknown for the linkage analysis. Of these, 2 had generalised spike-waves on EEG but no clinical seizures, 2 had generalised tonic-clonic seizures only, and one had generalised tonic-clonic seizures on awakening

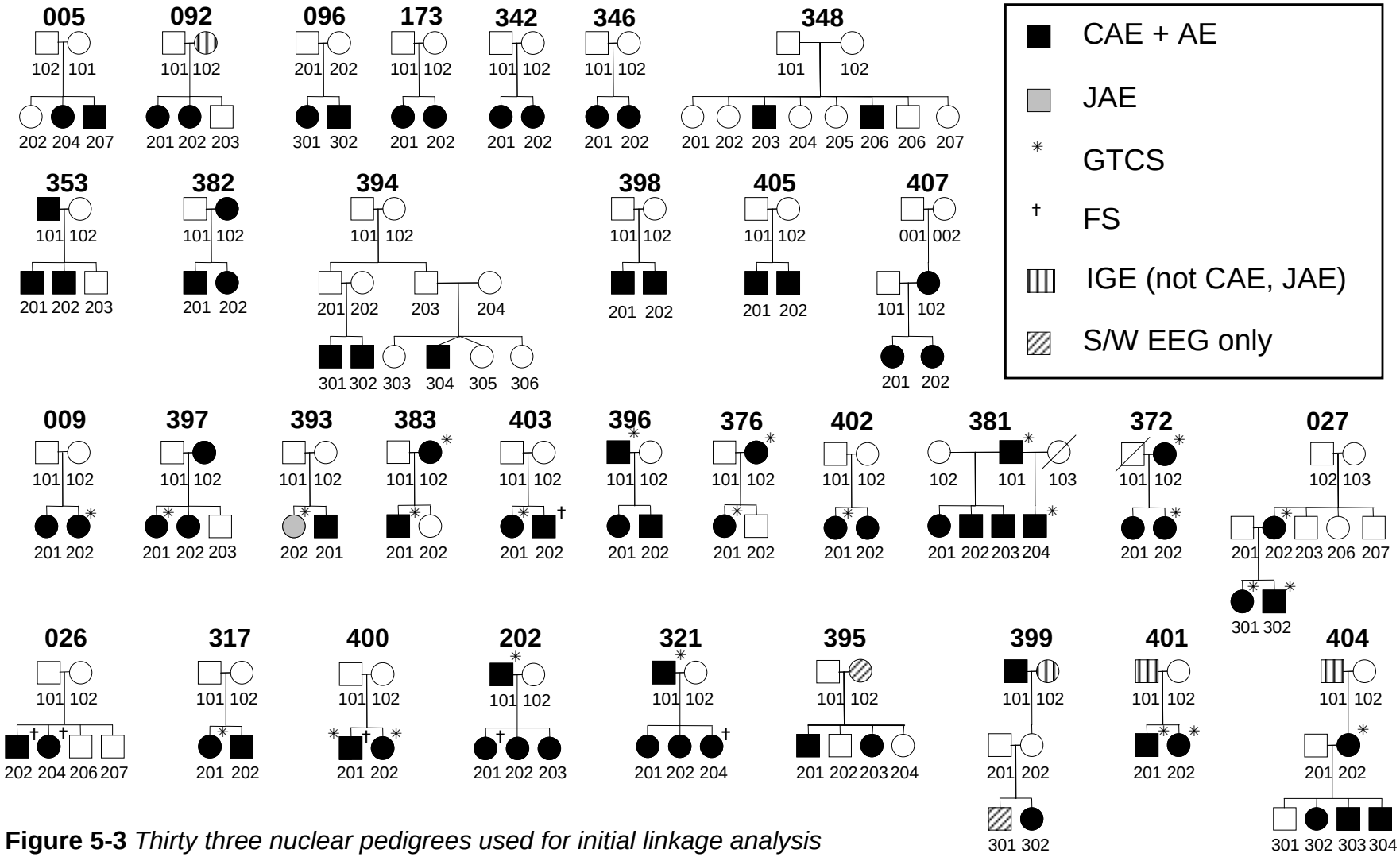



**Figure 5-3** *Thirty three nuclear pedigrees used for initial linkage analysis*

### 5.7.2 Pedigrees used for definitive linkage analysis

Following the preliminary linkage analysis, candidate gene loci showing positive linkage were tested in an expanded collection of families. Four of the original 33 families (families 376, 383, 398, 399) as well as 10 individuals from two other families (families 381, 394) were not included as new data became available which cast doubt on affectedness status as defined below. A total of 82 families were included and analysed in two ways with different criteria for affectedness status:

#### *i. Affectedness status 'CAE' using strict criteria*

A cohort of 53 nuclear pedigrees included two or more first degree relatives with a diagnosis of CAE according to strict criteria (Figure 5-4,Table 5-2). Twenty-two of the original 33 families were included. The 53 families included 275 individuals of which 120 with CAE were classified as 'affected'. Eleven individuals were classified as affection status 'unknown' (two AE, two with spike-wave EEG but no seizures, five with a different IGE and two with no clinical data available). DNA was not available on 19 individuals (15 unaffected, 2 affected and 2 affection status unknown). Additional clinical information is provided in Appendix 12.

#### *ii. Affectedness status 'absence' including CAE, JAE, AE*

A cohort of 82 families, comprising 418 individuals, included the 53 CAE families. The 29 additional families (Figure 5-5, Table 5-3) comprised two or more first degree relatives with a diagnosis of absence epilepsy consistent with CAE (section 5.3.1), JAE (section 5.3.2) or AE (section 5.3.3). The 82 families comprised a total of 418 individuals of whom 185 were classified as 'affected' (129 CAE, 28 JAE, and 28 AE). Twenty individuals were classified as affection status 'unknown'. DNA was not available on 29 individuals (21 unaffected, 3 affected and 5 affection status unknown). Additional clinical information is provided in Appendix 12.

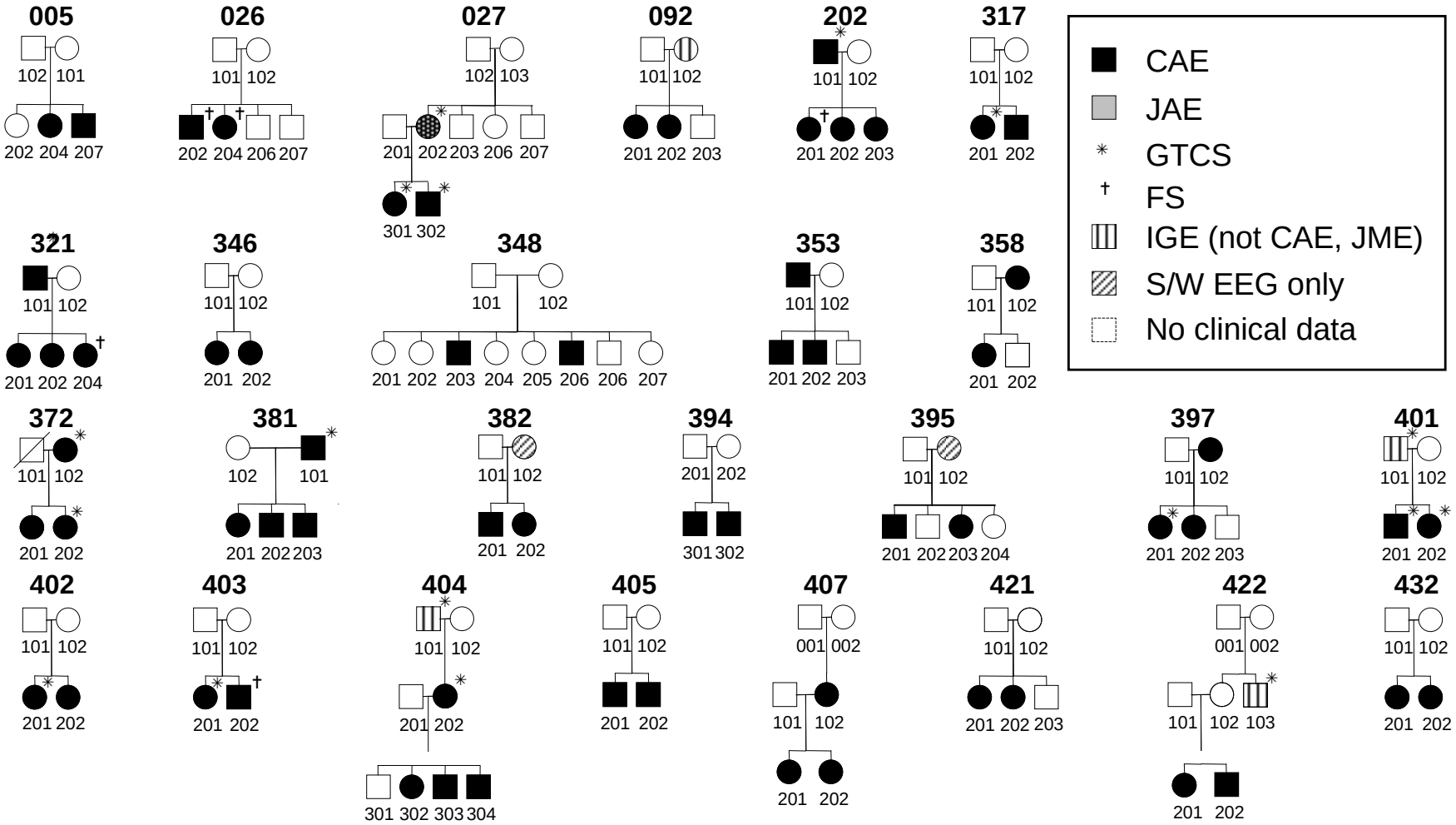



**Figure 5-4** *53 nuclear pedigrees used for definitive linkage analysis with affectedness status 'CAE' (continued on next page)*

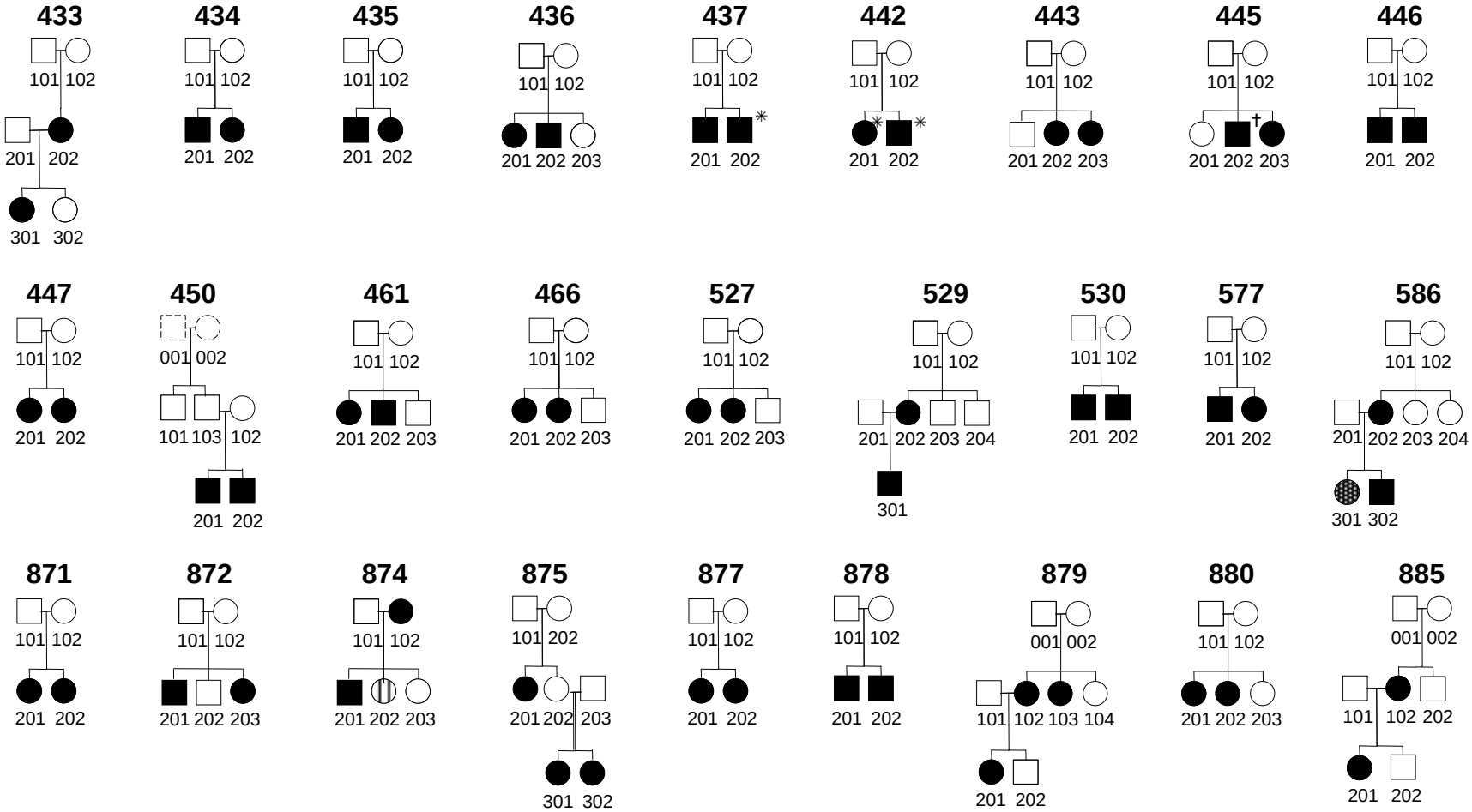



**Figure 5.4 continued** *53 nuclear pedigrees used for definitive linkage analysis with affectedness status 'CAE' (continued from previous page)*

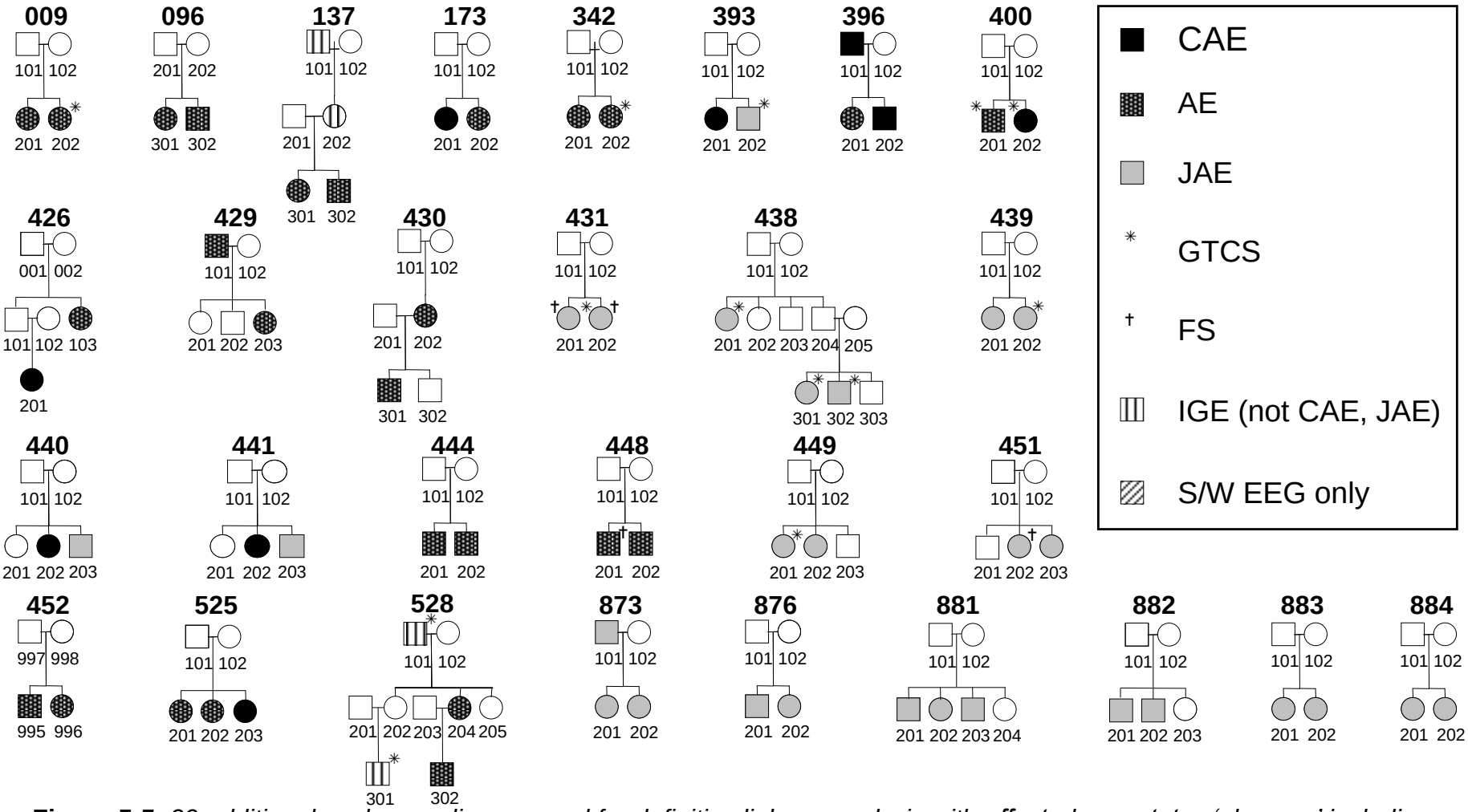



**Figure 5-5** *29 additional nuclear pedigrees used for definitive linkage analysis with affectedness status 'absence' including CAE, JAE, AE*

**Table 5-2** *Families used for CAE linkage. Only individuals with a diagnosis of CAE (classified 'affected') or AE, IGE or GSWD EEG (classified 'unknown') are shown.*

| Family | Individual | Sex | Epilepsy | Age at onset | Seizures | EEG |
|---|---|---|---|---|---|---|
| 5 | 204 | F | CAE | 4 | GSA | GSWD |
| | 207 | M | CAE | 5 | GSA | GSWD |
| 26 | 202 | M | CAE | 6 | GSA, FS | GSWD |
| | 204 | F | CAE | 6 | GSA, FS | GSWD |
| 27 | 202 | F | AE | 4 | GSA, GTCS | GSWD, PS |
| | 301 | F | CAE | 7 | GSA, GTCS | GSWD |
| | 302 | M | CAE | 7 | GSA, GTCS | GSWD |
| 92 | 102 | F | IGE | | GTCS | |
| | 201 | F | CAE | 2-12 | GSA | GSWD |
| | 202 | F | CAE | 2-12 | GSA | GSWD |
| 202 | 101 | M | CAE | 5 | GSA, GTCS | GSWD |
| | 201 | F | CAE | 3 | GSA, FS | GSWD |
| | 202 | F | CAE | 3 | GSA | GSWD |
| | 203 | F | CAE | 6 | GSA | GSWD |
| 317 | 201 | F | CAE | 5 | GSA, GTCS | GSWD |
| | 202 | M | CAE | 5 | GSA | GSWD |
| 321 | 101 | M | CAE | 9 | GSA, GTCS | GSWD |
| | 201 | F | CAE | 6 | GSA | GSWD |
| | 202 | F | CAE | 3 | GSA | GSWD |
| | 204 | F | CAE | 3 | GSA, FS | GSWD |
| 346 | 201 | F | CAE | 2 | GSA | GSWD |
| | 202 | F | CAE | 2-12 | GSA | GSWD |
| 348 | 203 | M | CAE | 3 | GSA | GSWD |
| | 206 | M | CAE | 3 | GSA | GSWD |
| 353 | 101 | M | CAE | 10 | GSA | GSWD |
| | 201 | M | CAE | 6 | GSA | GSWD |
| | 202 | M | CAE | 9 | GSA | GSWD |
| 358 | 102 | F | CAE | 4 | GSA | GSWD |
| | 201 | F | CAE | 5 | GSA | GSWD |
| 372 | 102 | F | CAE | 6 | GSA, GTCS | GSWD |
| | 201 | F | CAE | 4 | GSA | GSWD |
| | 202 | F | CAE | 5 | GSA, GTCS | GSWD |
| 381 | 101 | M | CAE | 2-12 | GSA, GTCS | GSWD |
| | 201 | F | CAE | 5 | GSA | GSWD |
| | 202 | M | CAE | 5 | GSA | GSWD |
| | 203 | M | CAE | 6 | GSA | GSWD |
| 382 | 102 | F | Unknown | | None | GSWD |
| | 201 | M | CAE | 7 | GSA | GSWD |
| | 202 | F | CAE | 6 | GSA | GSWD |
| 394 | 301 | M | CAE | 7 | GSA | GSWD |
| | 302 | M | CAE | 8 | GSA | GSWD |
| 395 | 102 | F | Unknown | | None | GSWD |
| | 201 | M | CAE | 4 | GSA | GSWD |
| | 203 | M | CAE | 3 | GSA | GSWD |
| 397 | 102 | F | CAE | 5 | GSA | GSWD |
| | 201 | F | CAE | 3 | GSA, GTCS | GSWD |
| | 202 | F | CAE | 3 | GSA | GSWD |

| Family | Individual | Sex | Epilepsy | Age at onset | Seizures | EEG |
|---|---|---|---|---|---|---|
| 401 | 101 | F | IGE | 20 | GTCS | |
| | 201 | M | CAE | 4 | GSA, GTCS | GSWD |
| | 202 | F | CAE | 3 | GSA, GTCS | GSWD |
| 402 | 201 | F | CAE | 12 | GSA, GTCS | GSWD |
| | 202 | F | CAE | 11 | GSA | GSWD |
| 403 | 201 | F | CAE | 3 | GSA, GTCS | GSWD |
| | 202 | M | CAE | 3 | GSA, FS | GSWD |
| 404 | 101 | M | IGE | 20 | GTCS | |
| | 202 | F | CAE | 6 | GSA, GTCS | GSWD |
| | 302 | F | CAE | 7 | GSA | GSWD |
| | 303 | M | CAE | 7 | GSA | GSWD |
| | 304 | M | CAE | 5 | GSA | GSWD |
| 405 | 201 | M | CAE | 6 | GSA | GSWD |
| | 202 | M | CAE | 3 | GSA | GSWD |
| 407 | 102 | F | CAE | 5 | GSA | GSWD |
| | 201 | F | CAE | 3 | GSA | GSWD |
| | 202 | F | CAE | 5 | GSA | GSWD |
| 421 | 201 | F | CAE | 2-12 | GSA | GSWD |
| | 202 | F | CAE | 2-12 | GSA | GSWD |
| 422 | 103 | M | IGE | 9 | GTCS | |
| | 201 | F | CAE | 6 | GSA | GSWD |
| | 202 | M | CAE | 12 | GSA | GSWD |
| 432 | 201 | F | CAE | 2-12 | GSA | GSWD |
| | 202 | F | CAE | 2-12 | GSA | GSWD |
| 433 | 202 | F | CAE | 2-12 | GSA | GSWD |
| | 301 | F | CAE | 2-12 | GSA | GSWD |
| 434 | 201 | M | CAE | 2-12 | GSA | GSWD |
| | 202 | F | CAE | 2-12 | GSA | GSWD |
| 435 | 201 | M | CAE | 2-12 | GSA | GSWD |
| | 202 | F | CAE | 2-12 | GSA | GSWD |
| 436 | 201 | F | CAE | 2-12 | GSA | GSWD |
| | 202 | M | CAE | 2-12 | GSA | GSWD |
| 437 | 201 | M | CAE | 2-12 | GSA | GSWD |
| | 202 | M | CAE | 2-12 | GSA, GTCS | GSWD |
| 442 | 201 | F | CAE | 2-12 | GSA, GTCS | GSWD |
| | 202 | M | CAE | 2-12 | GSA, GTCS | GSWD |
| 443 | 202 | F | CAE | 2-12 | GSA | GSWD |
| | 203 | F | CAE | 2-12 | GSA | GSWD |
| 445 | 202 | M | CAE | 2-12 | GSA, FS | GSWD |
| | 203 | F | CAE | 2-12 | GSA | GSWD |
| 446 | 201 | M | CAE | 2-12 | GSA | GSWD |
| | 202 | M | CAE | 2-12 | GSA | GSWD |
| 447 | 201 | F | CAE | 2-12 | GSA | GSWD |
| | 202 | F | CAE | 2-12 | GSA | GSWD |
| 450 | 201 | M | CAE | 2-12 | GSA | GSWD |
| | 202 | M | CAE | 2-12 | GSA | GSWD |
| 461 | 201 | F | CAE | 2-12 | GSA | GSWD |
| | 202 | M | CAE | 2-12 | GSA | GSWD |
| 466 | 201 | F | CAE | 2-12 | GSA | GSWD |
| | 202 | F | CAE | 2-12 | GSA | GSWD |
| 527 | 201 | F | CAE | 7 | GSA | GSWD |
| | 202 | F | CAE | 3 | GSA | GSWD |

| Family | Individual | Sex | Epilepsy | Age at onset | Seizures | EEG |
|---|---|---|---|---|---|---|
| 529 | 202 | F | CAE | 7 | GSA | GSWD |
| | 301 | M | CAE | 5 | GSA | GSWD |
| 530 | 201 | M | CAE | 5 | GSA | GSWD |
| | 202 | M | CAE | 5 | GSA | GSWD |
| 577 | 201 | M | CAE | 10 | GSA | GSWD |
| | 202 | F | CAE | 7 | GSA | GSWD |
| 586 | 202 | F | CAE | 5 | GSA | GSWD |
| | 301 | F | AE | 5 | GSAmy | GPSWD |
| | 302 | M | CAE | 7 | GSA | GSWD |
| 871 | 201 | F | CAE | 2-12 | GSA | GSWD |
| | 202 | F | CAE | 2-12 | GSA | GSWD |
| 872 | 201 | M | CAE | 2-12 | GSA | GSWD |
| | 203 | F | CAE | 2-12 | GSA | GSWD |
| 874 | 102 | F | CAE | 2-12 | GSA | GSWD |
| | 201 | M | CAE | 2-12 | GSA | GSWD |
| | 202 | F | IGE | 2-12 | GTCS | |
| 875 | 201 | F | CAE | 2-12 | GSA | GSWD |
| | 301 | F | CAE | 2-12 | GSA | GSWD |
| | 302 | F | CAE | 2-12 | GSA | GSWD |
| 877 | 201 | F | CAE | 2-12 | GSA | GSWD |
| | 202 | F | CAE | 2-12 | GSA | GSWD |
| 878 | 201 | M | CAE | 2-12 | GSA | GSWD |
| | 202 | M | CAE | 2-12 | GSA | GSWD |
| 879 | 102 | F | CAE | 2-12 | GSA | GSWD |
| | 103 | F | CAE | 2-12 | GSA | GSWD |
| | 201 | F | CAE | 2-12 | GSA | GSWD |
| 880 | 201 | F | CAE | 2-12 | GSA | GSWD |
| | 202 | F | CAE | 2-12 | GSA | GSWD |
| 885 | 102 | F | CAE | 2-12 | GSA | GSWD |
| | 201 | F | CAE | 2-12 | GSA | GSWD |

CAE: childhood absence epilepsy; JAE: juvenile absence epilepsy; AE: absence epilepsy; GSA: generalised absence seizure; GSAmy: generalised absence seizure with myoclonus; FS: febrile seizure; GSWD: generalised spike-wave discharges; GPSWD: generalised polyspike-wave discharges

**Table 5-3** *Additional families used for 'absence' linkage. Only individuals with a diagnosis of CAE or AE (classified 'affected') or IGE or GSWD EEG (classified 'unknown') are shown.*

| Family | Individual | Sex | Epilepsy | Age at onset | Seizures | EEG |
|---|---|---|---|---|---|---|
| 9 | 201 | F | AE | 4 | GSAmy | GPSWD |
| | 202 | F | AE | 2 | GSA, GTCS | GSWD, FD |
| 96 | 301 | F | AE | 7 | GSA | GPSWD |
| | 302 | M | AE | 6 | GSA | GPSWD |
| 137 | 101 | M | IGE | 18 | GTCS, MS | GSWD |
| | 202 | F | IGE | 12 | GTCS, MS | GSWD |
| | 301 | F | AE | 1 | GSAmy | GPSWD |
| | 302 | M | AE | 1 | GSAmy | GPSWD |
| 173 | 201 | F | CAE | 6 | GSA | GSWD |
| | 202 | F | AE | 5 | GSA | GPSWD |
| 342 | 201 | F | AE | 11 | GSAmy | GPSWD |
| | 202 | F | AE | <13 | GSAmy, GTCS | GPSWD |
| 393 | 201 | F | CAE | 8 | GSA | GSWD |
| | 202 | M | JAE | 15 | GSA, GTCS | GPSWD |
| 396 | 101 | M | CAE | 2-12 | GSA | GSWD |
| | 201 | F | AE | 3 | GSA | GPSWD |
| | 202 | M | CAE | 6 | GSA | GSWD |
| 400 | 201 | M | AE | 5 | GSA, GTCS | GPSWD |
| | 202 | F | CAE | 5 | GSA, GTCS | GSWD |
| 426 | 103 | F | AE | <13 | GSA | GPSWD |
| | 201 | F | CAE | 6 | GSA | GSWD |
| 429 | 101 | M | AE | <13 | GSA | GPSWD |
| | 203 | F | AE | <13 | GSA | GPSWD |
| 430 | 202 | F | AE | <13 | GSA | GPSWD |
| | 301 | M | AE | <13 | GSA | GPSWD |
| 431 | 201 | F | JAE | 13+ | GSA, FS | GPSWD |
| | 202 | F | JAE | 13+ | GSA, FS, GTCS | GPSWD |
| 438 | 201 | F | JAE | 13+ | GSA, GTCS | GPSWD |
| | 301 | F | JAE | 13+ | GSA, GTCS | GPSWD |
| | 302 | M | JAE | 13+ | GSA, GTCS | GPSWD |
| 439 | 201 | F | JAE | 13+ | GSA | GPSWD |
| | 202 | F | JAE | 13+ | GSA, GTCS | GPSWD |
| 440 | 202 | F | CAE | | GSA | GSWD |
| | 203 | M | JAE | 13+ | GSA | GPSWD |
| 441 | 202 | F | CAE | 2-12 | GSA | GSWD |
| | 203 | M | JAE | 13+ | GSA | GPSWD |
| 444 | 201 | M | AE | <13 | GSA | GPSWD |
| | 202 | M | AE | <13 | GSA | GPSWD |
| 448 | 201 | M | AE | <13 | GSA, FS | GPSWD |
| | 202 | M | AE | <13 | GSA | GPSWD |
| 449 | 201 | F | JAE | 13+ | GSA, GTCS | GPSWD |
| | 202 | F | JAE | 13+ | GSA | GPSWD |
| 451 | 202 | F | JAE | 13+ | GSA, FS | GPSWD |
| | 203 | F | JAE | 13+ | GSA | GPSWD |
| 452 | 995 | M | AE | <13 | GSA | GPSWD |
| | 996 | F | AE | <13 | GSA | GPSWD |

| Family | Individual | Sex | Epilepsy | Age at onset | Seizures | EEG |
|---|---|---|---|---|---|---|
| 525 | 201 | F | AE | 10 | GSA | GPSWD |
| | 202 | F | AE | 12 | GSA, GTCS | GSWD |
| | 203 | F | CAE | 6 | GSA | GSWD |
| 528 | 101 | M | IGE | 11 | GSA, GTCS | |
| | 204 | F | AE | 5 | GSA | GPSWD |
| | 301 | M | IGE | 2 | GSA, GTCS | Non-specific abnormalities |
| | 302 | M | AE | 1 | GSA | GSWD |
| 873 | 101 | M | JAE | 2-12 | GSA | GPSWD |
| | 201 | F | JAE | 2-12 | GSA | GPSWD |
| | 202 | F | JAE | 2-12 | GSA | GPSWD |
| 876 | 201 | M | JAE | 2-12 | GSA | GPSWD |
| | 202 | F | JAE | 2-12 | GSA | GPSWD |
| 881 | 201 | M | JAE | 2-12 | GSA | GPSWD |
| | 202 | F | JAE | 2-12 | GSA | GPSWD |
| | 203 | M | JAE | 2-12 | GSA | GPSWD |
| 882 | 201 | M | JAE | 2-12 | GSA | GPSWD |
| | 202 | M | JAE | 2-12 | GSA | GPSWD |
| 883 | 201 | F | JAE | 2-12 | GSA | GPSWD |
| | 202 | F | JAE | 2-12 | GSA | GPSWD |
| 884 | 201 | F | JAE | 2-12 | GSA | GPSWD |
| | 202 | F | JAE | 2-12 | GSA | GPSWD |

CAE: childhood absence epilepsy; JAE: juvenile absence epilepsy; AE: absence epilepsy; GSA: generalised absence seizure; GSAmy: generalised absence seizure with myoclonus; FS: febrile seizure; GSWD: generalised spike-wave discharges; GPSWD: generalised polyspike-wave discharges

## 5.8 Additional parent-child trios used for association analysis

Two hundred and seventeen additional parent-child trios were ascertained. Each trio comprised an affected child (204 CAE, 5 JAE and 8 AE) and two parents (2 CAE, 1JAE, 4 IGE, 427 unaffected). See Appendix 13 for clinical details. The complete cohort of pedigrees and trios was used for association analysis using intragenic SNPs, as described in Chapter 9.

## 5.9 Family with absence epilepsy and episodic ataxia

A collaboration was established with Dr SL Jaffe, Department of Neurology, Louisiana State University School of Medicine, Shreveport, USA. Dr Jaffe identified a North American family in which six members in three generations exhibited a typical primary generalized 3Hz spike-wave EEG abnormality (Table 5-4, Figure 5-6)(Imbrici, Jaffe et al. 2004). Five of these individuals exhibited clinical absence epilepsy with variable degrees of episodic and anticonvulsant drug induced cerebellar ataxia (AEA).

One individual had a typical 3Hz spike–wave EEG but had no clinical phenotype, and was neurologically normal (individual 319). Dr Jaffe had previously performed clinical, EEG and in some cases brain imaging on 11 genetically related individuals and one non-blood relative who had married into the family.

Two individuals were characterised in detail:

Individual 210 was diagnosed with EA2 with absence epilepsy. She presented at the age of 39 years with a history of 'dizzy' attacks from her teenage years. A typical attack consisted of sudden onset disequilibrium with associated headache and nausea, lasted between one and eight hours and could be precipitated by stress and anxiety. There was also a history of vacant episodes suggestive of absence seizures from late childhood which subsided in adult life. Typically these episodes lasted a few seconds with immediate recovery, but on some occasions, there were more prolonged atypical absence episodes. On one occasion a prolonged fugue-like state lasting several hours was accompanied by generalized 3 Hz spike–wave discharges. MRI brain scan demonstrated cerebellar vermian atrophy. Her symptoms remitted on a combination of acetazolamide and carbamazepine, although the ataxia was worsened by increased drug levels.

Individual 320 presented with a more severe younger onset phenotype similar to the case of EA2/epilepsy reported by Jouvenceau et al.

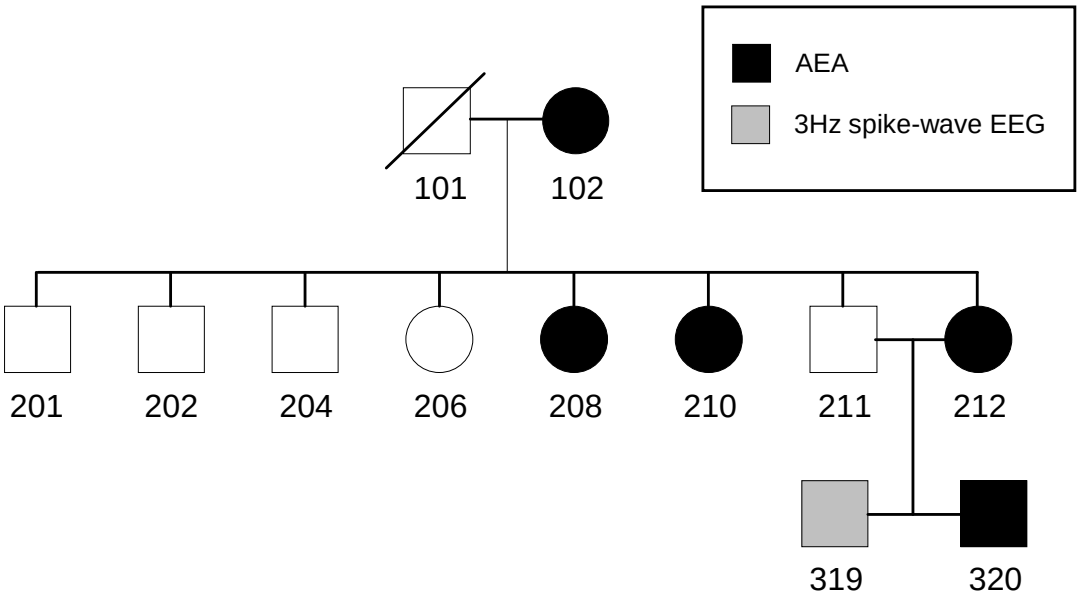
(Jouvenceau, Eunson et al. 2001). Typical brief absence seizures developed from age 4 years. Early motor milestones were delayed and he had learning difficulties. An initial EEG showed a generalised polyspike–wave abnormality; further EEGs demonstrated generalized 3 Hz spike–wave discharges. Seizures were difficult to control, and AEDs such as phenytoin at low doses would precipitate cerebellar ataxia with impaired consciousness at times. On one occasion, he developed extreme cerebellar ataxia and a decreased level of consciousness. Examination at the age of 10 years revealed moderate cerebellar signs.

**Table 5-4** *Family with absence epilepsy and ataxia*

| Individual | Sex | Age at onset | Seizures | Ataxia | EEG |
|---|---|---|---|---|---|
| 102 | F | Teens | AS | EA, CA | 3Hz SWD |
| 208 | F | 57 | - | CA | 3Hz SWD |
| 210 | F | 13 | AS | EA | 3Hz SWD |
| 212 | F | 10 | AS | EA, CAA | 3Hz SWD |
| 319 | M | - | - | - | 3Hz SWD |
| 320 | M | 4 | AS | EA, CA, CAA | 3Hz SWD |

AS: absence seizures; CA: clinical signs of cerebellar ataxia; CAA: significant worsening of cerebellar ataxia on exposure to AEDs; SWD: spike-wave discharges

**Figure 5-6** *Family with absence epilepsy and ataxia (AEA)*



# *Chapter 6 Laboratory Methods*

## 6.1 Summary

This chapter outlines the laboratory methodology employed from the arrival of the blood and buccal samples to the evaluation of the results. Standard solutions for molecular biological work were used.

Departmental laboratory procedures had previously been standardised and included the following criteria:

1. A copy of the departmental laboratory safety document was received and relevant safety issues were noted.

2. Sterile disposable gloves and white coats were worn at all times

3. Care was taken to ensure the accurate labelling of samples.

4. Standard precautions were taken to minimise the possibility of DNA contamination.

5. Disposable tips and eppendorf tubes were autoclaved prior to use.

## 6.2 Genomic DNA extraction from whole blood (Miller, Dykes et al. 1988)

5-20 ml blood samples were obtained and added to 50-200μl 0.34M EDTA or acid citrate dextrose (ACD). EBV transformed lymphocyte cell lines were established at CAMR Porton Down using blood samples from a subset of affected individuals (Appendix 14).

On arrival in the laboratory blood samples were decanted into a 50 ml labelled conical tube (Falcon) and centrifuged at 4$^{o}$C for 15 minutes at 1750g in an MSE Mistral 61 centrifuge. The plasma and cell fractions were then separated. DNA extraction was performed immediately on the cell fraction where possible, or else this was stored at -20$^{o}$C and thawed when required.

The solutions and protocol used for DNA extraction are described below:

Solution 1:

To make up 1 litre of concentration X3

| | |
|---|---|
| 10mls | 1M Tris HCL |
| 1.02g | MgCl2 ( 5ml 1M ) |
| 10mls | Triton X - 100 |
| 109.54g | Sucrose |
| 0.2g | Sodium Azide |

Make up to 1 Litre with distilled water.

Solution 2:

To make up a litre of 1X concentration

| | |
|---|---|
| 50mls | 1M Tris HCL |
| 40mls | 0.5M EDTA |
| 20g or 200mls | 10% SDS |

Solution 3

This is a supersaturated NaCl solution.

## Stratagene method for DNA extraction from whole blood

1. The blood samples were thawed where necessary and made up to 45-50 mls with distilled water. These solutions were left on ice for 5 minutes and then centrifuged at approximately 1500g for 15 minutes at 4oC.
2. The supernatants were removed by pouring gently to leave around 10mls.
3. 15ml of (3X) Solution 1 and 30 ml of distilled water were added and the samples mixed thoroughly to disperse pellet.
4. The samples were incubated on ice for 2 minutes.
5. The samples were centrifuged at approximately 1500g at 4 $^{o}$C for 15 minutes.
6. The supernatants were removed by pouring gently to leave a light pink pellet. If a dark red pellet was present steps 1 - 4 were repeated.
7. The pellets were resuspended in 11 mls of solution 2 and vortexed vigorously.
8. Pronase was added to yield a final concentration of 100μg/ml.
9. The samples were incubated at 60 o C with shaking for 1 hour or 37 $^{o}$C overnight.
10. The tubes were incubated on ice for 10 minutes.
11. 4ml of Solution 3 was added to each sample and the tubes were inverted several times to mix.
12. The samples were incubated on ice for 5 minutes.
13. The samples were centrifuged at approximately 1500g for 15 minutes at 4$^{o}$C.
14. The supernatants were transferred using a large bore pipette to sterile 50ml conical tubes that were labelled with the sample codes.
15. RNase was added to yield a final concentration of 20μg/ml.
16. The samples were incubated at 37$^{o}$C for 15 minutes.
17. Adding two volumes of 100% ethanol precipitated the DNA.
18. The DNA was removed by spooling out with a glass rod (or by centrifuging gently to leave a DNA pellet )
19. The DNA was rinsed on the rod in 70% ethanol and dried by briefly in air.
20. The DNA was resuspended in 500μg of Tris - EDTA buffer and stored at 4$^{o}$C.
21. Stock DNA was stored at -20$^{o}$C and working concentrations of 10ng/μl were made using sterile distilled water and stored at 4$^{o}$C.

The initial supernatant was kept until it was confirmed that DNA had been isolated. A modified Stratagene method, using smaller solution volumes, was used to extract DNA from small blood volumes (1-2ml). On occasions when DNA was not isolated a DNA extraction kit (Pharmacia Biotech) was used to isolate small amounts of DNA from the supernatant.

The chloroform phenol extraction method outlined below was used to improve the purity of DNA when necessary.

1. One volume of phenol was added to each DNA sample
2. The sample was centrifuged at high speed for 2 minutes.
3. The upper phase was pipetted into a separate clean eppendorf and the phenol phase was discarded into the phenol waste bucket.
4. One volume of chloroform:isoamyl alcohol (CHCL3.IAA 24:1) Was added to the sample
5. The sample was centrifuged at high speed for 1 minute.
6. The upper phase was pipetted into a clean eppendorf and the phenol phase was discarded into the chloroform / phenol waste bucket.
7. 1/10th of a volume of 3MNaOAC pH 5.2 was added to the sample.
8. Two volumes of 100% Ethanol were added to precipitate the DNA.
9. The DNA was suspended in TE as before.

## 6.3 Genomic DNA extraction from buccal cells

The cheek swabs were stored in a preservative solution of 0.05M EDTA, 0.5% SDS. If extraction could not be performed within 10 days, the swabs were frozen at -20 $^{o}$C and thawed for DNA extraction.

**Cheek swab DNA extraction method (provided by Dr Mark Thomas, Department of Biology, UCL):**

1. To each cheek swab in its EDTA/SDS solution, 0.8 ml of water/proteinase K mix (40µl of 10 mg/ml proteinase K in 20 ml of sterile distilled water) was added and incubated at 56°C for between 30 and 60 minutes minimum (*samples could be incubated for up to 3 hours*).
2. 0.6 ml of a phenol/chloroform (1:1) mix was placed in each of 24 numbered microfuge tubes in a rack. 0.8 ml of the EDTA/SDS solution from the cheek swab collection tube was added to the numbered microfuge tubes containing the phenol/chloroform. The remaining cheek swab solution was stored at -20$^{o}$C.
3. Samples were mixed and centrifuged for 10 minutes at maximum speed in a microfuge.
4. 0.6 ml of chloroform and 30 µl of 5 M NaCl were placed in each of 24 clean microfuge tubes in a rack. The aqueous (upper) phase from the centrifuged sample tubes was transferred into the numbered microfuge tubes containing chloroform and 5M NaCl (*care was taken to avoid transferring any of the white material at the interface between the two phases of the phenol/chloroform*).
5. Samples were mixed and centrifuged for 10 minutes at maximum speed in a microfuge.
6. 0.7 ml of chloroform was placed in each of 24 clean microfuge tubes in a rack. The aqueous (upper) phase from the centrifuged sample tubes was transferred into the numbered microfuge tubes containing chloroform (*care was taken to avoid transferring any of the white material at the interface between the two phases*).
7. Samples were mixed and centrifuged for 10 minutes at maximum speed in a microfuge.
8. 0.7 ml of isopropanol was placed in each of 24 clean screw top-microfuge tubes in a rack. The aqueous (upper) phase from the centrifuged sample tubes was transferred into the numbered microfuge tubes containing isopropanol (*care was taken to avoid transferring any of the white material at the interface between the two phases*).
9. The microfuge tubes were mixed and placed in a freezer for at least 2 hours (*samples could be stored in this state indefinitely*).
10. The microfuge tubes were centrifuged for 12 minutes at maximum speed in a microfuge.

11. The supernatant was carefully poured from the sample tubes and allowed to drain by placing the tube inverted at a 45˚ angle on some tissue paper for about 1 minute. This left the precipitated DNA stuck to the walls of the screw-top microfuge tubes.
12. To each sample tube, 0.8 ml of 70 % ethanol was added and the tubes were centrifuged for 10 minutes at maximum speed in a microfuge (*with the tubes are orientated so that the DNA pellet was outermost in the centrifuge rotor*).
13. The supernatant was carefully poured from the sample tubes and allowed to drain by placing the tube inverted at a 45˚ angle on some tissue paper for about 20 mins. This left the precipitated DNA on the walls of the screw-top microfuge tubes.
14. To each sample tube, 200 µl of TE (pH 9.0) was added. The tubes were incubated at 56˚C for 10 min in a water bath, mixing occasionally. They were then centrifuged briefly and stored at -20˚C.

## 6.4 Quantitation of DNA

Two methods of DNA quantitation were used: spectrophotometry and PicoGreen fluorescence.

### 6.4.1 Spectrophotometry

The quantity of genomic DNA that had been isolated was estimated using an M3 02 Spectrophotometer (Cam-Spec Ltd). The DNA was diluted to 1:200 and the optical density (OD) of this solution was then obtained at a wavelength of 260nm and 280nm.

At 260nm a reading of 1 OD corresponds to a DNA concentration of 50μg/ml. The concentration of the stock DNA could therefore be calculated using the formula:

DNA concentration in μg/ml = OD260nm X 50 X 200 μg/ml

The concentration of DNA obtained ranged from 50μg/ml (50ng/μl) to 500μg/ml (500ng/μl) indicating a total yield of DNA between 25μg and 250μg. The stock DNA was then diluted to make working DNA samples of 5-10ng/ml. The stock DNA was stored at -20$^{o}$C and the working DNA samples were stored at 4$^{o}$C.

The purity of the DNA sample was also measured by calculating the ratio OD260nm/OD280nm. Acceptable values for this ratio lie between 1.8 and 2.0. Values below this prevent accurate quantification of the concentration of nucleic acid and in these situations the DNA sample was further purified using a chloral/phenol extraction method as detailed above.

### 6.4.2 PicoGreen fluorescence

PicoGreen is a fluorescent dye that undergoes a marked fluorescence enhancement upon binding to double-stranded DNA (dsDNA). This can be

measured using a microplate fluorometer, and is an accurate method for determining dsDNA concentrations in solution (Singer, Jones et al. 1997). In collaboration with Professor Paul Mckeigue at the London School of Hygiene and Tropical Medicine, DNA quantitation was performed on samples in a 96-well plate format using a an automated fluorometer.

## 6.5 Whole genome amplification

For samples where the yield of DNA was very low, or where most of the stored DNA had been used previously, whole genome amplification was used to increase the DNA available. The method of multiple displacement amplification (MDA), which uses Phi29 DNA polymerase and random exonuclease primers to synthesize high fidelity genomic DNA, has been shown to yield an average DNA product > 10 kb in length with relatively uniform amplification across the genome and an error rate of 1 in $10^6$–$10^7$ (Esteban, Salas et al. 1993; Dean, Hosono et al. 2002). The DNA generated is an accurate representation of genomic DNA when used for SNP genotyping (Tranah, Lescault et al. 2003).

Whole genome amplification method using the TempliPhi Amplification Kit (Amersham Biosciences Piscataway, NJ) (Tranah, Lescault et al. 2003):

1. 1 µl DNA was diluted in 5 µl sample buffer in 0.2 ml tubes
2. Samples were incubated for 3 min at 95°C and then placed on ice.
3. 5 µl reaction buffer and 0.2 µl enzyme mix were added to each sample in 10 µl reactions.
4. DNA samples were incubated for 6 h at 30 ◦C.
5. After incubation, samples were heated to 65 ◦C for 10 min to inactive the enzyme.

DNA was quantified as described above.

## 6.6 Genotyping

### 6.6.1 Polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) was used to amplify genomic DNA for genotyping. Fluorescently labelled primers were selected to amplify DNA sequences encompassing simple sequence length polymorphisms (SSLPs). SSLPs were selected to encompass candidate genes, and were identified from genetic databases such as Genethon and the UCSC Genome Browser. Novel SSLPs were identified from the publicly available genomic DNA sequence and custom primers were designed using Primer3 software from the Whitehead Institute for Biomedical Research (Rozen and Skaletsky 2000)(Appendix 15). Commercially available primers were obtained from Research Genetics. Custom labelled primers were obtained from MWG Biotech UK Ltd. Primers were fluorescently labelled with one of three fluorescent dyes (HEX,TET and FAM) which appear as yellow, green and blue respectively under the ABI laser filter conditions used.

A 15 μl standard reaction volume used for genotyping and each reaction mixture contained 25-50ng of DNA (5-10μl). DNA samples for amplification were pipetted into wells of a 96 well microtitre plate (Hybaid). A mastermix incorporating the remaining ingredients of the reaction was then added using a dedicated DNA free pipette. The mastermix comprised:

| | | Final Concentration |
|---|---|---|
| PCR Buffer X 10 (Gibco) | 1.5μl | 1X |
| $MgCl_2$ 50mM (Gibco) | 0.6μl | 2.0mM |
| dNTPs 25mM | 0.3μl | 0.5mM |
| Forward Primer (8μM) | 0.5μl | 0.25μM |
| Reverse Primer (8μM) | 0.5μl | 0.25μM |
| Taq (5units/μl) | 0.05μl | 0.25units |
| Sterile distilled water | 1.55-6.55μl | |

The PCR/mastermix mixture was covered with one drop of mineral oil (Sigma) to prevent evaporation. The microtitre plate was protected with a thin self adhesive film and placed on an Omnigene thermocycler (Hybaid).

The standard PCR protocol for the thermocycler is as follows:

**1 cycle:**

| | | |
|---|---|---|
| Denaturation | 94$^{o}$C | 1 min 30 seconds |
| Annealing | 55$^{o}$C | 30 seconds |
| Extension | 72$^{o}$C | 30 seconds |

**30 cycles:**

| | | |
|---|---|---|
| Denaturation | 94$^{o}$C | 30 seconds |
| Annealing | 55$^{o}$C | 30 seconds |
| Extension | 72$^{o}$C | 30 seconds |

**1 cycle:**

| | | |
|---|---|---|
| Denaturation | 94$^{o}$C | 30 seconds |
| Annealing | 55$^{o}$C | 30 seconds |
| Extension | 72$^{o}$C | 2 minutes |

Gel electrophoresis was used to assess the success of the PCR. This enabled the PCR products to be visualised under ultraviolet light (Figure 6-1). The process comprised:

1. 5μl of the PCR product was added to 3μl of bromophenol blue loading buffer
2. The fragments were separated by electrophoresis at 100 volts on a 2% agarose minigel in 1XTBE buffer.
3. The gel was stained with ethidium bromide and visualised under ultraviolet light.
4. A standard 1kb ladder was used to confirm that the correct size product had been amplified.

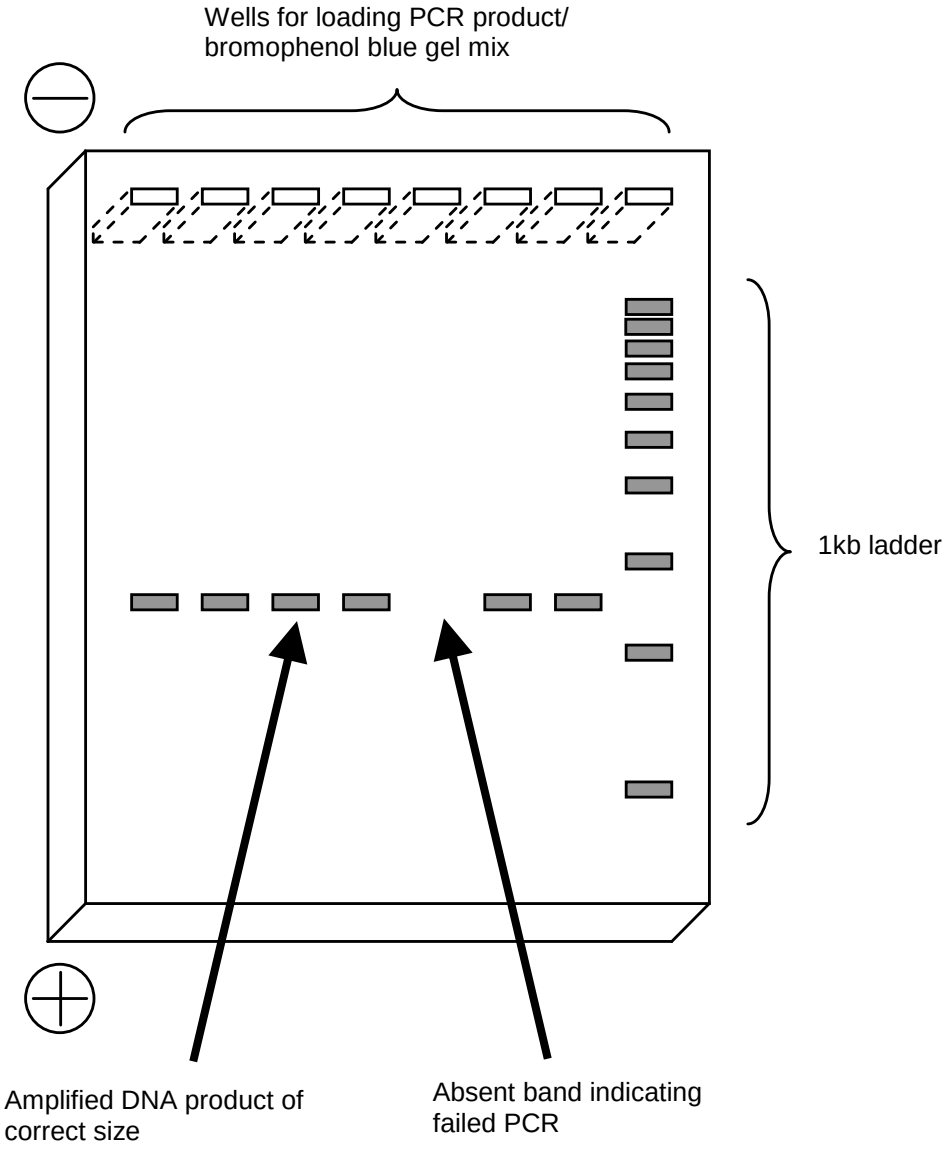



**Figure 6-1** *Agarose Gel showing size marker and amplified DNA fragments*

In situations where a specific product of the correct size was not identified the PCR was optimised by varying the annealing temperature or by performing a magnesium titration.

Once a PCR product had been successively amplified the 96 well plates were stored at 4$^{o}$C.

### 6.6.2 Polyacrylamide gel electrophoresis

Allele sizing of the PCR products was performed by electrophoresis on a denaturing 6% Polyacrylamide gel. A Sequagel reagent kit (National Diagnostics) was used to prepare the gel mix as follows:

| | |
|---|---|
| 32ml | Sequagel 6 mix |
| 8ml | Sequagel buffer |
| 300ml | 10% Ammonium Persulphate |

The glass plates were washed, rinsed with distilled water and allowed to dry before being assembled according to the manufacturer's instructions using 0.4mm spacers. The polyacrylamide solution was poured between the glass plates using the sliding plate method and the sharks tooth comb inserted. The gel was then left to polymerise for 2-3hours.

The polymerised gel was loaded into the ABI 373 and a plate check performed to establish that the plates and polyacrylamide were free from contamination. Freshly diluted 1X TBE buffer was prepared and added to the two buffer chambers placed at the bottom and the top of the gel plates prior to running the gel electrophoresis (Figure 6-2).

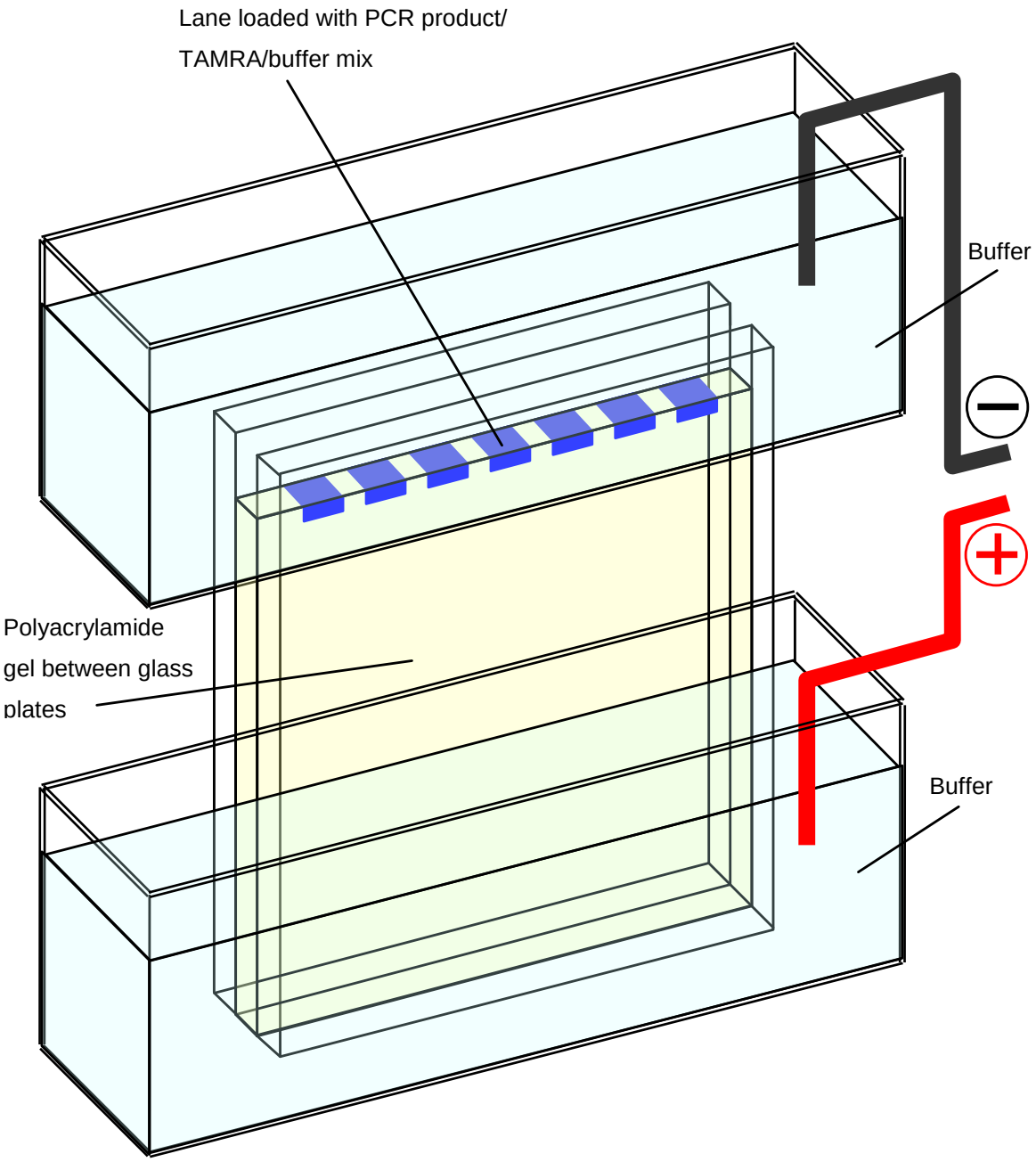



**Figure 6-2** *Polyacrylamide gel loaded into ABI373 sequencer*

### 6.6.3 Allele detection

Allele detection was performed on an Applied Biosystems 373A semi-automated sequencer (ABI373A) with GENESCAN software (Perkin Elmer). The main elements of allele detection were as follows.

1. The ABI373A can allow the simultaneous analysis of between 36 and 64 lanes. In general, a 48 well polyacrylamide gel was prepared.

2. Each lane can potentially contain up to 24 different markers. Typically 5-6 PCR products were loaded together in each lane. The PCR products to be analysed were pooled together according to their fluorescent label and allele sizes (Table 6-1).

3. A DNA size standard labelled with TAMRA was used. This shows up as red under standard laser filter conditions.

4. Each lane was loaded with a pooled aliquot containing: 0.1-0.2 ml of each PCR product; 0.5μl of 500-TAMRA standard; 4μl of blue formamide buffer (5 parts deionised formamide, 1 part 25mM EDTA containing 50 mg/ml blue dextran). The pooled PCR products were denatured at 94$^{o}$C for 3 minutes before being placed on ice and loaded into the wells.

5. Gel electrophoresis was used to separate the PCR products. The Genotyping gels were run under standard conditions using GENESCAN software for 3-5 hours at 620 volts using filter set B.

6. Laser based technology was used to size the PCR products in comparison to the standard (Figure 6-3).

7. Once the GENESCAN run was completed the information was extracted from the gel using the GENESCAN and GENOTYPER software as described below.

**Table 6-1** *Example of microsatellite pool in a genescan*

| LOCUS | ASSAY | LABEL | SIZE |
|---|---|---|---|
| D3S3582 | AFMa231xe9 | HEX | 220-236 |
| D3S1588 | AFM287yd9 | TET | 212-236 |
| D3S1289 | AFM198yf2 | FAM | 197-215 |
| D1S212 | AFM212xb10 | TET | 105-125 |
| D1S2848 | AFM348tg1 | FAM | 105-123 |

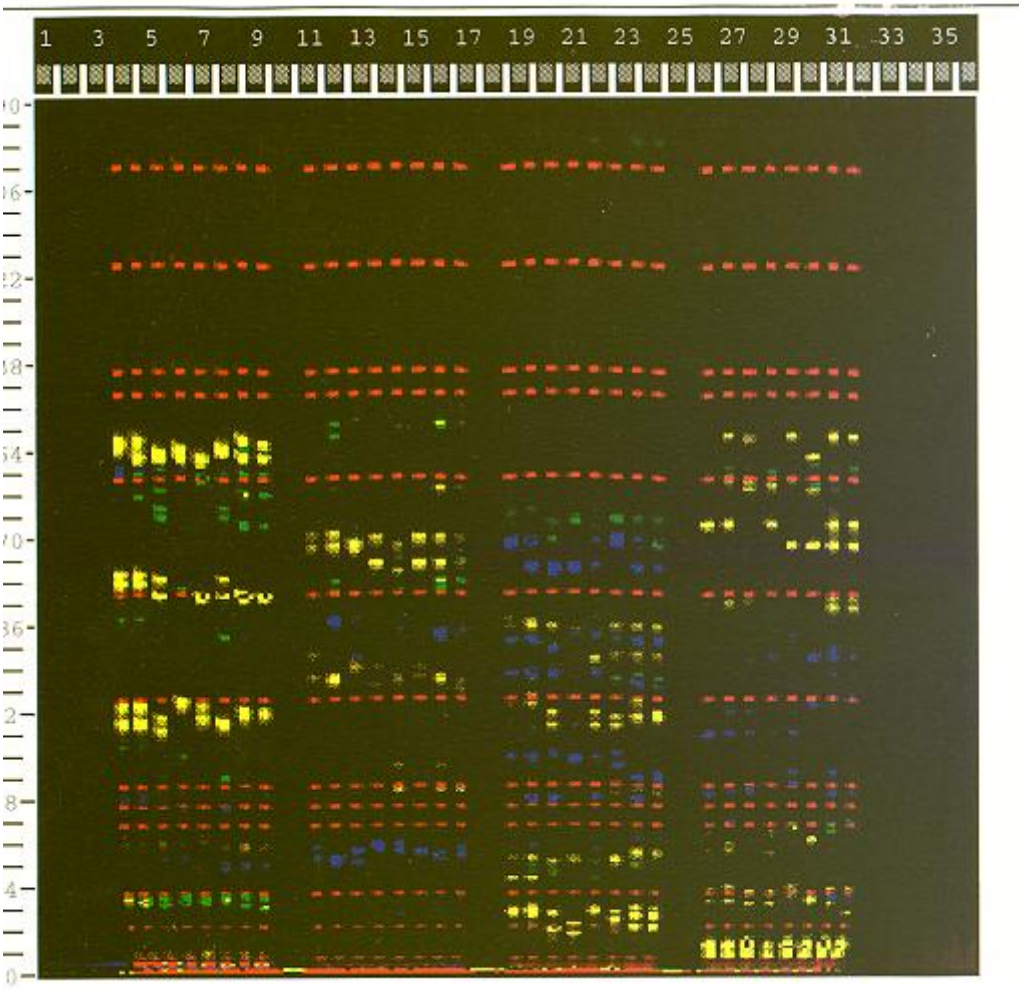



**Figure 6-3:** *Genotyper gel comparing allele sizes to marker sizes. The yellow, green and blue fluorescent dyes indicate the DNA fragments that have been amplified. These are sized by comparison to the size standard that is labelled in red.*

**GENESCAN (Applied Biosystems) program.**

The details of the TAMRA labelled molecular weight standard were entered. The fluorescently labelled PCR product alleles in each lane of the gel were aligned with the TAMRA standard and the allele size information contained within each lane extracted. GENESCAN uses the size standard to calculate the allele sizes of the PCR products. The GENESCAN algorithms automatically identify and size each PCR product peak relative to the internal size standard, as well as providing peak area and peak height information. This data can then be imported into GENOTYPER for further analysis.

**GENOTYPER (Applied Biosystems) program.**

The GENOTYPER software uses the size and fluorescent labelling characteristics of each SSLP marker to analyse the data from the genescan gel. The results for each marker can be viewed in the form of an electropherogram, The allele sizes for each marker are calculated and output in tabular form, and can be checked manually against the electropherograms (Figure 6-4).

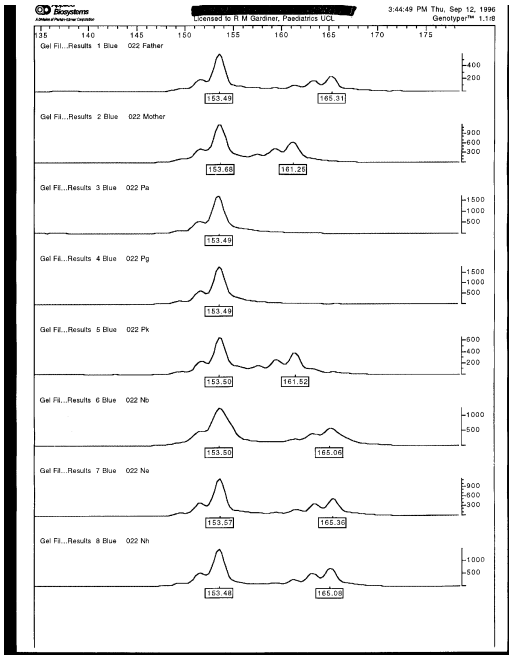



**Figure 6-4** *Allele Sizes. The x axis represents allele size and the y axis the intensity of the fluorescence recorded.*

## 6.7 Single nucleotide polymorphism analysis

Several methods for SNP typing were explored, including melting curve analysis of SNPs (McSNP®) (6.7.1), TaqMan® (section 6.7.1), Pyrosequencing® (section 6.7.3) and Amplifluor® (section 4.3.8). Initial SNP typing by Taqman and Pyrosequencing was performed by the SNP genotyping services at MRC Geneservice, HGMP. The majority of SNPs were typed by KBiosciences (http://www.kbioscience.co.uk) using Amplifluor® and TaqMan® chemistries.

### 6.7.1 McSNP

McSNP combines restriction enzyme digestion with melting curve analysis (Figure 6-5). Restriction enzyme digestion is used to discriminate SNP alleles by selectively cleaving DNA. In the instances where a SNP does not cause a restriction enzyme site difference (approximately 50% of SNPs), a restriction site for one of the SNP alleles can be introduced by designing one of the PCR primers with a mismatch near the 3' end. This increases the proportion of SNPs suitable for restriction enzyme analysis to about 98%. Melting curve analysis is a technique for detecting double-stranded DNA fragments of varying lengths. The melting temperature, Tm, of a particular fragment is dependent on the fragment length, sequence composition, and GC content. The DNA fragments are slowly heated in the presence of a DNA intercalating dye, SYBR Green I, which fluoresces in the presence of double stranded DNA. When the melting temperature of a particular fragment is reached, the fluorescence decreases. Real-time fluorescent monitoring is performed with a DASH machine (section 4.3.6) which produces a melting curve for the DNA fragments in the sample. The pattern of melting curve peaks reflects the DNA fragments present. McSNP is suited for high-throughput SNP genotyping.

In collaboration with Professor Paul Mckeigue at the London School of Hygiene and Tropical Medicine, attempts were made to establish an

automated SNP genotyping facility using the McSNP protocol, with a software package and DASH machine from ThermoHybaid. Despite significant time and effort spent on this project, unfortunately several problems were encountered which ultimately led to alternative SNP typing methods being used. The main problems were as follows:

1. Robot malfunction. Two robots were installed for automated handling of DNA samples and reagents in 96 well plate formats. Software and hardware problems resulted in prolonged robot malfunction.

2. Assay design software. Software was provided by Hybaid to identify/create restriction sites for differentiating SNP alleles and to design PCR primers. This software never functioned successfully and was unable to design assays for the majority of SNPs.

3. SNP assays that were successfully designed were run on the DASH machine, but the melting curves obtained could not differentiate the SNP alleles.

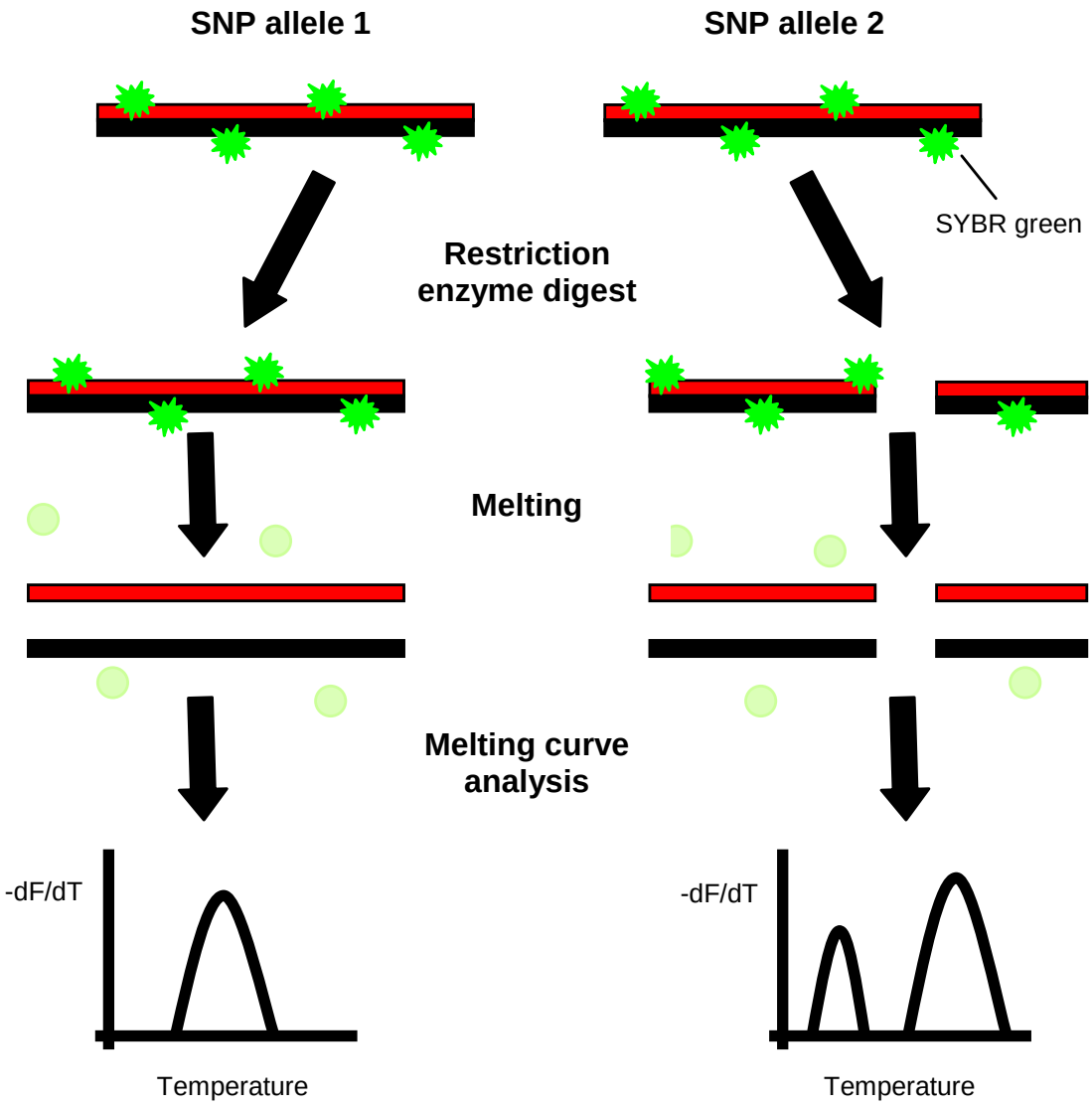



**Figure 6-5** *Melting curve SNP analysis (McSNP)*

### 6.7.2 TaqMan

TaqMan uses the principle of fluorescence resonance energy transfer (FRET), as described in section 4.3.7. The TaqMan assay involves alternative allele-specific probes which are used during PCR amplification (Figure 6-6). Each probe is doubly labelled with a 3' fluorescent reporter and a 5' universal quencher. During the annealing phase of the PCR reaction the probe hybridises to the target DNA, and during the extension phase, the 5'-3' exonuclease activity of Taq DNA polymerase degrades only perfectly matched, annealed probes which separates the quencher from the reporter and produces fluorescence. Mismatched probes are displaced from the target without degradation. Allele-specific probes are labelled with different fluorophores.

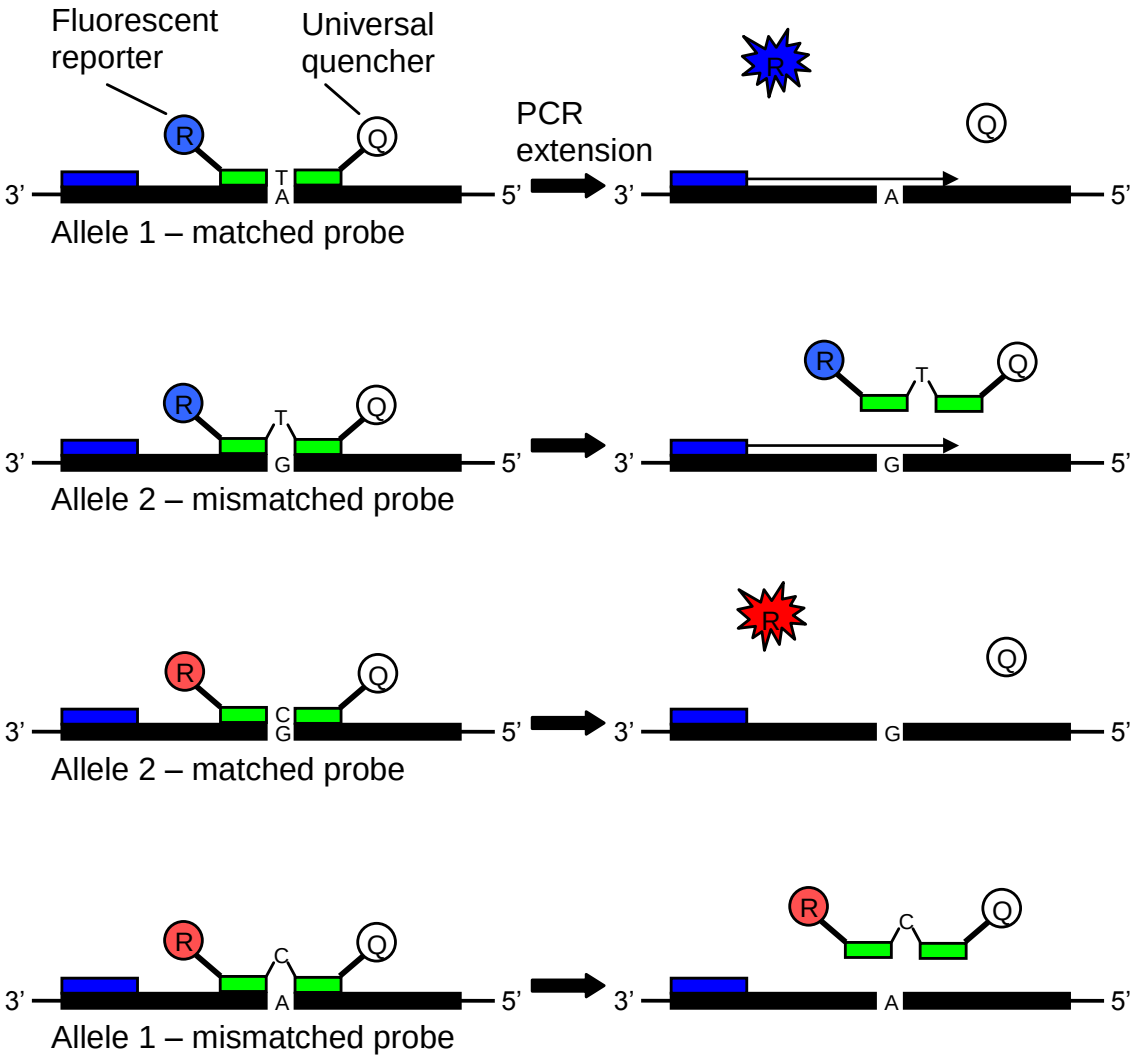



**Figure 6-6** *TaqMan SNP genotyping*

### 6.7.3 Pyrosequencing

Pyrosequencing detects the pyrophosphate which is released whenever a dNTP is incorporated during primer extension (Figure 6-7). A primer is used that anneals to the target DNA immediately adjacent to the SNP. Primer extension by DNA polymerase only occurs with the nucleotide complementary to the nucleotide at the site of the SNP. The released pyrophosphate triggers enzymatic conversion of luciferin to oxyluciferin which generates visible light. The four nucleotides are sequentially added and degraded by apyrase to determine short DNA sequences flanking a SNP (or several closely spaced SNPs).

In this study pyrosequencing was used predominantly for validation of novel SNPs in a panel of control DNAs.

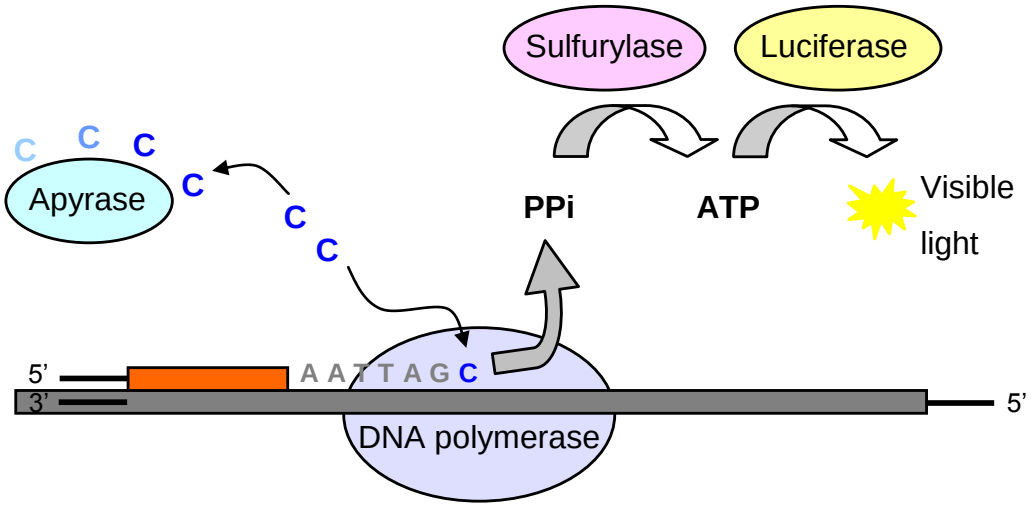



**Figure 6-7** *Pyrosequencing reaction*

## 6.8 Sequencing of Candidate Genes

Exonic DNA and surrounding intronic and promoter sequence DNA of *CACNG3* was sequenced in a subset of affected individuals. The genomic sequence had been published previously. Primers were designed for amplification of the selected sequences using Primer3 software (Table 6-2). PCR conditions were optimised for each pair of primers. The thermocycler conditions were as outlined previously (section 6.6.1).

**Table 6-2** *PCR primers used to amplify CACNG3 exons for sequencing Primers were designed using Primer3 software http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi. The PCR products for CACNG3EX1A and CACNG3EX1B overlap by 69 bp. The PCR products for CACNG3EX4A and CACNG3EX4B overlap by 85 bp.*

| Primer Name | Exon | Forward | Reverse | Product size (bp) | Exon size (bp) |
|---|---|---|---|---|---|
| *CACNG3EX1A*(F/R) | 1a | aagcagacctaacttcacctgc | aggtccgtggtgctgaag | 968 | 1411 |
| *CACNG3EX1B*(F/R) | 1b | aagcctaggcgttaagacga | ctccaaacacccaccagaat | 992 | |
| *CACNG3EX2*(F/R) | 2 | gttgatgactgatcccgttg | atgatcctcccaccttagcc | 951 | 84 |
| *CACNG3EX3*(F/R) | 3 | agaatgggcgtttgagagc | atgagggtctcactgtgctg | 964 | 141 |
| *CACNG3EX4A*(F/R) | 4a | ggcagaggcgtgtatcttgt | ccatgcaacctcagagacat | 948 | 1064 |
| *CACNG3EX4B*(F/R) | 4b | cccgtctgaactgacctctg | aactctgtggaggagagtgct | 918 | |

### 6.8.1 Purification of PCR products

Amplified PCR products were purified using the Quiagen PCR Purification Kit. This utilises a column containing a silica gel membrane which binds DNA in high-salt buffer and elutes the DNA with low-salt buffer or water. This purification procedure removes primers, nucleotides, enzymes, mineral oil, salts, agarose, ethidium bromide, and other impurities from DNA samples. The Kit was used according to the manufacturer's instructions.

### 6.8.2 Dideoxy-DNA sequencing

The majority of sequencing was carried out on an automated ABI 377 DNA sequencer by the sequencing service at the Advanced Biotechnology Centre (ABC), Imperial College London. Sequencing was also performed at UCL on the ABI 373 sequencer. In both cases the dideoxy-DNA sequencing method was used, as described in section 4.3.3.

### 6.8.3 Sequencing using ABI 373A

1. Sequencing kit. The TaqFS Dye Terminator Cycle Sequencing Kit (Perkin Elmer) was used. The kit contains AmpliTaq DNA Polymerase FS, dNTPs, dideoxynucleotide terminators and buffer in a premix and was used according to manufacturer's instructions.

2. Sequencing cycle. The sequencing cycle was performed using a Cetus thermocycler (Perkin Elmer) with the following settings:

   **25 cycles:**

   | | | |
   |---|---|---|
   | Denaturing | 96$^{o}$C | for 30sec |
   | Annealing | 50$^{o}$C | for 15sec |
   | Extension | 60$^{o}$C | for 4 minutes |
   | Hold | 4$^{o}$C | |

3. Purification of products. The products were purified using ethanol precipitation and the pellets were re-suspended in a blue dextran loading buffer. Pellets were stored at –20$^{o}$C prior to being re-suspended in loading buffer if necessary.

4. Sequencing gel. The sequencing gel was prepared from the following ingredients:

   | | |
   |---|---|
   | Urea | 30g |
   | 40% Acrylamide (Bio Rad) | 9ml |
   | dd.H2O | 23ml |
   | Amberlite mixed bed resin | 0.5g |
   | 10% Ammonium persulphate | 300μl |
   | TEMED | 33μl |

The urea, acrylamide, amberlite mixed bed resin and distilled water were mixed together at room temperature for 30 minutes before being filtered through a 0.2 micron cellulose acetate filter and degassed. The TBE, ammonium persulphate and TEMED were then added to this solution to enable polymerisation and the gel was poured between the sequencing plates. The gel was then left to set for 2-3hours.

5. Gel electrophoresis. ABI sequencing software was used. Prior to sample loading, the gel was pre-run for 1 hour in the ABI 373a using 1X TBE buffer. A comb with the required number of lanes was inserted into the gel and the DNA sample for sequencing was loaded. The Gel was run overnight for 12 hours using filter set A.

6. Both forward and reverse sequences were obtained to minimise sequencing errors.

### 6.8.4 Sequence analysis

The output from sequencing at both UCL and Imperial College was obtained as a four colour sequence electropherogram which could be checked manually and edited for sequencing errors (Figure 6-8). The Sequence Navigator program (Perkin Elmer) was used to compare the published sequence with the sequence of the sample DNA fragments.

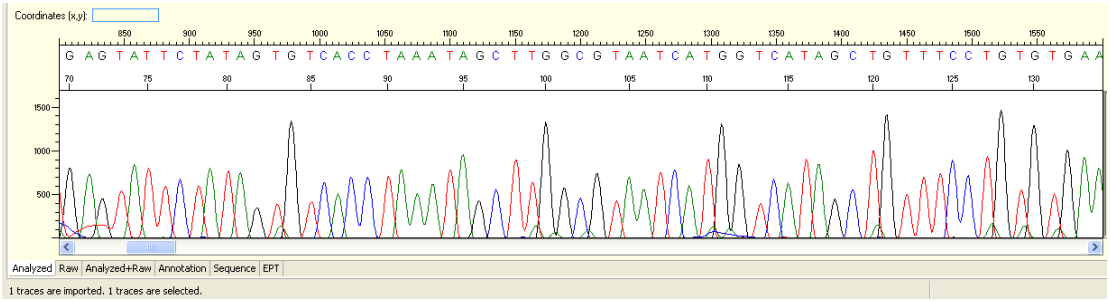



**Figure 6-8** *Sequence navigator output.*

# *Chapter 7 Candidate gene selection*

## 7.1 Introduction

Thirty-four ion channel genes were selected for testing in the initial 33 absence families by linkage analysis (Table 7-1). These included voltage dependent calcium channels, $GABA_A$ and $GABA_B$ receptors. The *ECA1* locus on chromosome 8q24 was also tested (Fong, Shah et al. 1998; Sugimoto, Morita et al. 2000). Genes were selected on the following basis:

- Expression pattern in brain
- A role in the generation of spike-wave seizures
- Potential targets for anti-absence drugs
- Genes known to cause spike-wave seizures in rodents or epilepsies in humans which include absence seizures as part of the phenotype
- Genes implicated in previous linkage and association studies.

## 7.2 Voltage dependent calcium channels (VDCC)

Seventeen VDCC subunit gene regions were selected for linkage testing in the 33 families. These markers represented genes encoding seven $\alpha$ subunits, one $\alpha 2\delta$ subunit, one $\beta$ subunit and eight $\gamma$ subunits. Twenty-one SSLPs were selected to encompass the relevant loci. Five of these markers were intragenic polymorphisms: *D19S1150* (*CACNA1A*), *D3S1568* (*CACNA2D2*), *A1GUCL* (*CACNA1G*), *D22S426* (*CACNG2*), and *D2S2236* (*CACNB4*).

VDCCs are involved in several calcium dependent processes, including entry of calcium ions into excitable cells, muscle contraction, hormone or neurotransmitter release, and gene expression. The channel activity is

directed by a pore-forming alpha-1 subunit (Figure 3-3). The auxiliary subunits beta, gamma and alpha-2-delta regulate channel activity. There are at least 10 classes of alpha-1 subunits: $\alpha$1A, B, C, D, E, F, G, H, I and S; 4 beta subunits: $\beta$1, 2, 3, 4; 8 gamma subunits: $\gamma$1,2,3,4,5,6,7,8; 4 alpha-2-delta subunits: $\alpha 2\delta$1, 2, 3, 4. There is strong evidence for a role for VDCCs in the generation of absence seizures. VDCCs are involved in thalamocortical spike-wave activity (section 3.5.1). Mutations in VDCC genes cause absence seizures in humans and mice (sections 3.5.2, 3.5.4). The anti-absence drug ethosuximide may have an action on the T-type Calcium current (section 3.5.3).

### *7.2.1 CACNA1A*

*CACNA1A* encodes the $\alpha$1A subunit which gives rise to P and/or Q-type calcium currents. These are high-voltage activated (HVA) currents. *CACNA1A* is expressed specifically in brain and is mainly found in the cerebellum, cerebral cortex, thalamus and hypothalamus. P/Q type calcium channels mediate the entry of calcium ions into excitable cells and are involved in calcium-dependent processes including muscle contraction, hormone or neurotransmitter release, gene expression, cell motility, cell division and cell death. Mutations in *CACNA1A* cause a range of neurological disorders, including familial hemiplegic migraine (Ophoff, Terwindt et al. 1996), episodic ataxia type 2 (Ophoff, Terwindt et al. 1996), spinocerebellar ataxia type 6 (Zhuchenko, Bailey et al. 1997) and absence epilepsy with episodic ataxia (Jouvenceau, Eunson et al. 2001). *Cacna1a* mutations cause the tottering phenotype in mice (section 3.5.2). A significant association has been demonstrated between a *CACNA1A* intragenic SNP and patients with IGE (Chioza, Wilkie et al. 2001).

### *7.2.2 CACNA1B*

*CACNA1B* encodes the $\alpha$1B subunit which gives rise to the N-type calcium current, a HVA current. It is highly expressed in the central nervous system.

Calcium channels containing α1B subunit may play a role in directed migration of immature neurons (Komuro and Rakic 1992).

### 7.2.3 CACNA1D

*CACNA1D* encodes the α1D subunit which gives rise to the L-type calcium current, a HVA current. *CACNA1D* is expressed in brain, mainly in hippocampus, basal ganglia, habenula and thalamus. It is also expressed in pancreatic islet cells. Functionally, L-type calcium channels in suprachiasmatic neurons contribute to the generation of spontaneous oscillations in membrane potential, which are tightly coupled to spike generation (Pennartz, de Jeu et al. 2002).

### 7.2.4 CACNA1E

*CACNA1E* encodes the α1E subunit which gives rise to the R-type calcium current, a HVA current. *CACNA1E* is expressed in brain and kidney. Calcium channels containing α1E subunits may be involved in the modulation of firing patterns of neurons and synaptic plasticity (Breustedt, Vogt et al. 2003). R-type currents are inhibited by Lamotrigine (Hainsworth, McNaughton et al. 2003). *EFHC1*, mutations of which can cause JME, was shown to enhance the R-type current in vitro (Suzuki, Delgado-Escueta et al. 2004). This enhancement was partly reversed by mutations associated with JME, suggesting that disruption of R-type calcium current homeostasis is involved in seizure generation.

### 7.2.5 CACNA1G

*CACNA1G* encodes the α1G subunit which gives rise to the T-type (transient, low-threshold) calcium current, a low-voltage activated (LVA) current. *CACNA1G* is highly expressed in brain, mainly amygdala, subthalamic nuclei, cerebellum and thalamus. It is also moderately expressed in heart; with low expression in placenta, kidney and lung, colon

and bone marrow. T-type channels serve pacemaking functions in both central neurons and cardiac nodal cells, support calcium signalling in secretory cells and vascular smooth muscle (Yunker and McEnery 2003). They have a major role in modulation of neuronal oscillations, resonance, and rebound burst firing. *CACNA1G* knockout mice are resistant to the generation of spike wave discharges in response to $GABA_B$ receptor activation (Kim, Song et al. 2001). The thalamocortical relay neurons of the knockout mice lacked the burst mode firing of action potentials, but showed the normal pattern of tonic mode firing. The *CACNA1G* mediated t-type calcium current appears to modulate the intrinsic firing pattern in the thalamocortical pathway, which generates the $GABA_B$ receptor-mediated spike-wave discharges typical of absence seizures.

### *7.2.6 CACNA1H*

*CACNA1H* encodes the $\alpha$1H subunit which gives rise to a T-type calcium current. It is expressed in brain, including thalamic reticular neurons, basal ganglia, olfactory bulb, hippocampus, caudal hypothalamus sensory ganglia, pituitary, and dentate gyrus (Talley, Cribbs et al. 1999). It is also expressed in kidney, liver, and heart. Sequencing of *CACNA1H* found 12 nonsynonymous SNPs in CAE patients (Chen, Lu et al. 2003). Functional characterisation of 5 of these SNPs in rat channels found increased calcium influx during physiological activation and, for two SNPs, an increased propensity to channel opening (Khosravani, Altier et al. 2004). Thirteen SNPs were introduced into human channels for functional characterisation. Eleven SNPs altered some aspect of channel gating, in at least one case in a manner that would lead to enhanced burst firing of neurons (Vitko, Chen et al. 2005).

### *7.2.7 CACNA1I*

*CACNA1I* encodes the $\alpha$1I subunit which gives rise to a T-type calcium current. It is expressed predominantly in brain, thalamic reticular neurons, subthalamic nucleus, basal ganglia, olfactory bulb, hippocampus, and caudal hypothalamus. Alternative splicing of *CACNA1I* appears to provide a

mechanism for controlling the latency and duration of low-threshold spikes in neurones (Murbartian, Arias et al. 2004).

### *7.2.8 CACNB4*

*CACNB4* encodes the β4 subunit of the VDCC. The β subunit contributes to the function of the calcium channel by increasing peak calcium current, shifting the voltage dependencies of activation and inactivation, modulating G protein inhibition and controlling α-1 subunit membrane targeting(Dolphin 2003). *CACNB4* is expressed predominantly in the brain (cerebellum) and kidney. *Cacnb4* mutations cause the lethargic phenotype in mice, characterised by absence seizures and ataxia (section 3.5.2). *CACNB4* was screened for mutations in small pedigrees with familial epilepsy and ataxia (Escayg, De Waard et al. 2000). A premature-termination mutation R482X was identified in a patient with JME, and a missense mutation identified both in a German family with generalized epilepsy and praxis-induced seizures and in a French Canadian family with episodic ataxia.

### *7.2.9 CACNG1-8*

These genes encode the γ subunits of the VDCC. They may also act as AMPA receptor regulatory proteins. *CACNG4*, *CACNG5* and *CACNG1* occur in a cluster on chromosome 17. *CACNG7*, *CACNG8* and *CACNG6* occur in a cluster on chromosome 19. *CACNG1* is predominantly expressed in muscle, whereas *CACNG2*, *CACNG3* and *CACNG4* are almost exclusively expressed in brain (*CACNG2*: cerebellum, cerebral cortex, hippocampus and thalamus; *CACNG3*: cerebral cortex and amygdala; *CACNG4*: basal ganglia) (Moss, Dolphin et al. 2003). *CACNG5* is expressed in CNS as well as kidney, thymus, prostate and testis. *CACNG6* and *CACNG7* are widely expressed in multiple tissues including brain. *CACNG8* is expressed only in brain. The overall effect of the various γ subunits on VDCC activity appears to be a small downregulation of calcium channel activity (Black 2003). The stargazer mouse, characterised by ataxia, periodic head elevation with

upward gaze, and seizures including behavioural arrest associated with spike-wave discharges, is caused by mutations in *Cacng2* (section 3.5.2). Studies of stargazer have shown that the γ2 subunit interacts with AMPA receptors and regulates their delivery of AMPA receptors to the postsynaptic cell membrane (Chen, Chetkovich et al. 2000). It is likely that other γ subunits are also involved in protein targeting.

### *7.2.10 CACNA2D2*

*CACNA2D2* encodes the α2δ2 subunit of the VDCC. It is expressed in brain as well as lung, testis, heart, and pancreas. The α2δ subunits appear to stimulate functional expression by increasing the number of functional calcium channels at the plasma membrane. Mutations in *Cacna2d2* were found to underlie the ducky mouse phenotype as well as two further mouse mutants du(2J) and Entla (Barclay, Balaguero et al. 2001; Brill, Klocke et al. 2004). These all exhibit generalised spike-wave discharges on EEG associated with behavioural arrest.

## 7.3 $GABA_A$ receptors ($GABA_ARs$)

Markers for ten $GABA_AR$ subunit genes, which occur in three clusters on chromosomes 5q, 4p and 15q, were also tested for linkage. These comprised five α subunits (α1, α2, α4, α5, α6), three β subunits (β1, β2, β3) and two γ subunits (γ2, γ3). Six polymorphic markers were typed in all individuals.

Functional $GABA_ARs$ are formed by the assembly of multiple subunit subtypes (α1–α6, β1–β3, γ1–γ3, δ, ε, π, υ, ανδ ρ1–ρ3) into a pentamer. $GABA_ARs$ control opening of the chloride channel. When GABA binds to the receptor complex, the channel is opened and chloride anions enter the neuron. $GABA_ARs$ are the primary mediators of fast inhibitory synaptic transmission in the central nervous system. They are strongly implicated in the generation of abnormal spike-wave activities in cortical and thalamic

neurons (section 3.5.1). $GABA_A$Rs are associated with binding sites for benzodiazepine and barbiturate AEDs (antiepileptic drugs) in the form of a receptor complex.

$GABA_A$Rs have frequently been shown to play a role in animal models of seizures. Examples include: Injection of various antagonists of the $GABA_A$ receptor induced spike-and-wave discharges in GAERS (Vergnes, Boehrer et al. 2000); Disruption of the *gabrb3* gene, contained within the Angelman Syndrome deletion region, in mice produces EEG abnormalities and seizures as well as behavioural abnormalities (DeLorey, Handforth et al. 1998); Application of the $GABA_A$ receptor antagonist bicuculline significantly increased the magnitude of SWD-related burst firing in thalamocortical neurones in the WAG/Rij strain of rats (Staak and Pape 2001); Selective inhibition of the expression of the $GABA_A$ receptor $\gamma 2$ subunit in the rat hippocampus by means of antisense oligonucleotides leads to spontaneous electrographic seizures (Karle, Woldbye et al. 1998); In seizure-prone (Fast kindling) and seizure-resistant (Slow kindling) rat models, both the mRNA and protein levels of the major $GABA_A$R $\alpha$-subunit expressed in adult brain ($\alpha 1$), as well as those highly expressed during development ($\alpha 2$, $\alpha 3$, and $\alpha 5$), were differentially expressed in both models compared with normal controls(Poulter, Brown et al. 1999).

Mutations in several $GABA_A$R genes have now been identified in human IGEs, including GEFS+ and a phenotype of childhood absence epilepsy and febrile seizures (*GABRG2*, section 2.4.6) and autosomal dominant JME (*GABRA1*, section 2.5.1). A heterozygous variant in *GABRD* (Glu177Ala) is associated with GEFS+, and this variant results in a decreased $GABA_A$ receptor current amplitude (Dibbens, Feng et al. 2004).

## 7.4 $GABA_B$ receptors ($GABA_BRs$)

Seven markers were typed to test the two $GABA_B$ receptor genes, *GABBR1* and *GABBR2* on chromosomes 6p21.3 and 9q22.1 respectively.

The functional $GABA_BR$ is a heterodimer consisting of two seven-transmembrane proteins, *GABBR1* and *GABBR2*. Unlike $GABA_A$ receptors, $GABA_B$ receptors do not form ion channels, but address second messenger systems through the binding and activation of G-proteins. Pre-synaptic $GABA_BRs$ inhibit neurotransmitter release by down-regulating high-voltage activated calcium channels, whereas postsynaptic $GABA_BRs$ decrease neuronal excitability by activating a prominent inwardly rectifying potassium (Kir) conductance that underlies the late inhibitory postsynaptic potentials (Bettler, Kaupmann et al. 2004). Human *GABABR2* and *GABABR1* mRNAs are exclusively expressed in brain and differentially expressed in the amygdala, putamen, and caudate nucleus (Martin, Russek et al. 1999).

*GABBR1*-null mice exhibit spontaneous seizures as well as hyperalgesia, hyperlocomotor activity, and memory impairment (Schuler, Luscher et al. 2001). In lethargic mice, antagonists of the $GABA_B$ receptor suppress seizures, whereas agonists of $GABA_B$ receptors exacerbate them, suggesting enhanced $GABA_B$ receptor-mediated synaptic responses may underlie absence seizures in lethargic mice (Hosford, Clark et al. 1992). In two experimental animal models of generalized absence seizures (one in which the seizures are induced by GHB (Gamma-hydroxybutyric acid) and the other in which the seizures are induced by administration of low-dose pentylenetetrazole), all specific $GABA_B$-receptor antagonists produced blockade of experimental absence seizures; pretreatment with $GABA_B$-receptor agonists resulted in generalized absence status epilepticus lasting for hours (Snead 1996).

## 7.5 *ECA1* locus

The *ECA1* locus on 8q24 was tested using the marker *D8S534* (section 3.5.5).

**Table 7-1** *Candidate genes tested and markers used. A1GUCL is a novel marker (see Appendix 15)*

| Gene Class | Gene | Subunit | Location | Markers | Genomic DNA (kb) | Exonic DNA (kb) | Exons | Marker position (distances in kb) |
|---|---|---|---|---|---|---|---|---|
| Voltage-dependent Calcium Channels | *CACNA1A* | $\alpha$1A | 19p13.3 | *D19S1150* | 299 | 7.8 | 48 | *CACNA1A*(Intron 7: *D19S1150*) |
| | *CACNA1B* | $\alpha$1B | 9q24 | *D9S158* | 244 | 7 | 46 | *D9S158*-1723-*CACNA1B* |
| | *CACNA1D* | $\alpha$1D | 3p14-p21 | *D3S3582, D3S1568, D3S1588, D3S1289* | 317 | 7.6 | 49 | *D3S3582*-5010-*CACNA2D2*(Intron 7:*D3S1568*)-2988-*CACNA1D*-252-*D3S1588*-381-*D3S1289* |
| | *CACNA2D2* | $\alpha 2\delta 2$ | | | 141 | 5.5 | 39 | |
| | *CACNA1E* | $\alpha$1E | 1q31-q32 | *D1S212, D1S2848* | 318 | 9.7 | 48 | *D1S212*-3370-*CACNA1E*-4064-*D1S2848* |
| | *CACNA1G* | $\alpha$1G | 17q22 | *A1GUCL* | 66 | 7.8 | 38 | *CACNA1G*(Intron 2: *A1GUCL*) |
| | *CACNA1H* | $\alpha$1H | 16p12-p13.3 | *D16S521, D16S3024, D16S420* | 68 | 8.1 | 35 | *D16S521*-1109-*CACNA1H*-382-*D16S3024*-22549-*D16S420*-30-CACNG3 |
| | *CACNG3* | $\gamma$3 | | | 107 | 2.7 | 4 | |
| | *CACNA1I* | $\alpha$1I | 22q12-q13 | *D22S283, D22S426, D22S1156, D22S284* | 119 | 10 | 37 | *D22S283*-209-*CACNG2*(Intron 3: *D22S426*)-1283-*D22S1156*-1585-*CACNA1I*-231-*D22S284* |
| | *CACNG2* | $\gamma$2 | | | 139 | 1.5 | 4 | |
| | *CACNB4* | $\beta$4 | 2q23 | *D2S2236, D2S2299* | 261 | 3.2 | 14 | *CACNB4*(Intron 12: *D2S2236*)-681-*D2S2299* |
| | *CACNG5* | $\gamma$5 | 17q24 | *D17S807, D17S1821* | 12 | 1.3 | 4 | *D17S807*-14-*CACNG5*-80-*CACNG4*-11-*CACNG1*-233-*D17S1821* |
| | *CACNG4* | $\gamma$4 | | | 68 | 3.4 | 4 | |
| | *CACNG1* | $\gamma$1 | | | 8 | 0.8 | 5 | |
| | *CACNG7* | $\gamma$7 | 19q13.4 | *D19S572* | 20 | 1.8 | 4 | *D19S572*-390-*CACNG7*-20-*CACNG8*-9-*CACNG6* |
| | *CACNG8* | $\gamma$8 | | | 30 | 1.6 | 5 | |
| | *CACNG6* | $\gamma$6 | | | 20 | 1.4 | 4 | |

**Table 7-1 continued**. *Candidate genes tested and markers used.*

| Gene Class | Gene | Subunit | Location | Markers | Genomic DNA (kb) | Exonic DNA (kb) | Exons | Marker position (distances in kb) |
|---|---|---|---|---|---|---|---|---|
| $GABA_A$ receptors | *GABRA1* | α1 | 5q31.1-q33.2 | *D5S2112, D5S422* | 52 | 3.7 | 11 | *D5S2112*-4038-*GABRB2*-138-*GABRA6*-145-*GABRA1*-168-*GABRG2*-571-*D5S422* |
| | *GABRA6* | α 6 | | | 16 | 1.7 | 9 | |
| | *GABRB2* | β2 | | | 254 | 2 | 11 | |
| | *GABRG2* | γ2 | | | 88 | 3.9 | 10 | |
| | *GABRA2* | α2 | 4p13-p12 | | 140 | 2.4 | 10 | *GABRA2*-529-*GABRA4*-38-*GABRB1*(Intron 8:*GABRB1(GATA)*) |
| | *GABRA4* | α4 | | *GABRB1(GATA)* | 75 | 11 | 9 | |
| | *GABRB1* | β1 | | | 395 | 3.4 | 9 | |
| | *GABRB3* | β3 | 15q11-q13 | *GABRB3(CA), 155CA2, A55CA1* | 227 | 3 | 9 | *GABRB3*(CA)-56-*GABRB3*(Intron 6:*155CA2*)-84-*A55CA1*-8-*GABRA5*-720-*GABRB3* |
| | *GABRA5* | α5 | | | 17 | 0.8 | 6 | |
| | *GABRG3* | γ3 | | | 52 | 0.9 | 5 | |
| $GABA_B$ receptors | *GABBR1* | | 6p21.3 | *D6S265, D6S478, D6S1281* | 31 | 4.4 | 23 | *D6S1281*-4273-*GABBR1*-205-*D6S478*-113-*D6S265* |
| | *GABBR2* | | 9q22.1 | *D9S1816, D9S1809, D9S272, D9S176* | 421 | 5.8 | 19 | *D9S1816*-281-*D9S1809*-2492-*GABBR2*-280-*D9S272*-307-*D9S176* |
| Unknown | *ECA1* | | 8q24 | *D8S534* | | | | *ECA1(D8S534)* |

# *Chapter 8 Linkage analysis*

All linkage programmes were run via the website of the Research and Bioinformatic Divisions, UK Human Genome Mapping Project (HGMP) based in Hinxton, Cambridge, or run directly on a stand alone PC using freely available software.

## 8.1 Files needed for linkage analysis

The linkage analysis software requires several input files. The two files required universally are the pedigree file, which provides data about pedigree structure and members, and the locus data file which provides information about the genetic loci typed in the pedigree. Additional files are required for certain analysis software, and these will be described in the relevant section.

### 8.1.1 Pedigree file

The pedigree file provides information about the structure of each pedigree, the sex and affection status of each individual, and genotyping data for the marker loci tested. The pattern of the pedigree file is shown in Table 8-1.

**Table 8-1** *Structure of linkage pedigree file. The header row is shown for information only and does not form part of the pedigree file. The pedigree must contain both parents or neither. Two marker loci are shown for illustration.*

| **Family ID** | **Subject ID** | **Father** 0 if not in pedigree | **Mother** 0 if not in pedigree | **Gender** 1=male 2=female | **Affection status** 0=unknown 1=unaffected 2=affected | **Alleles at marker locus 1** 0=unknown | | **Alleles at marker locus 2** 0=unknown | |
|---|---|---|---|---|---|---|---|---|---|
| 5 | 101 | 0 | 0 | 2 | 1 | 2 | 3 | 0 | 0 |
| 5 | 102 | 0 | 0 | 1 | 1 | 1 | 10 | 1 | 2 |
| 5 | 202 | 102 | 101 | 2 | 1 | 1 | 2 | 1 | 2 |
| 5 | 204 | 102 | 101 | 2 | 2 | 2 | 10 | 0 | 0 |
| 5 | 207 | 102 | 101 | 1 | 2 | 1 | 2 | 1 | 2 |
| 26 | 101 | 0 | 0 | 1 | 1 | 3 | 4 | 1 | 1 |
| 26 | 102 | 0 | 0 | 2 | 1 | 2 | 4 | 1 | 2 |
| 26 | 202 | 101 | 102 | 1 | 2 | 2 | 3 | 1 | 1 |
| 26 | 204 | 101 | 102 | 2 | 2 | 2 | 3 | 1 | 1 |
| 26 | 206 | 101 | 102 | 1 | 1 | 2 | 4 | 1 | 1 |

### 8.1.2 Locus data file

The locus data file provides information about the loci (disease and marker) included in the pedigree file and the type of analysis to be performed. A typical locus data file contains 3 sections, as follows:

```
(section 1)
3 0 0 5
0 0.0 0.0 0
1 2 3
(sections 2)
1 2
0.995 0.005
1
0.001 0.5 0.5
3 11  # D16S420
0.091 0.091 0.091 0.091 0.091 0.091 0.091 0.091 0.091 0.091 0.09
3 3  # CACNG3EX1CA
0.33 0.33 0.34
(section 3)
0  0
0.0 0.0025
1  0.05  0.4
```

**Section 1.** This contains general information on loci and locus order.

**Line 1** contains 4 parameters: the number of loci to be coded; a risk locus (0=no risk calculations to be performed); sex-linked or autosomal data (0=autosomal); and a program code (5 indicates MLINK is to be used for analysis).

**Line 2** contains 4 parameters: a mutation locus (0=mutation rate zero); male mutation rate; female mutation rate; linkage disequilibrium (0 if loci are assumed to be in linkage equilibrium)

**Line 3** gives the order of the loci along the chromosome

**Section 2.** This section of the locus data file contains detailed information for each locus.

**Line 1** contains 2 parameters: a code to indicate the locus type (1=affection locus, 3=numbered alleles); the number of alleles

The subsequent lines for each locus depend on the locus type.

For an affection locus, the subsequent lines are:

- **Line 2** gives the allele frequency for each allele
- **Line 3** gives the number of liability classes, useful if the disease has an age dependent penetrance
- **Line 4** gives the penetrance values for each genotype. For two alleles, 1 and 2, the genotype order is 11 (penetrance close to 0 if this is the wild type allele), 12 (penetrance ≈1 for a fully penetrant dominant mendelian disorder, ≈ 0 for a recessive disease) and 22 (penetrance ≈1 for a dominant or recessive mendelian disorder)

For numbered alleles, the subsequent line is:

- **Line 2** gives the allele frequency for each allele

**Section 3.** The third section of the locus data file determines what analysis is to be performed.

**Line 1** defines whether there is a sex-difference in recombination fractions (0=no sex difference) and whether an interference parameter is to be used (0=no interference).

**Line 2** specifies the initial recombination fraction(s) between the loci. For n loci, n-1 recombination fractions are required. To measure the LOD score at theta=0 the first value should be 0.

**Line 3** depends on the program being used. For MLINK, the first parameter defines which recombination fractions to vary on subsequent evaluations (1=first recombination fraction only). The second parameter is the value to increment the recombination fraction by each time, and the third parameter is the value at which to stop.

## 8.2 Parameters for linkage analysis

Parametric linkage analysis was performed under the assumption of AD inheritance with 50% and 70% penetrance. These parameters correspond to the available data on mendelian segregation in this trait. They are conservative and the low penetrance reduces the risk of false negative results. A disease allele frequency of 0.01 and a phenocopy rate of 0.0001 were assumed. A disease allele frequency of 1% allows the variant to be classified as a polymorphism which would be consistent with the so-called common disease/common variant hypothesis (although this is a conservative estimate of disease allele frequency). A phenocopy rate of 0.0001 reflects the likely heterogeneous architecture of CAE.

## 8.3 Power of family resource to detect linkage – SLINK software

The power of the family resource to detect linkage was calculated using SLINK (Ott 1989; Weeks, Ott et al. 1990). SLINK randomly simulates multilocus marker data under the condition of linkage or non-linkage and then analyses the results. The simulation algorithm first calculates the conditional probabilities of all the possible multilocus genotypes in a pedigree and randomly assigns one of these genotypes (g1) to individual 1. Taking into account the genotype just generated, the multilocus genotype conditional probabilities are recalculated and a genotype (g2) randomly assigned to individual 2. This process of successive risk calculations continues until multilocus genotypes have been assigned to all individuals in the pedigree. The user defines the number of replicates to be performed.

SLINK calculates the maximum LOD under conditions of homogeneity or heterogeneity and at various recombination fractions. The maximum LOD

score at any recombination fraction is also calculated, as well as the power of the dataset to produce LOD scores of a given magnitude.

SLINK requires 5 input files:

1. Simdata.dat is a standard locus data file which contains information about the allele frequencies of the affection locus and the markers under study.

2. Simped.dat is a standard pedigree file which contains information about the pedigree including each individual's availability code as defined below:

| Code | Trait | Markers |
|---|---|---|
| 0 | Known disease phenotype as defined | Marker genotype unavailable |
| 1 | Use simulated disease phenotype | Simulate |
| 2 | Known disease phenotype as defined | Simulate |
| 3 | Use simulated disease phenotype | Marker genotype unavailable |

3. Slinkin.dat contains information about the way in which the simulation is to be conducted. This includes the number of replicates to be simulated and the proportion of heterogeneity (unlinked families). The format of slinkin.dat is as follows:

57 (a random number seed)
100 (the number of simulated pedigrees to produce)
1 (indicates the affection locus is the first locus in the pedigree and locus data files)
0.0 (the proportion of unlinked families)

4. Limit.dat provides LOD score values for which the simulation program will report how many replicates produce a LOD score exceeding these. The format of limit.dat is a single line as follows:

2 3 4 (the number of replicates with a LOD exceeding 2, 3 and 4 will be calculated)

5. Datafile.dat is a standard locus data file very similar to simdata.dat, which is used by the MSIM program to analyse the output data from SLINK and UNKNOWN. The recombination fraction in the penultimate line is changed to 0.0, so that the starting recombination for calculation of LOD scores is zero.

The simulation is run in three stages:

1. SLINK – this produces the simulated marker data
2. UNKNOWN – this checks through the pedigrees, eliminates impossible genotypes and formats the data for MSIM
3. MSIM – this is a modified version of MLINK (Section 8.5) which calculates lod scores for each simulated data set.

When SLINK generates data under heterogeneity, the LOD scores calculated by MSIM are too low, and the estimates of theta are too high, as MSIM assumes homogeneity ($\alpha$=1). ELODHET is an additional program which analyses the LODFILE.DAT file created by MSIM under heterogeneity.

### 8.3.1 Simulation with thirty-three absence pedigrees (section 5.7.1)

Assuming locus homogeneity and autosomal dominant inheritance with penetrance of 70%, the estimated probability of obtaining a LOD score greater than 3 at $\theta$=0 (as candidate genes are being analysed) was 95.6%. Allowing for heterogeneity with 70% of the families linked, there was a 45.1% probability of obtaining a LOD score greater than 3, and a 73.9% probability of obtaining a LOD score greater than 2.

### 8.3.2 Simulation with fifty-three CAE pedigrees (section 5.7.2)

Assuming locus homogeneity, the estimated probability of obtaining a LOD score greater than 3 at $\theta$=0 (as candidate genes are being analysed) was 99.4%. Allowing for heterogeneity with 70% of the families linked, there was a 59.3% probability of obtaining a LOD score greater than 3, and an 82.0% probability of obtaining a LOD score greater than 2.

## 8.4 Preparation of linkage pedigree file and checking for mendelian errors – GAS software

GAS (Genetic Analysis System, © Alan Young 1993-1998) version 2.3 was used to prepare the linkage pedigree file and to check for mendelian errors in the marker data. GAS requires a modified locus data file (.gas), a pedigree file (.ped) and an allele data file (.siz) (Appendix 16). The GAS program is able to read genotypic data in the allele data file, given in terms of the lengths of CA-repeats, and to process it into a form suitable for linkage analysis. The locus data file includes a command to read allele-size family data and to convert allele size into an allele number. An adaptive binning method is used to determine which alleles are the same and which are different and to label alleles according to three parameters:

| | |
|---|---|
| diffsize | (two alleles differing by more than this are different) |
| samesize | (two alleles differing by less than this are the same) |
| orderfirst | (alleles are labelled in the order they are read) |

An additional parameter, global, ensures the same bins are used for every family in the dataset. This is important if the data is to be used for association analysis.

When GAS cannot decide how to score an ambiguous allele the options are given to put the ambiguous allele into a nearby bin, to create a new bin or to score the allele as unknown.

The pedigree file provides information about the structure of each pedigree, the sex and affection status of each individual, but contains no genotyping data.

GAS checks the allele data for consistency - ie. that the alleles are sensible, distinct and do not imply illegitimacy. An output log file gives a warning if the allele lengths are outside of defined limits, or if a child has a genotype inconsistent with that of their parents and any full siblings. When inconsistencies are encountered, the original electropherogram data must be checked to identify the source of the error so that necessary corrections can be made. The GAS program is then rerun.

The GAS output file is a standard linkage pedigree file including genotyping data for the marker loci tested.

## 8.5 Two-point linkage analysis – LINKAGE software

MLINK version 5.2, part of the LINKAGE package, was used to perform multilocus two point linkage analysis (Lathrop and Lalouel 1984; Lathrop, Lalouel et al. 1984). MLINK calculates two-point linkage between the disease locus and a marker locus. For each pedigree, the logarithm of the likelihood at each value of the recombination fraction, $\theta$, is calculated and the sum of these is the total log likelihood for the pedigree set. The difference between the total log likelihood at a particular value of $\theta$ and the log likelihood at $\theta$=0.5 gives the overall LOD score.

The input files for MLINK are the locus data file (section 8.1.2) and the pedigree file (section 8.1.1). The pedigree file (.ped) was first preprocessed using MAKEPED, which renumbers consecutively the subjects within each pedigree, assigns a ‘proband’ for each pedigree, and reformats the data. The

modified pedigree file (.ppd) was next run through UNKNOWN, which checks through the pedigrees, eliminates impossible genotypes (although these should previously have been identified by GAS) and formats the data for MLINK.

MLINK requires the two output files from UNKNOWN as well as the locus data file, and then calculates the log likelihoods and LOD scores. In order for MLINK to analyse each marker in the locus data file in turn, the linkage control program (LCP) was used. LCP is programmed to automatically copy and rename files, extract data for a subset of loci from the pedigree and locus files and call up MLINK to analyse the data. The linkage reporting program (LRP) was then used to generate a table of LOD scores from the output of LCP. LRP produces the LODs at each value of $\theta$ for each pedigree as well as the total LOD.

## 8.6 Two-point linkage analysis with locus heterogeneity – TABLE software.

LRP simply adds the LODs from different pedigrees in order to obtain an overall LOD at each value of $\theta$. This assumes locus homogeneity. The individual pedigree LODs can also be combined under the assumption that only a proportion ($\alpha$) show linkage, ie locus heterogeneity. The TABLE program can be used to analyse the output from LCP and calculate the total LOD under the assumption of both homogeneity and heterogeneity (HLOD). The HLOD is shown at each value of $\theta$ together with the value of $\alpha$ that maximises it.

## 8.7 Parametric and non-parametric multipoint linkage analysis – GENEHUNTER software

GENEHUNTER 2.1 was used to perform multipoint linkage analysis(Kruglyak, Daly et al. 1996). This allows multiple linked loci to be analysed simultaneously. This compensates for uninformative markers at particular meioses and for missing marker data by extracting data simultaneously from neighbouring markers and allowing construction of marker haplotypes.

The input files for GENEHUNTER are a pedigree file (.ped) which has been generated by GAS, and a modified locus data file (.dat) (Appendix 17) which includes the genetics distance between each marker specified.

GENEHUNTER generates a multipoint parametric LOD score calculated using the transmission model specified in the locus data file. A heterogeneity LOD is also generated across the map together with the value of α that maximises it.

GENEHUNTER also calculates a non-parametric or model-free linkage score (NPL score) together with its associated p-value. The NPL score is a measure of allele sharing identical by descent between either affected sib pairs ($NPL_{pairs}$) or between all affected members in a pedigree ($NPL_{all}$).

A graphical representation of the multipoint LOD score and NPL score across the map is generated in postscript form.

GENEHUNTER also maps the information content, a measure of the proportion of the total inheritance information extracted by the available marker data, as well as constructing the most likely marker haplotypes.

## 8.8 Parametric and non-parametric multipoint linkage analysis with two trait loci – GENEHUNTER-TWOLOCUS software

In the initial analysis of candidate genes, two loci were consistent with linkage to the disease locus. Absence epilepsy is likely to be a multigenic trait, and it was therefore decided to perform a two-trait-locus study by testing both loci simultaneously with GENEHUNTER-TWOLOCUS (Strauch, Fimmers et al. 2000). This uses two unlinked marker maps. The genetic positions of both trait loci are varied on their respective marker maps, as in GENEHUNTER. The inheritance vector approach is used to determine how many alleles a set of individuals share identical by descent. A parametric (LOD) and nonparametric (NPL) score is generated for the two-locus trait model.

GENEHUNTER-TWOLOCUS first loads three input files, two standard locus data files (one for each of the marker maps) and a third file specifying two-disease-locus-model parameters for LOD score analysis. The third file contains the allele frequencies for each of the disease loci, the number of liability classes and two-locus, parent-of-origin specific penetrances (Appendix 18). For each liability class, a four-by-four matrix with two-locus, parent-of-origin specific penetrances is specified. The rows of the matrix represent the genotype at disease locus 1, in the order w/w, m/w, w/m, m/m, where w and m represent the wild-type and mutant alleles respectively and the paternally inherited allele is given first. The columns of the matrix represent the genotype at disease locus 2 ordered in the same way. The penetrance matrices used are illustrated in Figure 8-1. Three parametric models were used. The multiplicative model tests the hypothesis that only a simultaneous dominant mutant allele at the trait loci on both chromosome 15 and 16 leads to the disease. The heterogeneity model assumes that a dominant mutant allele at either trait loci is sufficient to cause the disease. The additive model combines the multiplicative and heterogeneity models, so that a dominant mutant allele at either trait loci can cause the disease, but

the likelihood of disease is significantly increased if mutant alleles are present on both chromosomes 15 and 16.

GENEHUNTER-TWOLOCUS then scans two standard pedigree files, containing the marker genotypes for the first and second marker map respectively. The scan command also requires a parameter specifying the position of the first disease locus on the first marker map which is held fixed during the analysis.

The sequence of commands to run GENEHUNTER-TWOLOCUS is as follows:

```
load chr16.dat chr15.dat model.dat
skip large off                (include large pedigrees)
haplo on                      (construct haplotypes)
photo chr1516.out             (save session to file)
scan chr16.ped chr15.ped 0.23 (fix first disease locus at
                              position 0.23)
total stat het                (combine all pedigrees with
                              heterogeneity)
```

**Figure 8-1** *Penetrance matrices used in GENEHUNTER-TWOLOCUS. The paternally inherited allele is always given first.*

*a) Multiplicative model. The mutant allele at both disease loci produces a dominant effect, as long as at least one mutant allele is present at each locus.*

| | | **Disease locus 2** | | | |
|---|---|---|---|---|---|
| | | w/w | m/w | w/m | m/m |
| **Disease locus 1** | w/w | 0.00 | 0.00 | 0.00 | 0.00 |
| | m/w | 0.00 | 0.90 | 0.90 | 0.90 |
| | w/m | 0.00 | 0.90 | 0.90 | 0.90 |
| | m/m | 0.00 | 0.90 | 0.90 | 0.90 |

*b) Heterogeneity model. The mutant allele at either disease locus exerts a dominant effect, so that either is sufficient to cause the disease.*

| | | **Disease locus 2** | | | |
|---|---|---|---|---|---|
| | | w/w | m/w | w/m | m/m |
| **Disease locus 1** | w/w | 0.00 | 0.90 | 0.90 | 0.90 |
| | m/w | 0.90 | 0.90 | 0.90 | 0.90 |
| | w/m | 0.90 | 0.90 | 0.90 | 0.90 |
| | m/m | 0.90 | 0.90 | 0.90 | 0.90 |

*c) Additive model. A dominant mutant allele at either trait loci can cause the disease, but the likelihood of disease is significantly increased if mutant alleles are present at both loci.*

| | | **Disease locus 2** | | | |
|---|---|---|---|---|---|
| | | w/w | m/w | w/m | m/m |
| **Disease locus 1** | w/w | 0.00 | 0.45 | 0.45 | 0.45 |
| | m/w | 0.45 | 0.90 | 0.90 | 0.90 |
| | w/m | 0.45 | 0.90 | 0.90 | 0.90 |
| | m/m | 0.45 | 0.90 | 0.90 | 0.90 |

# *Chapter 9 Association analysis*

All association analysis programmes were run via the website of the Research and Bioinformatic Divisions, UK Human Genome Mapping Project (HGMP) based in Hinxton, Cambridge, or run directly on a stand alone PC.

## 9.1 Transmission disequilibrium test with SSLPs – GENEHUNTER software

Linkage disequilibrium in informative trios within the pedigrees used for linkage was tested using the TDT command in GENEHUNTER (PART 11.1). Only transmissions from heterozygote parents are counted, including the cases where one parent is missing but the included parent and offspring are both distinct heterozygotes. The TDT was performed on single markers as well as marker pairs, marker triplets and marker quadruplets using the TDT2, TDT3 and TDT3 commands respectively.

## 9.2 Pedigree disequilibrium test - PDT software

The PDT is an extension of the TDT to use all potentially informative data in extended pedigrees for analysis of linkage disequilibrium (Martin, Monks et al. 2000; Martin, Bass et al. 2001). The PDT is a valid test of linkage and association in general pedigrees even in the presence of population stratification and calculates two alternatives of the test statistic. The AVE PDT weights each individual family's contribution equally, regardless of pedigree size and complexity while the SUM PDT weights families proportional to these factors. The input files for PDT are a standard locus data file and a pedigree file which has been processed through MAKEPED. The output data includes the number of informative families used, the number of informative triads used, the number of discordant sib pairs used,

the test result for each alleles and a global statistic for the marker locus. PDT was performed using both individual markers and SNP haplotypes.

## 9.3 Selection of SNPs

SNPs were chosen at ~10kb intervals encompassing the promoter through to the 3' UTR of *CACNG3* (Figure 9-1). The SNPs have been numbered from one to 23 for ease of reference. One of these is a novel SNP (SNP6) identified via sequencing of a subset of the nuclear pedigrees; the remaining 22 can be found on the NCBI SNP database (Table 9-1). These SNPs were typed in the entire resource.

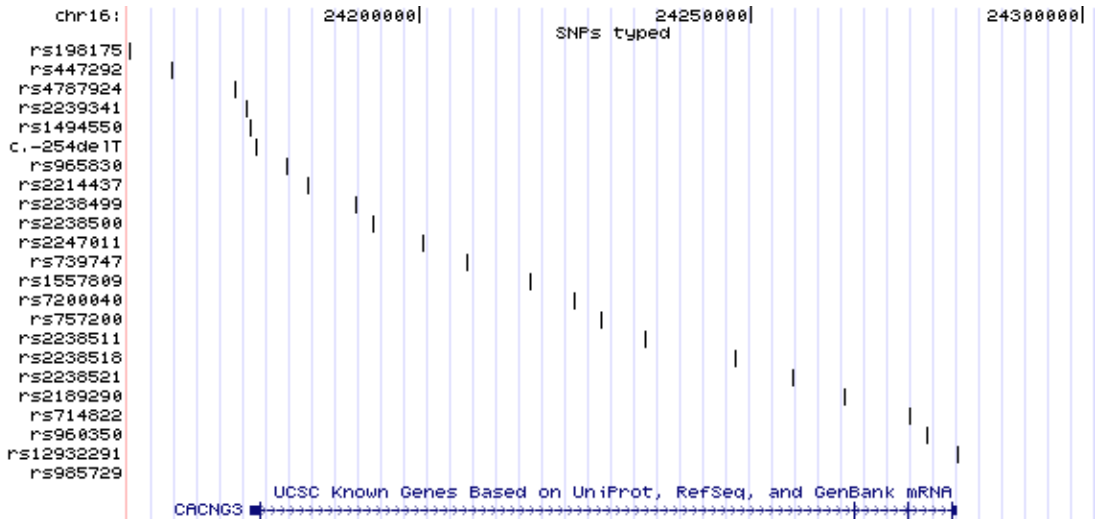



**Figure 9-1** *Position of 23 SNPs type in the entire family resource of AE pedigrees and trios. Image modified from the UCSC Genome Browser (http://genome.cse.ucsc.edu/cgi-bin/hgGateway)*

**Table 9-1** *Details of CACNG3 SNPs used for association analysis*

| SNP NUMBER | refSNP ID | MINOR ALLELE FREQUENCY (allele 2 unless otherwise indicated) | Allele 1 | Allele 2 |
|---|---|---|---|---|
| SNP1 | rs198175 | 0.42[2](allele1) | G | A |
| SNP2 | rs447292 | 0.30[1] | G | A |
| SNP3 | rs4787924 | 0.47[2](allele1) | A | G |
| SNP4 | rs2239341 | 0.21[1] | A | C |
| SNP5 | rs1494550 | 0.26[1] | T | C |
| SNP6 | c-254delT | 0.07[3] | T | delT |
| SNP7 | rs965830 | 0.47[2](allele1) | T | G |
| SNP8 | rs2214437 | 0.50[1] | T | A |
| SNP9 | rs2238499 | 0.20[1] | G | A |
| SNP10 | rs2238500 | 0.46[1](allele1) | G | A |
| SNP11 | rs2247011 | 0.26[1] | A | G |
| SNP12 | rs739747 | 0.27[1](allele1) | A | T |
| SNP13 | rs1557809 | 0.20[1] | T | C |
| SNP14 | rs7200040 | 0.35[1](allele1) | A | G |
| SNP15 | rs757200 | 0.25[1] | C | T |
| SNP16 | rs2238511 | 0.34[1] | T | C |
| SNP17 | rs2238518 | 0.46[1] | T | C |
| SNP18 | rs2238521 | 0.26[1] | G | A |
| SNP19 | rs2189290 | 0.36[1] | T | C |
| SNP20 | rs714822 | 0.14[1] | C | G |
| SNP21 | rs960350 | 0.38[1] | T | C |
| SNP22 | rs12932291 | 0.10[2] | C | G |
| SNP23 | rs985729 | 0.48[1](allele1) | T | C |

[1] = data derived from NCBI SNP database; [2] = data derived from KBioscience validation; [3] = data derived from study population

## 9.4 Linkage disequilibrium map construction, identification of haplotype block structure– HAPLOVIEW software

Haploview was used to construct an LD map across regions of interest using the SNPs typed in that region (Barrett, Fry et al. 2005). Haploview generates a graphical representation of LD using a colour scheme to represent the strength of the LD. Blocks were defined as a solid spine of LD ie the first and last marker in a block are in strong LD with all intermediate markers (allowing for one slight mismatch). The intermediate markers are not necessarily in strong LD with one another. Haplotype blocks were defined using a value of D' = 0.7 as the criteria for strong LD.

The PDT was then performed using both individual SNPs and the SNP haplotypes identified. Each haplotype was treated as a single allele at a multi-allelic locus, as PDT cannot simultaneously analyse multiple loci.

Data was input into Haploview in standard LINKAGE pedigree file format.

## 9.5 Association analysis using tagging SNPs.

Haploview was also used to select a set of tagging SNPs. Two strategies were used. The first strategy selects tagging SNPs which will define all haplotypes above a certain (adjustable) frequency threshold. A frequency of 1% was used. The second strategy uses aggressive tagging based on Tagger (de Bakker, Yelensky et al. 2005). This combines a pairwise and a multimarker approach. Initially a minimal set of markers is selected such that all alleles to be captured are correlated at an $r^2$ greater than an adjustable threshold with a marker in that set. An $r^2$ threshold of 0.8 was used. Multi-marker tests constructed from the set of markers chosen as pairwise tags are then used to try to capture SNPs which could not be captured in the pairwise step. Finally the tag list is minimised where possible by replacing certain tags with multimarker tests. Multimarker tests are constructed only from SNPs

which are in strong LD with each other, as measured by a pairwise LOD score. This avoids overfitting. The default cutoff of LOD=3.0 was used.

The PDT result obtained with the tagging SNPs was compared with the haplotype PDT.

# *Chapter 10 Sequencing of CACNG3*

Exonic DNA and surrounding intronic and promoter sequence DNA of *CACNG3* was sequenced in 73 individuals (Figure 10-1). Cases were predominantly chosen from families which demonstrated linkage to the *CACNG3* locus. These consisted of 46 individuals with CAE selected from 43 of the families used for CAE linkage (including 3 sib pairs); 25 individuals (7 CAE, 5 JAE and 13 AE) selected from 23 of the additional families included in the AE linkage (including 2 sib pairs); 2 individuals with CAE selected from 2 of the trios included in the SNP based association analysis (Table 10-1). A full list of the individuals selected for sequencing is given in Appendix 19.

Following the SNP based association analysis, further sequencing of approximately 34kb of genomic DNA (chromosome 16, NCBI Build 36.1 position 24155960- 24190949) encompassing SNPs 1-9 from 24 unrelated affected individuals was performed. Cases were chosen whose haplotypes demonstrated the most significant disease association.

Identified variants were assessed for any potential functional affect by searching for predicted regulatory motifs contained within the TransFac and Biobases databases via the Softberry NSITE portal. The FPROM program, also via the Softberry portal, was used to predict the position of potential promoters and enhancers. FPROM identifies characteristics of functional motifs and oligonucleotide compositions of potential transcription start positions (Solovyev and Shahmuradov 2003). The GeneSplicer program was used to predict whether any variant might affect the splicing of the gene by identifying exon-intron boundaries and scoring them (Pertea, Lin et al. 2001). The program ESEfinder was used to predict the presence of any exonic splicing enhancers (ESEs) in exon 1 (Cartegni, Wang et al. 2003) Exonic enhancers serve as binding sites for specific serine/arginine-rich (SR) proteins. SR proteins bound to ESEs can promote exon definition by directly recruiting the splicing machinery and by antagonizing the action of nearby silencer elements. SR proteins are characterised by RNA-recognition motifs (RRM) and a distinctive C-terminal domain highly enriched in RS dipeptides (the RS domain). Substrate specificity is determined by the RRMs, which

mediate sequence-specific binding to the RNA, whereas the RS domain appears to be involved mainly in protein-protein interactions. ESEfinder identifies putative binding sites for four SR-rich proteins.

Variants were also checked for sequence conservation between humans, Pan troglodytes (chimpanzees), Mus musculus (house mouse) and Rattus Norvegicus (brown rat) using standard BLAST analysis via the ENSEMBL website.

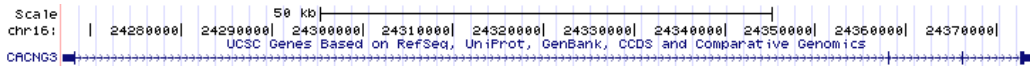



**Figure 10-1** *Genomic structure of CACNG3. Image modified from the UCSC Genome Browser (http://genome.cse.ucsc.edu/cgi-bin/hgGateway).*

**Table 10-1** *Summary of individuals selected for CACNG3 sequencing.*

| | | Families: | | |
|---|---|---|---|---|
| | | 43 CAE linkage families (3 sib pairs) | 23 AE linkage families (2 sib pairs) | 2 trios |
| Individuals: | CAE | 46 | 7 | 2 |
| | JAE | | 5 | |
| | AE | | 13 | |

# PART 3 Results

# *Chapter 11 Initial investigation of candidate gene loci in absence pedigrees*

The initial cohort of 33 families (Figure 5-3) was used to test 30 candidate gene loci, represented by markers localising to seventeen VDCC subunit genes, ten $GABA_AR$ subunit genes, two $GABA_B$ receptor genes and the *ECA1* locus on 8q24, and analysed by by linkage and association analysis.

Two-point LOD scores are presented for genes tested by a single SSR polymorphism. Loci at which LOD scores were positive were tested for allelic association using the Pedigree Disequilibrium Test (PDT). Where genes were tested by two or more markers, the multipoint LOD scores are presented at the recombination fractions corresponding to the known gene locations.

## 11.1 Voltage-dependent calcium channel (VDCC) subunit genes

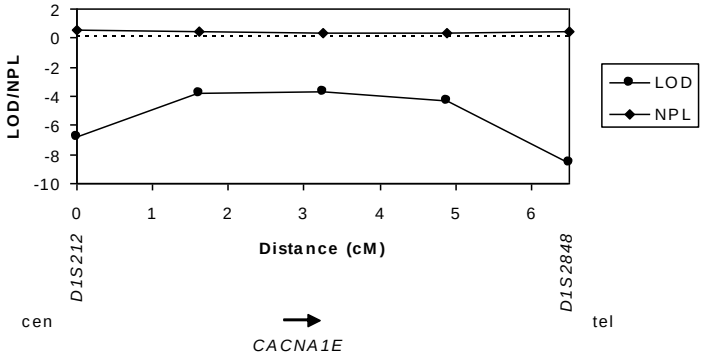
Six loci intragenic or in close proximity to eight VDCC subunit genes were tested using single SSLPs. The two-point LOD scores assuming locus homogeneity at various recombination fractions, $\theta$, are presented in Table 11-1**.** The three intragenic markers, *D19S1150* (*CACNA1A*), *D3S1568* (*CACNA2D2*), *A1GUCL* (*CACNA1G*), all gave exclusionary LOD scores at $\theta$=0. The HLODs at these loci were zero, and inspection of the segregation of alleles showed, as expected, only a small minority of pedigrees consistent with linkage. The SSRs selected to test the VDCC subunits *CACNA1B* (*D9S158*) and the *CACNG6*, *CACNG7*, *CACNG8* cluster (*D19S572*) both map to within 5cM of the genes under investigation. Both markers gave exclusionary LOD scores up to a recombination fraction of $\theta$=0.1. The LOD score at *D16S420*, which lies 31 kb telomeric to *CACNG3*, was -3.34 at $\theta$=0. The HLOD however was 0.55 ($\alpha$=0.35) and the NPL score 1.21. The PDT showed no evidence of allelic association at *D16S420*.

Six SSLP loci encompassing ten VDCC subunits were tested using between two and four SSLP polymorphisms to encompass each locus. The four markers on chromosome 3p14-p21 (*D3S3582, D3S1568, D3S1588, D3S1289)* gave a LOD score of –11.5 at the value of $\theta$ corresponding to the location of *CACNA1D* (Figure 11-1(c)). Four markers on chromosome 22q12-q13 (*D22S1156, D22S284, D22S283, D22S426)* gave a LOD score of –4.54 at the $\theta$ corresponding to *CACNG2*, and a LOD score of –6.54 at the $\theta$ corresponding to *CACNA1I* (Figure 11-2(c)). *CACNB4* was tested using two markers on chromosome 2q23 (*D2S2236, D2S2299)* giving a LOD score of –6.9 at the relevant recombination fraction (Figure 11-1(b)). The HLODs at each of these loci were zero. The cluster of three VDCC $\gamma$ subunits (*CACNG5, CACNG4, CACNG1*) on chromosome 17 is encompassed by two markers (*D17S807, D17S1821*) (Figure 11-2(b)). The maximum (least negative) LOD score between these two markers was –6.9, with an HLOD of zero. *CACNA1E* was tested using two markers on chromosome 1q31-q32 (*D1S212, D1S2848)* which gave a LOD score of –3.6 at the approximate location of *CACNA1E* (Figure 11-1(a)). The two markers on chromosome 16p12-p13.3 (*D16S521, D16S3024)* gave a LOD score of –6.7 at the $\theta$ corresponding to *CACNA1H* (Figure 11-2(a)).

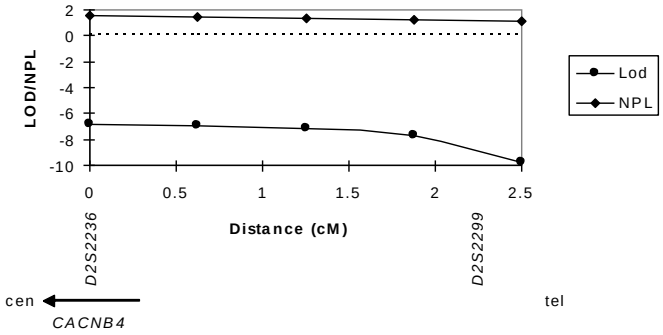
**Table 11-1** *Two-point LOD scores for 33 CAE pedigrees at six single SSLP loci to which eight voltage-dependent calcium channel subunit genes map. D19S1150, A1GUCL and D3S1568 are intragenic markers.*

| | | | LOD score at recombination fraction, θ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Gene | Location | Marker | 0.00 | 0.01 | 0.05 | 0.10 | 0.20 | 0.30 | 0.40 |
| *CACNA1A* | 19p13.3 | *D19S1150* | -15.16 | -11.86 | -7.22 | -4.63 | -2.02 | -0.77 | -0.18 |
| *CACNA1B* | 9q24 | *D9S158* | -13.09 | -9.75 | -5.53 | -3.38 | -1.39 | -0.55 | -0.18 |
| *CACNA1G* | 17q22 | *A1GUCL* | -6.48 | -4.81 | -2.70 | -1.65 | -0.68 | -0.25 | -0.05 |
| *CACNA2D2* | 3p14-21 | *D3S1568* | -7.60 | -5.47 | -2.60 | -1.19 | -0.10 | 0.16 | 0.11 |
| *CACNG3* | 16p12-p13.1 | *D16S420* | -3.34 | -2.06 | -0.40 | 0.34 | 0.67 | 0.42 | 0.09 |
| *CACNG6* | | | | | | | | | |
| *CACNG7* | 19q13.4 | *D19S572* | -11.30 | -8.48 | -4.95 | -3.10 | -1.28 | -0.45 | -0.10 |
| *CACNG8* | | | | | | | | | |

**a) Chromsome 1q31-32: *CACNA1E***




**b) Chromosome 2q23: *CACNB4***




**c) Chromosome 3p14-21: *CACNA1D*, *CACNA2D2***

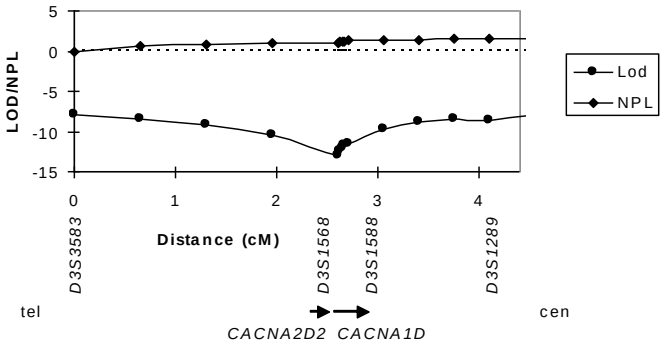
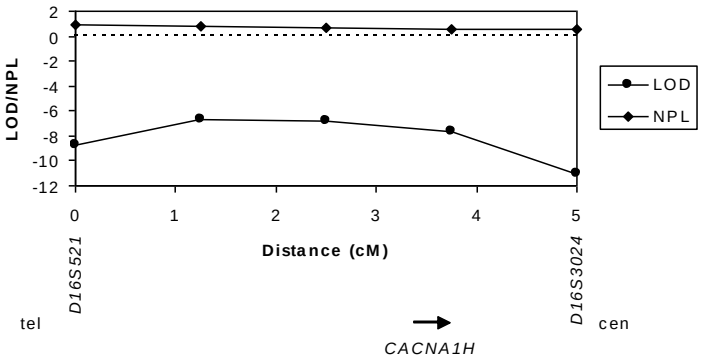
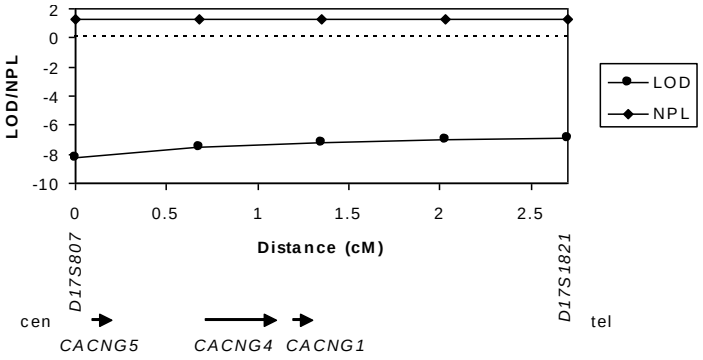



**Figure 11-1** *Thirty-three absence pedigrees: multipoint parametric and non-parametric linkage analysis at VDCC loci CACNA1E, CACNB4, CACNA1D, CACNA2D2*

**a) Chromosome 16p13.3: *CACNA1H***




**b) Chromosome 17q14: *CACNG1, CACNG4, CACNG5***




**c) Chromosome 22q12-13: *CACNA1I, CACNG2***

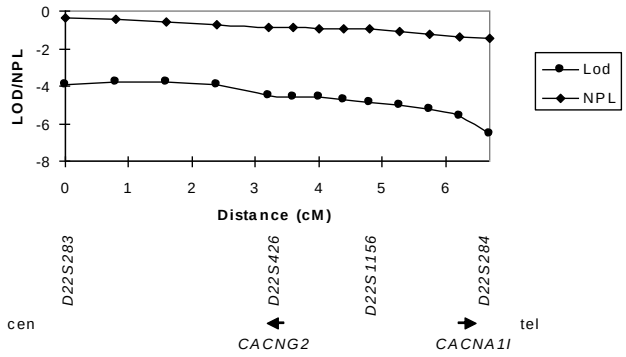



**Figure 11-2** *Thirty-three absence pedigrees: multipoint parametric and non-parametric linkage analysis at VDCC loci CACNA1H, CACNG1, CACNG4, CACNG5, CACNA1I, CACNG2*

## 11.2 $GABA_AR$ subunit genes

The cluster of four $GABA_AR$ subunits (*GABRB2, GABRA6, GABRA1, GABRG2*) on chromosome 5q31.1-q33.2, were tested using two markers (*D5S2112, D5S422)* spanning the genes. The maximum LOD score across this region was –11.8, thus excluding all four genes (Figure 11-3). *GABRB1* on chromosome 4p13-p12 was tested using an intragenic tetranucleotide repeat polymorphism (*GABRB1(GATA)*). The LOD score at $\theta$=0 was –11.6. *GABRA4* and *GABRA2* lie telomeric to *GABRB1*, at physical distances of 190 kb and 607 kb respectively. The LOD score at $\theta$=0.5 from *GABRB1* was –4.8, thus also excluding *GABRA4* and *GABRA* (Table 11-2). Three SSR loci (*GABRB3(CA), 155CA2, A55CA1)* were used to test the cluster of $GABA_AR$ subunits (*GABRA5, GABRB3, GABRG3*) on chromosome 15q11-q13. Although the maximum LOD across this region was –5.6, the HLOD was 0.69 at the *GABRB3* intragenic marker (*155CA2*) with a corresponding alpha of 0.33. Similarly, the HLOD was 0.78 at *A55CA1* (adjacent to the 5' end of *GABRA5*) with a corresponding alpha of 0.35 (Figure 11-4). The corresponding NPL scores at these two markers were both 2.49 (p=0.005). *GABRG3* physically maps just outside the region encompassed by the three SSR loci, approximately 23 kb telomeric to *GABRA5*. Examination of the haplotypes showed that 20 of the 33 families were *consistent* with linkage at the *A55CA1* locus. Thus a proportion of families could not be excluded from linkage to *GABRA5*, *GABRB3* or *GABRG3*. The PDT did not demonstrate allelic association at these loci.

**Table 11-2** *Two-point LOD scores for 33 CAE pedigrees at GABRB1(GATA) at which three $GABA_AR$ subunit genes map. GABRB1(GATA) is intragenic to GABRB1*

| | | | LOD score at recombination fraction, θ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Gene | Location | Marker | 0.00 | 0.01 | 0.05 | 0.10 | 0.20 | 0.30 | 0.40 |
| *GABRB1* | | | | | | | | | |
| *GABRA4* | 4p13-12 | *GABRB1(GATA)* | -11.56 | -8.55 | -4.77 | -2.91 | -1.20 | -0.44 | -0.11 |
| *GABRA2* | | | | | | | | | |

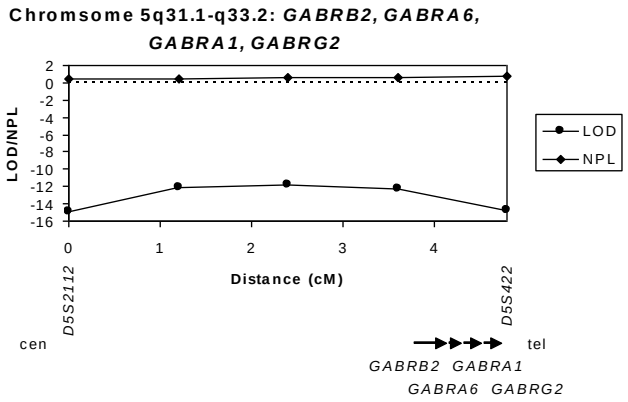



**Figure 11-3** *Thirty-three absence pedigrees: multipoint parametric and non-parametric linkage analysis at chromosome region 5q31.1-33.2 encompassing GABRB2, GABRA6, GABRA1, GABRG2*

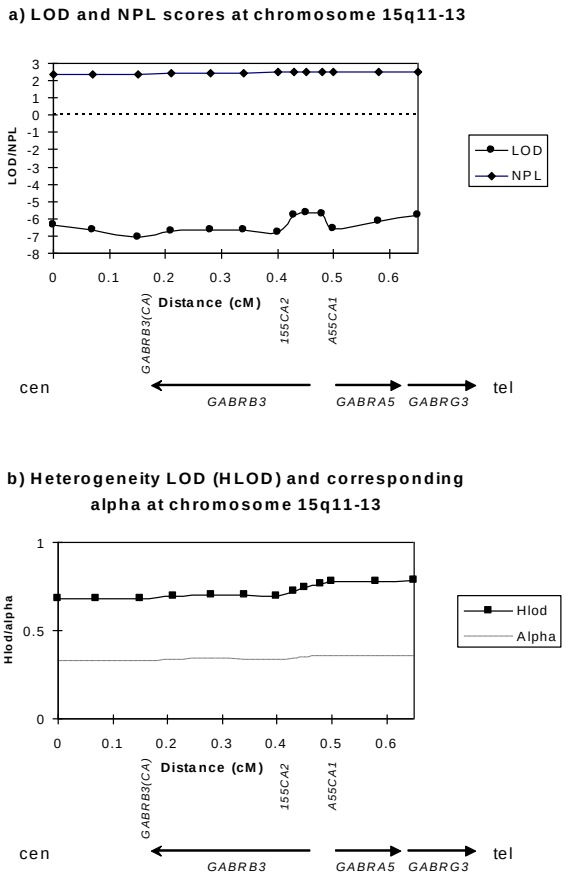



**Figure 11-4** *Thirty-three absence pedigrees: multipoint parametric and non-parametric linkage analysis at chromosome region 15q11-13 encompassing GABRB3, GABRA5, GABRG3*

## 11.3 $GABA_BR$ subunit genes

The $GABA_B$ receptor gene *GABBR1* on chromosome 6p21.3 was tested using three markers *(D6S625, D6S478, D6S1281)* spanning approximately 3cM. *GABBR1* lies between *D6S478* and *D6S1281*. The LOD score at the $\theta$ corresponding to *GABBR1* was -18.16 with an HLOD of zero (Figure 11-5(a)). *GABBR2* on chromosome 9q22.1 was tested using four markers (*D9S1816, D9S1809, D9S272, D9S176). GABBR2* lies between *D9S1809* and *D9S272*, and the LOD score at this location was approximately -10 with an HLOD of zero (Figure 11-5**(**b)).

## 11.4 *ECA1* locus

The *ECA1* locus on 8q24 was tested using the marker *D8S534*. The LOD score was –17.03 and HLOD equal to zero.

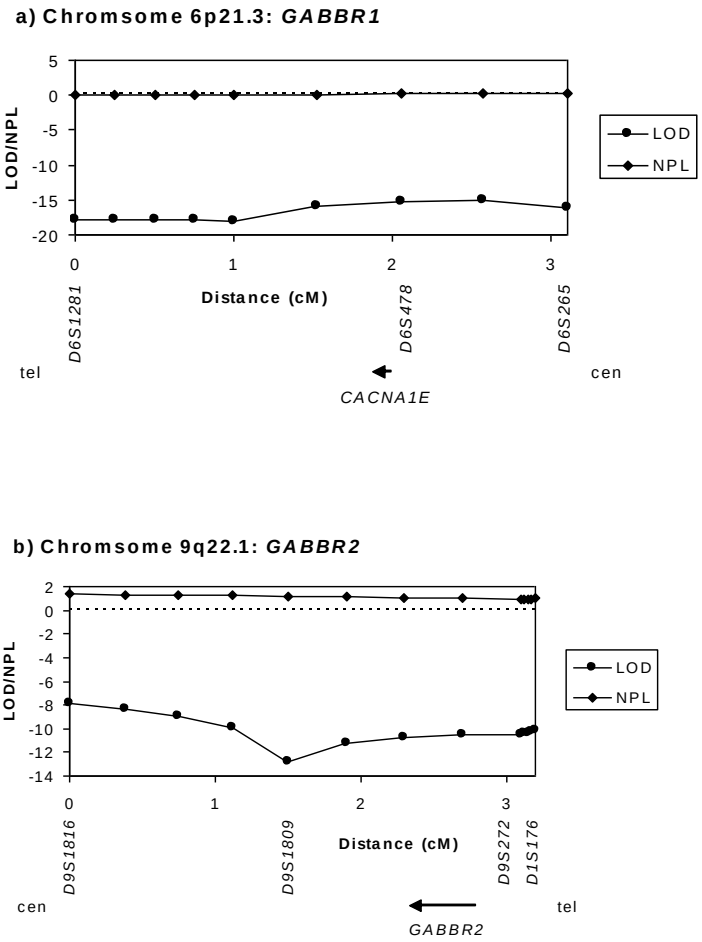



**Figure 11-5** *Thirty-three absence pedigrees: multipoint parametric and non-parametric linkage analysis at chromosome regions encompassing GABBR1 and GABBR2*

# *Chapter 12 Further investigation of candidate gene loci implicated by initial linkage analysis*

Following the preliminary linkage analysis, candidate gene loci showing positive linkage were tested in the expanded collection of 53 CAE families, 82 absence families and 217 parent-child trios. Maximum HLODs are summarised in Table 12-1. Maximum NPL scores are summarised in Table 12-2.

## 12.1 Voltage-dependent calcium channel gene: *CACNG3*

*CACNG3* showed evidence of possible linkage with a positive HLOD in the initial cohort of 33 absence pedigrees. *CACNG3* was therefore tested by linkage in the larger cohorts of CAE and AE pedigrees using three SSLP markers spanning the gene, *D16S420* and two novel SSLPs, *CACNG3EX1CA* and *UCL10321* (Figure 12-1, Appendix 15). *D16S420* lies 31kb 5' (telomeric) to *CACNG3*. *CACNG3EX1CA* is a *CA* repeat polymorphism which lies 73bp 5' to *CACNG3* exon 1. *UCL10321* is a *CA* repeat 135kb 3' (centromeric) to *CACNG3.*

### 12.1.1 Fifty-three CAE families: linkage analysis of *CACNG3*

Parametric linkage analysis was performed in the 53 CAE families using GENEHUNTER assuming AD inheritance with penetrance of 0.7 and 0.5. Criteria for affectedness was 'childhood absence epilepsy' as defined in section 5.3.1. The LOD scores remained negative in this larger cohort of CAE families: maximum LOD= -2.3 with penetrance=0.7 and LOD= -1.6 with penetrance=0.5 (Figure 12-2(a)). However, the HLOD increased to 3.22 (alpha=0.58) with penetrance of 0.7, and 3.03 (alpha=0.61) with penetrance of 0.5 (Figure 12-2(b,c)). HLOD was maximal between *CACNG3EX1CA* and *UCL10321*. The NPL score also increased to a maximum of 2.59 (p=0.003).

The transmission disequilibrium test performed on each SSLP in turn produced a significant association with the disease phenotype for allele 1 of *CACNG3EX1CA*, which was transmitted from a heterozygous parent to an affected offspring 45 times, and untransmitted 21 times ($\chi^2$=8.73, p=0.003) (Appendix 20a). Allele 2 of *CACNG3EX1CA* was associated with a protective effect (transmitted 22, untransmitted 44, $\chi^2$=7.33, p=0.007). Two alleles of *UCL10321* also showed borderline significant associations, one positive association with the disease phenotype (transmitted 9, untransmitted 2, $\chi^2$=4.45, p=0.035), and one negative association with the disease phenotype (transmitted 27, untransmitted 49, $\chi^2$=6.37, p=0.012).

The PDT also produced a significant association with the disease phenotype for allele 1 of *CACNG3EX1CA* (p=0.005) and a negative association for allele 2 (p=0.01) (Appendix 22a). One allele of UCL10321 also showed a negative association (p=0.005).

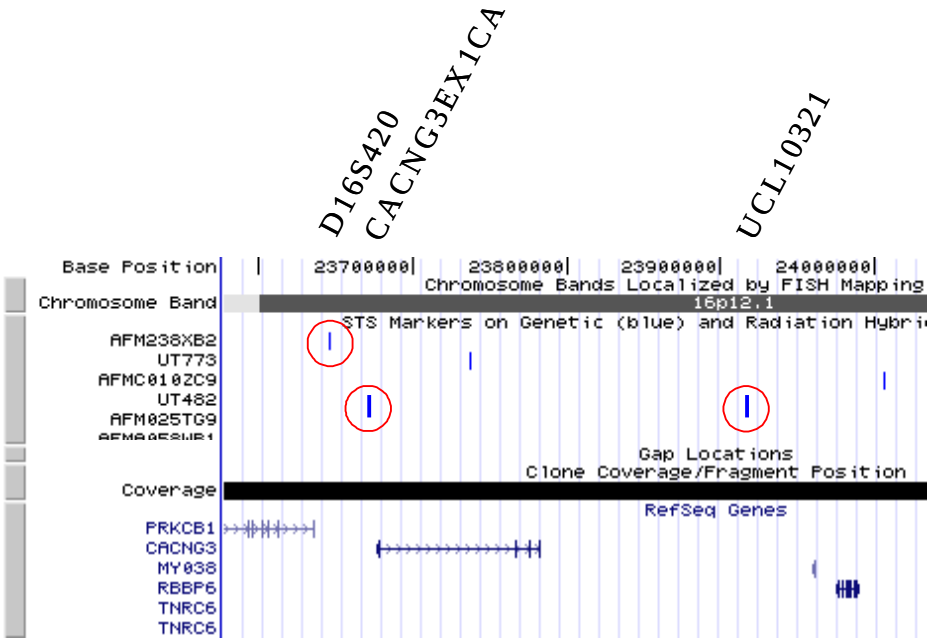



**Figure 12-1** *Position of three SSLPs, D16S420, CACNG3EX1CA and UCL10321 relative to the VDCC gene, CACNG3. Image modified from the UCSC Genome Browser (http://genome.cse.ucsc.edu/cgi-bin/hgGateway)*

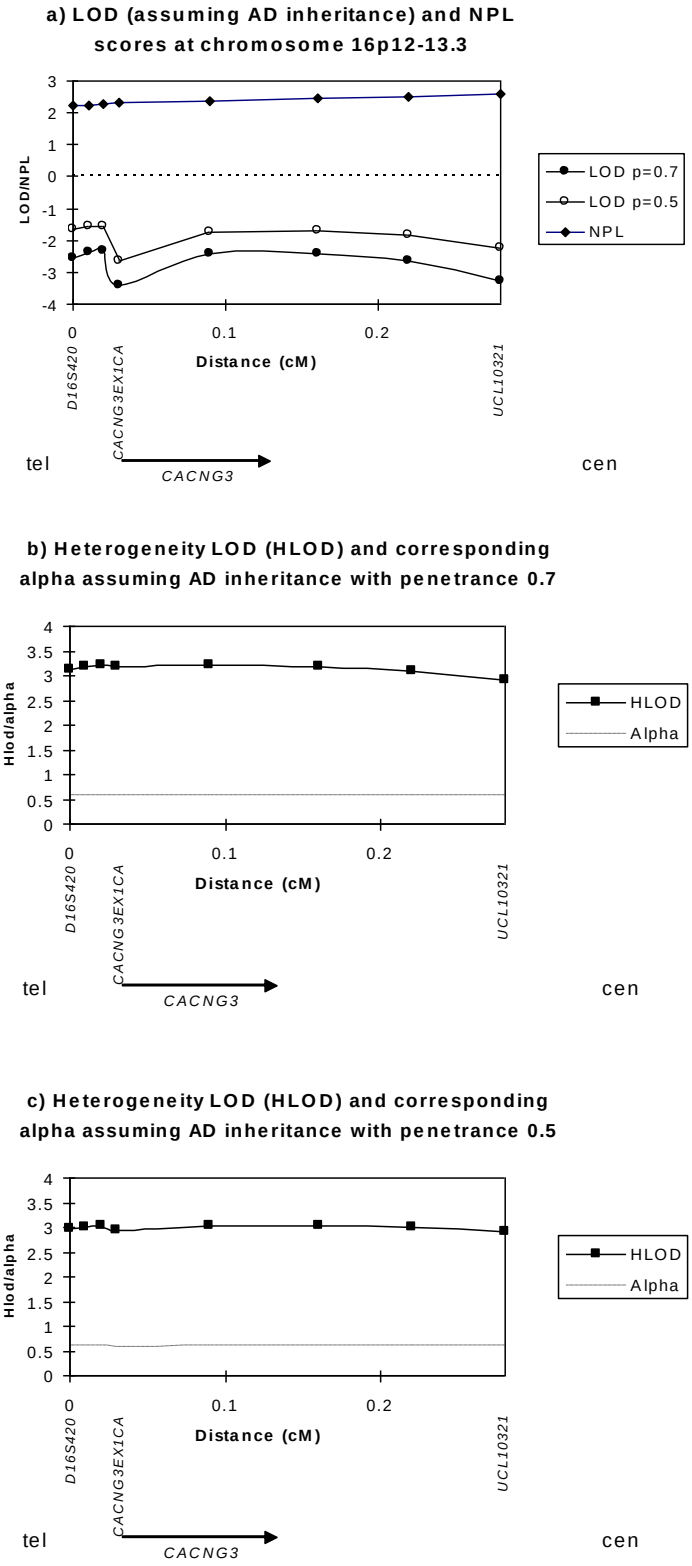



**Figure 12-2** *Fifty-three CAE pedigrees: Multipoint parametric (AD inheritance, penetrance 0.7 and 0.5) and non-parametric linkage analysis at CACNG3 locus on chromosome 16p12-13.3*

### 12.1.2 Eighty-two absence families: linkage analysis of *CACNG3*

*CACNG3* was then tested in the eighty-two absence families by linkage analysis with criteria for affectedness CAE, JAE or AE as defined in section 1.1. Parametric linkage analysis was performed using GENEHUNTER assuming AD inheritance with penetrance of 0.7 and 0.5. The homogeneity LODs were negative with penetrances of 0.7 and 0.5 (maximum LODs -5.6 and -4.5 respectively) (Figure 12-3(a)). The maximum HLOD with penetrance 0.7 was 3.44 (alpha=0.52) (Figure 12-3(b)). The maximum HLOD with penetrance 0.5 was 3.36 (alpha=0.55) (Figure 12-3(c)). HLOD was maximal between markers *CACNG3EX1CA* and *UCL10321*. The maximum NPL score was 2.74 (p=0.002) at marker *UCL10321*.

The transmission disequilibrium test performed on each SSLP in turn again produced a significant result for allele 1 of *CACNG3EX1CA*, which was transmitted from a heterozygous parent to an affected offspring 63 times, and untransmitted 36 times ($\chi^2$=7.36, p=0.007) (Appendix 20b). Allele 2 of *CACNG3EX1CA* was associated with a weakly protective effect (transmitted 38, untransmitted 63, $\chi^2$=6.19, p=0.013).

The PDT also showed a positive association with the disease phenotype for allele 1 of *CACNG3EX1CA* (p=0.009) and a negative association for allele 2 (p=0.017) (Appendix 21b).

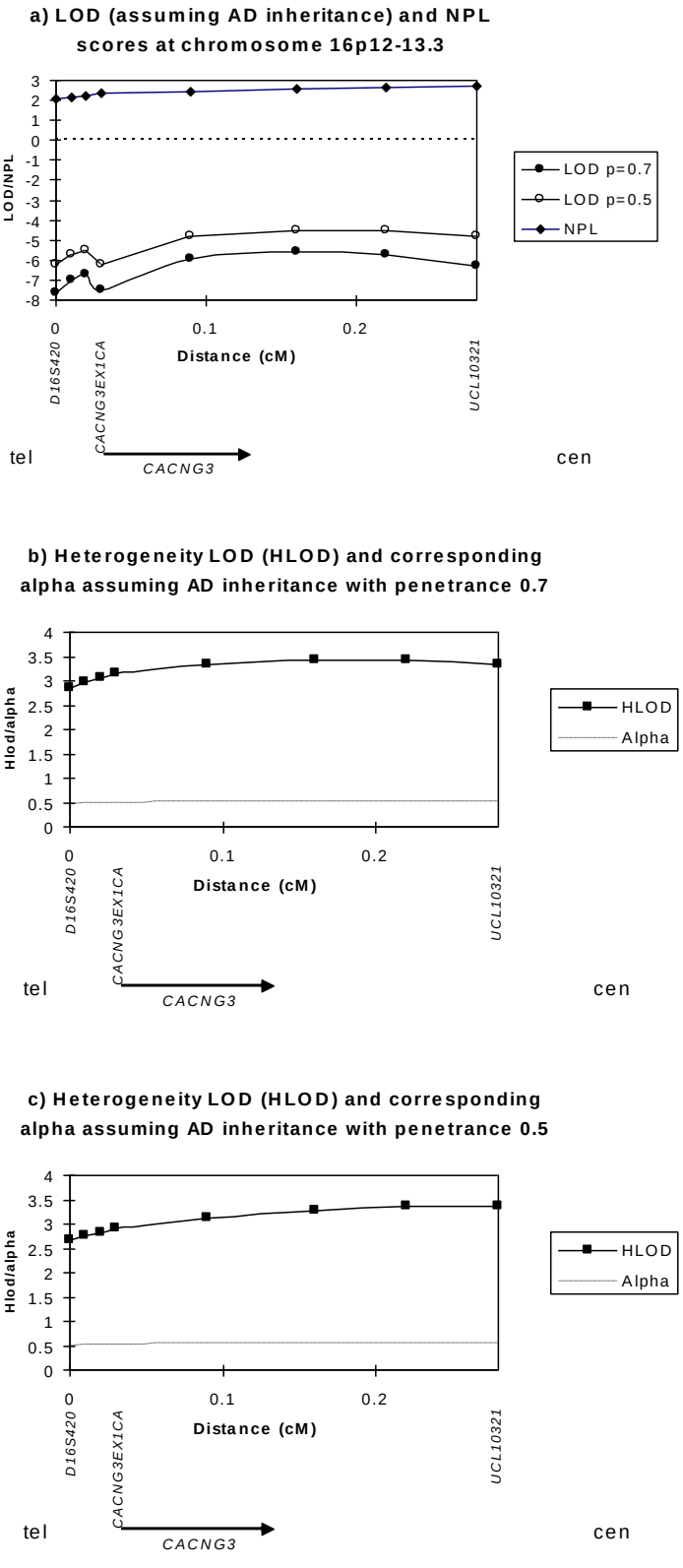



**Figure 12-3** *Eighty-two Absence pedigrees: Multipoint parametric (AD inheritance, penetrance 0.7 and 0.5) and non-parametric linkage analysis at CACNG3 locus on chromosome 16p12-13.3*

## 12.2 $GABA_AR$ subunit gene cluster on chromosome 15q11-13

The $GABA_AR$ subunits *GABRBB3, GABRA5, GABRG3* also showed evidence of weak linkage with a positive HLOD in the initial cohort of 33 absence pedigrees. This cluster of $GABA_AR$ subunits on chromosome 15q was therefore tested by linkage in the larger cohorts of CAE and AE pedigrees. Four SSLP markers were used; *GABRB3(CA), 155CA2, A55CA1* as used in the preliminary linkage (Table 7-1), and an additional marker *D15S156* which encompasses the telomeric end of the cluster and lies 29kb 3' to *GABRG3* (Figure 12-4).

### *12.2.1* Fifty-three CAE families: linkage analysis of *GABRB3, GABRA5, GABRG3*

Parametric linkage analysis was performed in the 53 CAE families using GENEHUNTER assuming AD inheritance with penetrance of 0.7 and 0.5. Criteria for affectedness was 'childhood absence epilepsy' as defined in section 5.3.1. The LOD scores remained very negative in the larger cohort of CAE families: maximum LOD= -9.4 with penetrance=0.7 and maximum LOD= -6.9 with penetrance=0.5 (Figure 12-5(a)). The HLOD increased to a maximum of 0.97 (alpha=0.32) with penetrance of 0.7, and 1.18 (alpha=0.38) with penetrance of 0.5 (Figure 12-5(b,c). The peak HLOD occurred at marker *155CA2* which is intragenic to *GABRB3*.

The NPL score increased to a maximum of 3.05 (p=0.0006) at marker *D15S156* (Figure 12-5(a)). The discrepancy between the HLOD and NPL scores suggested that the data may be more consistent with a non-dominant mode of inheritance. Parametric linkage analysis was therefore performed assuming AR inheritance with penetrance of 0.7 and 0.5 (Figure 12-6). The LOD scores remained negative, but the HLODs increased to a maximum of 1.34 (alpha=0.34) with penetrance of 0.7, and 1.84 (alpha=0.46) with penetrance of 0.5. The peak HLODs now occurred adjacent to marker *D15S156*.

The transmission disequilibrium test performed on each SSLP in turn identified one allele of *155CA2* which may be associated with a protective effect, as it was only transmitted once from a heterozygous parent to an affected offspring, but untransmitted 10 times ($\chi^2$=7.36, p=0.007) (Appendix 21a). A borderline significant protective effect was also associated with another allele of *155CA2* (transmitted 8, untransmitted 20, $\chi^2$=5.14, p=0.02) and one allele of *A55CA1* (transmitted 19, untransmitted 37, $\chi^2$=5.79, p=0.02). Another allele of *A55CA1* showed a weak association with the disease phenotype (transmitted 14, untransmitted 5, $\chi^2$=4.26, p=0.04).

The PDT also showed a negative association with the disease phenotype for two alleles of *155CA2* (p=0.02 and 0.03) and a negative association for one allele of A55CA1 (p=0.01) (Appendix 23a)

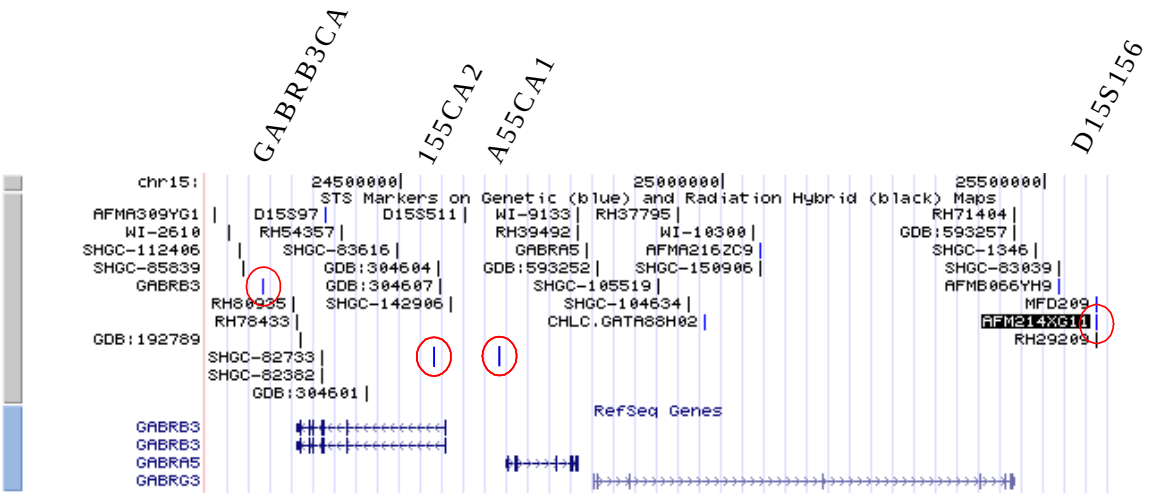



**Figure 12-4** *Position of four SSLPs, GABRB3CA, 155CA2, A55CA1 and D15S156 relative to the GABAAR gene cluster, GABRB3, GABRA5 and GABRG3 on chromosome 15q11-13 . Image modified from the UCSC Genome Browser (**http://genome.cse.ucsc.edu/cgi-bin/hgGateway**)*

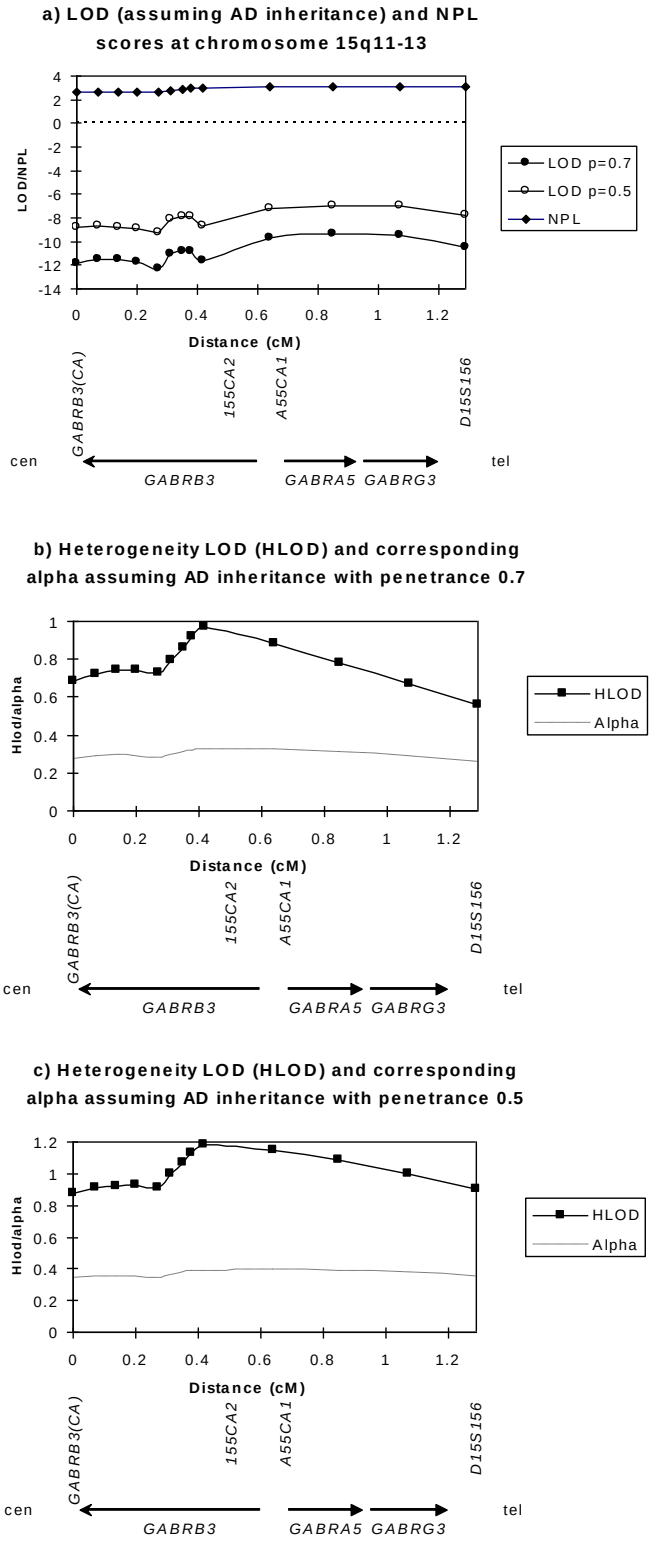



**Figure 12-5** *Fifty-three CAE pedigrees: Multipoint parametric (AD inheritance, penetrance 0.7 and 0.5) and non-parametric linkage analysis at $GABA_AR$ gene cluster, GABRB3, GABRA5 and GABRG3 on chromosome 15q11-13*

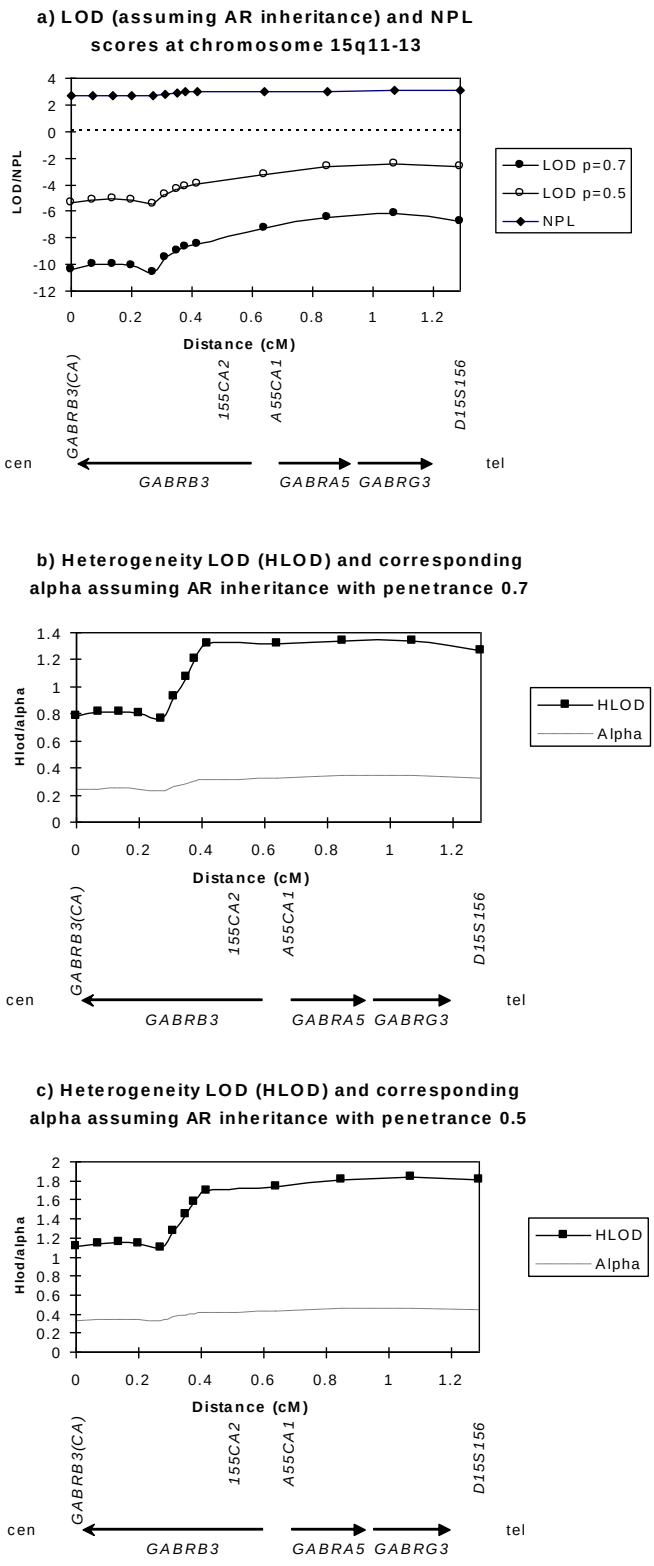



**Figure 12-6** *Fifty-three CAE pedigrees: Multipoint parametric (AR inheritance, penetrance 0.7 and 0.5) and non-parametric linkage analysis at GABAAR gene cluster, GABRB3, GABRA5 and GABRG3 on chromosome 15q11-13*

### *12.2.2* Eighty-two Absence families: linkage analysis of *GABRBB3, GABRA5, GABRG3*

*GABRBB3, GABRA5* and *GABRG3* were then tested in the eighty-two absence families by linkage analysis with criteria for affectedness CAE, JAE or AE as defined in section 1.1. Parametric linkage analysis was performed using GENEHUNTER assuming AD inheritance with penetrance of 0.7 and 0.5. The maximum LOD scores remained very negative in the larger cohort of AE families: maximum LOD= -12.0 with penetrance=0.7 and LOD= -9.1 with penetrance=0.5 (Figure 12-7(a)). The maximum HLOD was 2.04 (alpha=0.40) with penetrance of 0.7, and 2.31 (alpha=0.45) with penetrance of 0.5 (Figure 12-7(b,c)). The peak HLOD again occurred at marker *155CA2*, intragenic to *GABRB3*.

LOD scores were also calculated assuming AR inheritance (Figure 12-8). The homogeneity LODs remained very negative throughout, and the maximum HLODs were 1.83 (alpha=0.28) with penetrance 0.7, and 2.07 (alpha=0.35) with penetrance 0.5, at the position of marker *155CA2.*

The maximum NPL score was now 3.74 (p=.0.00004), also at the intragenic *GABRB3* marker *155CA2.*

The transmission disequilibrium test performed on each SSLP in turn identified one allele of *GABRB3CA* which showed a weakly positive association with the disease phenotype (transmitted 14, untransmitted 5, $\chi^2$=4.26, p=0.04), and one allele of *A55CA1* with a weakly negative effect (transmitted 26, untransmitted 46, $\chi^2$=5.56, p=0.02) (Appendix 21b).

The PDT also showed a negative association for one allele of marker *A55CA1* (p=0.02) (Appendix 23b).

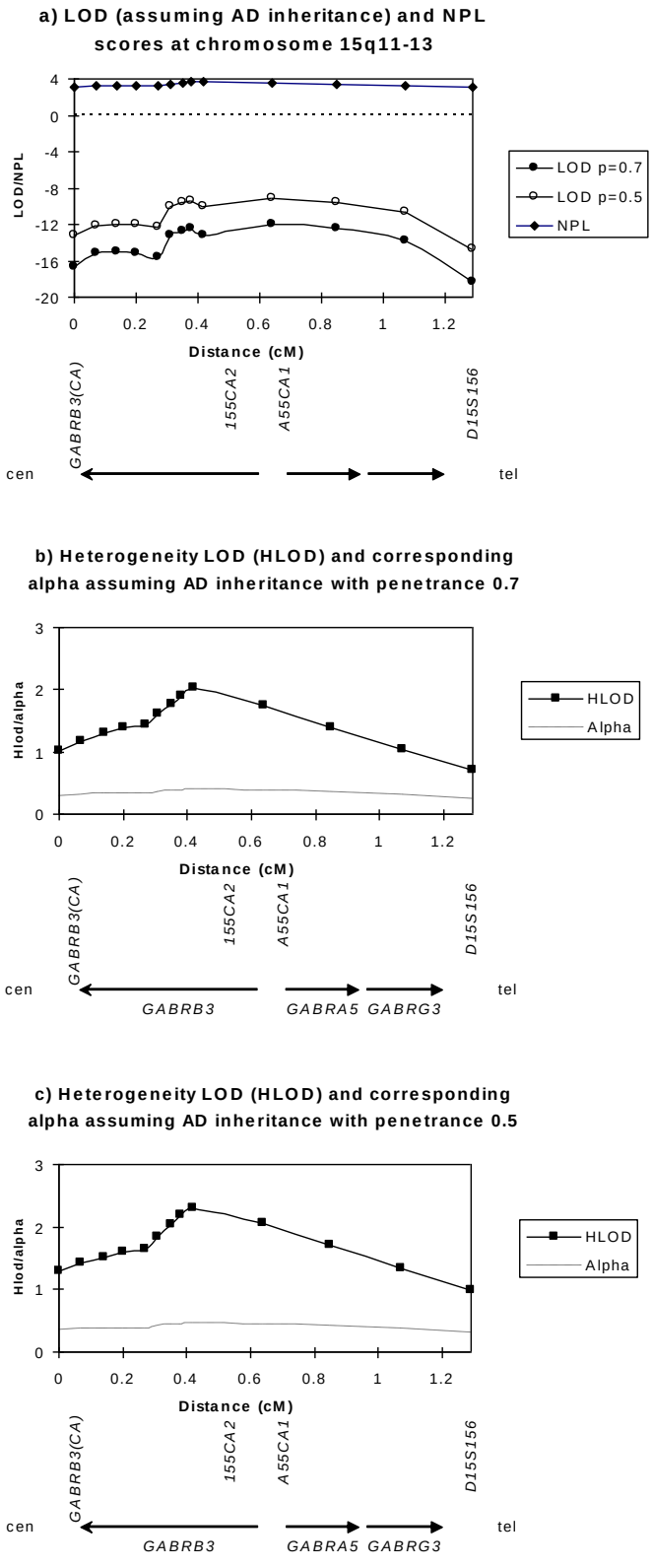



**Figure 12-7** *Eighty-two Absence pedigrees: Multipoint parametric (AD inheritance, penetrance 0.7 and 0.5) and non-parametric linkage analysis at $GABA_AR$ gene cluster, GABRB3, GABRA5 and GABRG3 on chromosome 15q11-13*

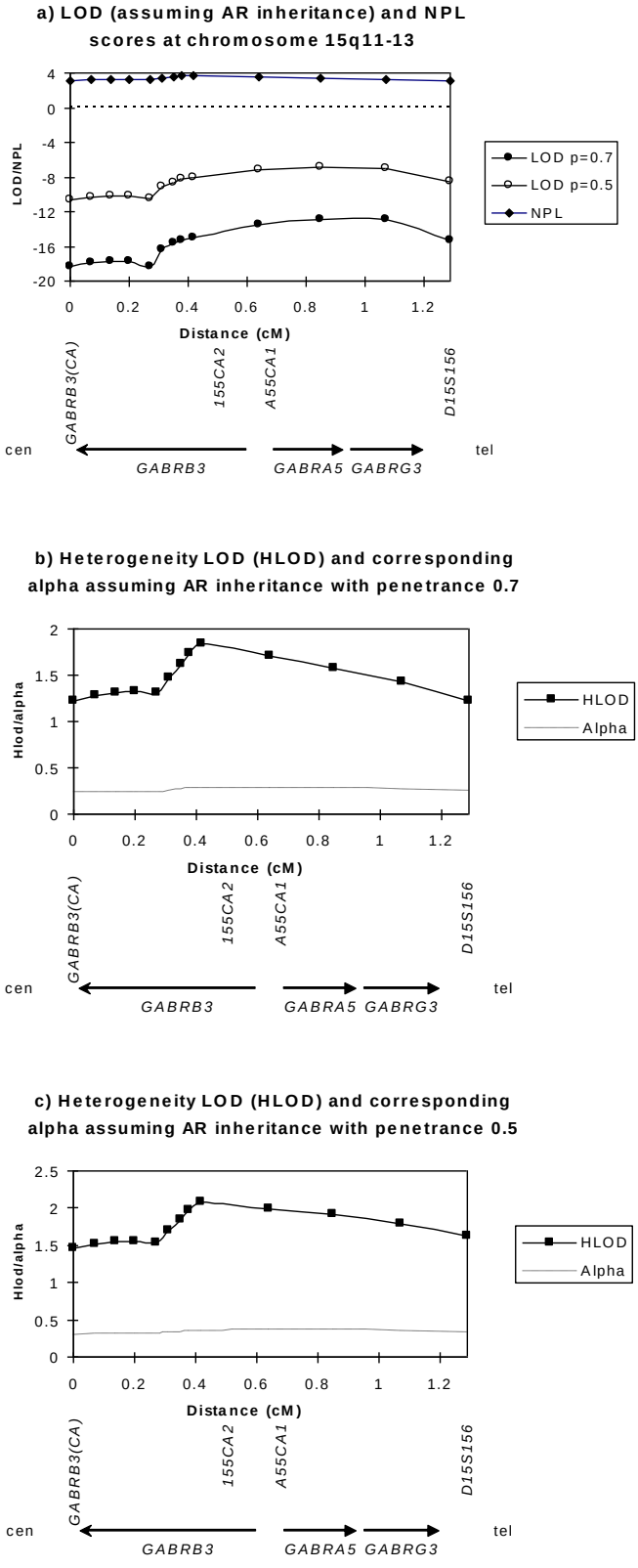



**Figure 12-8** *Eighty-two Absence pedigrees: Multipoint parametric (AR inheritance, penetrance 0.7 and 0.5) and non-parametric linkage analysis at $GABA_AR$ gene cluster, GABRB3, GABRA5 and GABRG3 on chromosome 15q11-13*

**Table 12-1** *Maximum HLODs at CACNG3 on chromosome 16p12-13.3 and GABRB3, GABRA5, GABRG3 on chromosome 15q11-13) for two phenotypes, CAE and Absence*

| Locus tested | Patient resource | Mode of inheritance | Disease allele penetrance | Maximum HLOD | Alpha | Information | Marker position |
|---|---|---|---|---|---|---|---|
| *CACNG3 (16p12-13.3)* | 53 CAE pedigrees | AD | 0.7 | 3.22 | 0.58 | 0.87 | *CACNG3EX1CA* |
| | | | 0.5 | 3.03 | 0.61 | 0.87 | *CACNG3EX1CA-UCL10321* |
| | 82 Absence pedigrees | AD | 0.7 | 3.44 | 0.53 | 0.85 | *CACNG3EX1CA-UCL10321* |
| | | | 0.5 | 3.36 | 0.55 | 0.88 | *UCL10321* |
| *GABRB3*, *GABRA5, GABRG3 (15q11-13)* | 53 CAE pedigrees | AD | 0.7 | 0.97 | 0.32 | 0.95 | *155CA2* |
| | | | 0.5 | 1.18 | 0.38 | 0.95 | *155CA2* |
| | | AR | 0.7 | 1.34 | 0.34 | 0.84 | *A55CA1-D15S156* |
| | | | 0.5 | 1.84 | 0.46 | 0.84 | *A55CA1-D15S156* |
| | 82 Absence pedigrees | AD | 0.7 | 2.04 | 0.40 | 0.97 | *155CA2* |
| | | | 0.5 | 2.31 | 0.45 | 0.97 | *155CA2* |
| | | AR | 0.7 | 1.83 | 0.28 | 0.97 | *155CA2* |
| | | | 0.5 | 2.07 | 0.35 | 0.97 | *155CA2* |

**Table 12-2** *Maximum NPL scores at CACNG3 on chromosome 16p12-13.3 and GABRB3, GABRA5, GABRG3 on chromosome 15q11-13) for two phenotypes, CAE and Absence*

| Locus tested | Patient resource | Maximum NPL | p-value | Information | Marker position |
|---|---|---|---|---|---|
| *CACNG3 (16p12-13.3)* | 53 CAE pedigrees | 2.59 | 0.00303 | 0.91 | *CACNG3EX1CA* |
| | 82 Absence pedigrees | 2.74 | 0.00195 | 0.88 | *UCL10321* |
| *GABRB3*, *GABRA5, GABRG3 (15q11-13)* | 53 CAE pedigrees | 3.05 | 0.00059 | 0.90 | *D15S156* |
| | 82 Absence pedigrees | 3.75 | 0.00004 | 0.97 | *155CA2* |

## 12.3 Multilocus linkage analysis

The CACNG3 locus on chromosome 16p12-p13.1 and the GABAAR subunit cluster on chromosome 15q11-13 were tested simultaneously using GENEHUNTER-TWOLOCUS. Three dominant digenic models were tested (penetrance matrices shown in Figure 8-1). The results of the two-locus parametric and non-parametric linkage analysis are compared with the single-locus linkage results in Table 12-3 and Table 12-4.

### 12.3.1 Fifty-three CAE families: two-locus linkage analysis

In the 53 CAE families, the multiplicative model produced the maximum two-locus HLOD, 4.84 (alpha =0.54) at marker *D16S420* on chromosome 16 and *155CA2* on chromosome 15, corresponding to the 5' end of *CACNG3* and the 5' end of *GABRB3* respectively (Figure 12-9(a)). The heterogeneity and additive models yielded maximum HLOD scores of 3.02 and 4.01, respectively. The maximum NPL score was 3.99 (p=0.000012) at marker *UCL10321* on chromosome 16, which lies 3' of *CACNG3*, and *D15S156* on chromosome 15, lying 3' of *GABRG3* (Figure 12-9(b)). This p value reflects a test of the null hypothesis that both disease loci are unlinked to their corresponding marker maps.

### 12.3.2 Eighty-two Absence families: two-locus linkage analysis

The multiplicative model again produced the maximum two-locus HLOD, 5.71 (alpha=0.47) between markers *CACNG3EX1CA* and *UCL10321* on chromosome 16 (centromeric to *CACNG3*), and at marker *155CA2* on chromosome 15 (5' end of *GABRB3*) (Figure 12-10(a)). The heterogeneity and additive models produced maximum HLOD scores of 2.82 and 4.67, respectively. The maximum NPL was 4.57 (p=0.0000007459) at *UCL10321* on chromosome 16 and *155CA2* on chromosome 15 (Figure 12-10(b)).

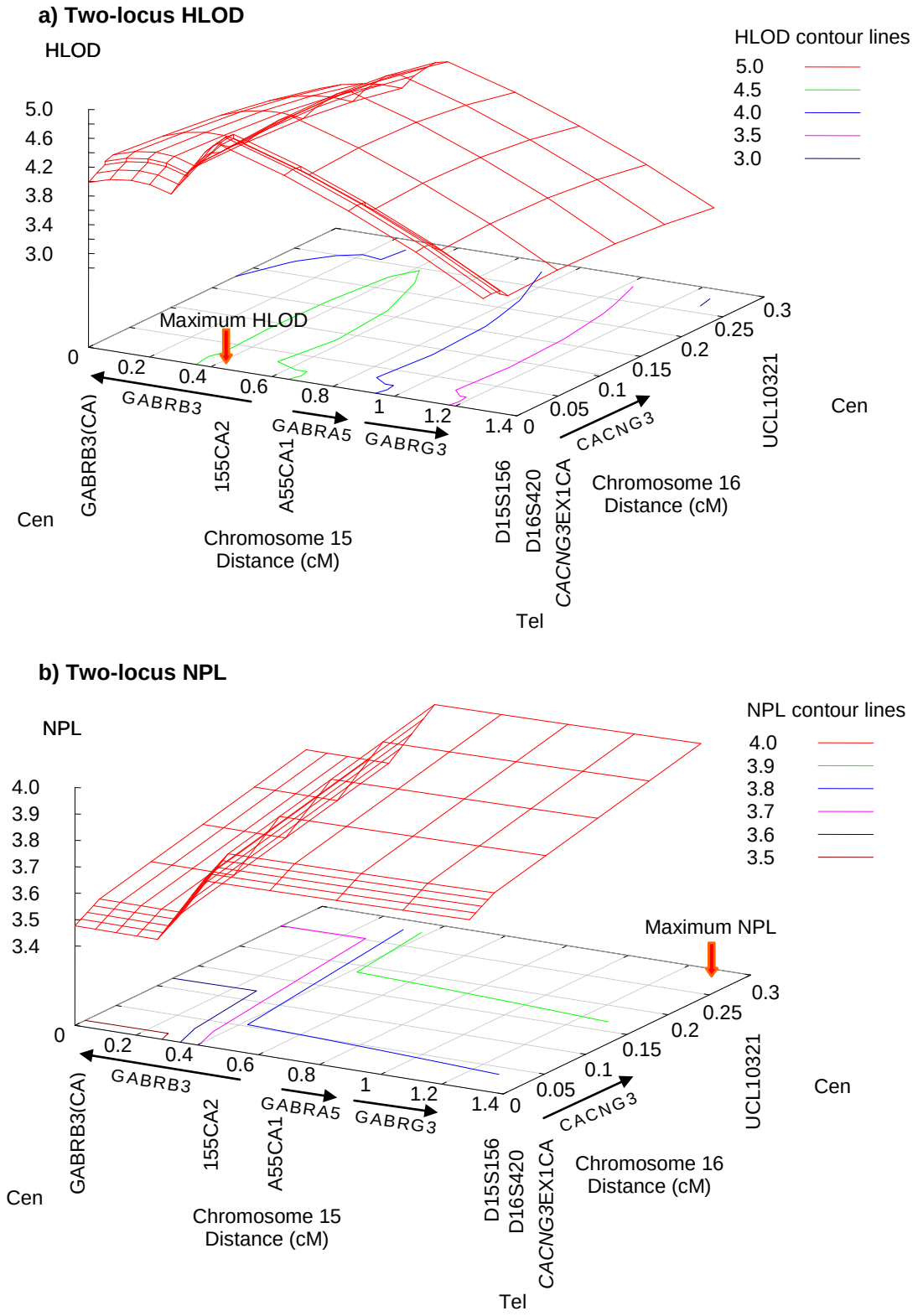



**Figure 12-9** *Fifty-three CAE pedigrees: Two-locus (a) parametric (multiplicative model) and (b) non-parametric linkage analysis at CACNG3 locus on chromosome 16p12-13.3 and GABA$_A$R gene cluster, GABRB3, GABRA5 and GABRG3 on chromosome 15q11-13*

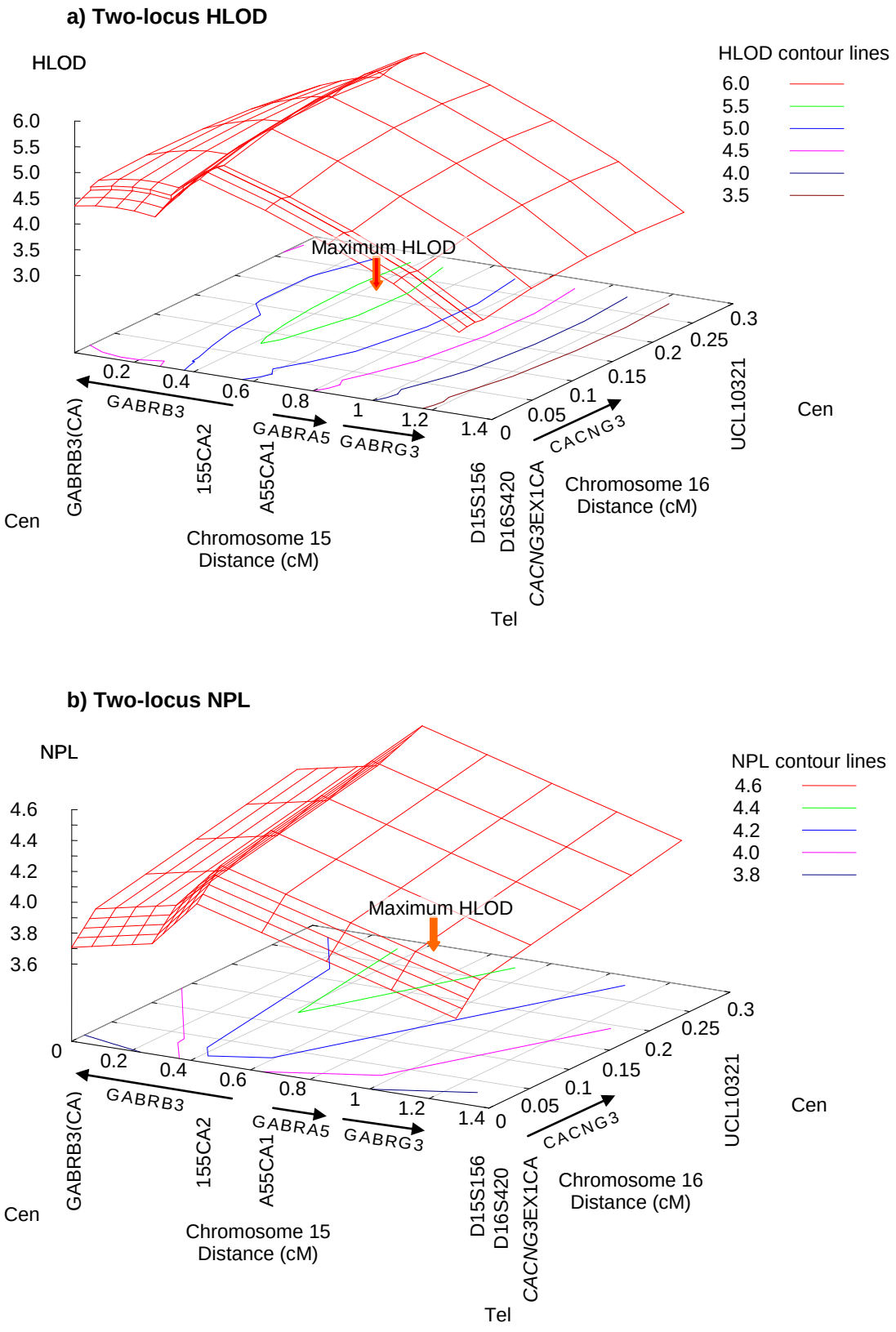



**Figure 12-10** *Eighty two Absence pedigrees: Two-locus (a) parametric (multiplicative model) and (b) non-parametric linkage analysis at CACNG3 locus on chromosome 16p12-13.3 and GABA$_A$R gene cluster, GABRB3, GABRA5 and GABRG3 on chromosome 15q11-13*

**Table 12-3** *Comparison of parametric linkage results from 2-locus and single-locus analyses for CACNG3 and the 15q GABA$_A$ receptor gene cluster*

| Loci tested | Patient resource | Max HLOD | Alpha | Info | Marker position chromosome 16 | Marker position chromosome 15 |
|---|---|---|---|---|---|---|
| Two loci 16p & 15q | *Multiplicative model:* | | | | | |
| | 53 CAE pedigrees | 4.84 | 0.54 | 0.91 | *D16S420* | *155CA2* |
| | 82 Absence pedigrees | 5.71 | 0.47 | 0.90 | *CACNG3EX1CA-UCL10321* | *155CA2* |
| | *Heterogeneity model:* | | | | | |
| | 53 CAE pedigrees | 3.02 | 0.69 | 0.92 | *D16S420* | *155CA2* |
| | 82 Absence pedigrees | 2.82 | 0.56 | 0.91 | *D16S420* | *GABRB3(CA)* |
| | *Additive model:* | | | | | |
| | 53 CAE pedigrees | 4.01 | 0.99 | 0.85 | *D16S420* | *A55CA1-D15S156* |
| | 82 Absence pedigrees | 4.67 | 0.81 | 0.91 | *CACNG3EX1CA-UCL10321* | *155CA2* |
| *CACNG3 (16p12-13.3)* | 53 CAE pedigrees | 3.22 | 0.58 | 0.87 | *CACNG3EX1CA* | |
| | 82 Absence pedigrees | 3.44 | 0.53 | 0.85 | *CACNG3EX1CA-UCL10321* | |
| *GABRB3*, *GABRA5, GABRG3 (15q11-13)* | 53 CAE pedigrees | 1.84 | 0.46 | 0.84 | | *A55CA1-D15S156* |
| | 82 Absence pedigrees | 2.31 | 0.45 | 0.97 | | *155CA2* |

**Table 12-4** *Comparison of non-parametric linkage results from 2-locus and single-locus analyses for CACNG3 and the 15q $GABA_A$ receptor gene cluster*

| Loci tested | Patient resource | Max NPL | p-value | Info | Marker position chromosome 16 | Marker position chromosome 15 |
|---|---|---|---|---|---|---|
| Two loci 16p & 15q | 53 CAE pedigrees | 3.99 | 0.000012 | 0.89 | *UCL10321* | *D15S156* |
| | 82 Absence pedigrees | 4.57 | 7.46E-07 | 0.91 | *UCL10321* | *155CA2* |
| *CACNG3 (16p12-13.3)* | 53 CAE pedigrees | 2.59 | 0.00303 | 0.91 | *CACNG3EX1CA* | |
| | 82 Absence pedigrees | 2.74 | 0.00195 | 0.88 | *UCL10321* | |
| *GABRB3*, *GABRA5, GABRG3 (15q11-13)* | 53 CAE pedigrees | 3.05 | 0.00059 | 0.9 | | *D15S156* |
| | 82 Absence pedigrees | 3.75 | 0.00004 | 0.97 | | *155CA2* |

# *Chapter 13 Sequencing of CACNG3 coding sequence*

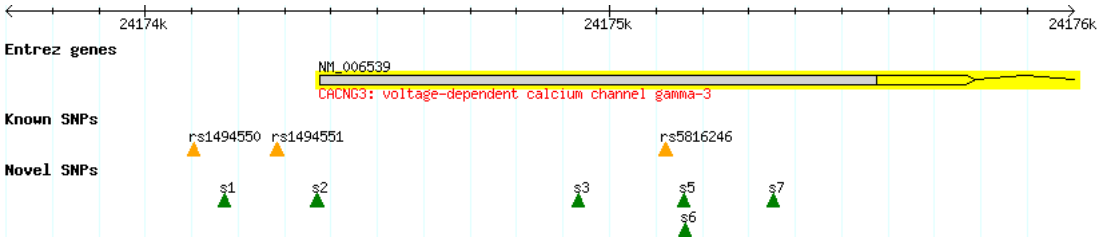
Direct sequencing of exonic DNA and surrounding intronic and promoter sequence DNA of *CACNG3* in 73 individuals identified 34 sequence variants (Figure 13-1). Four were upstream of the gene, five in the 5'UTR, six in intron 1, five in intron 2, nine in intron 3, one in exon 4, two in the 3'UTR and two downstream of the gene. Sixteen were novel variants, and the remaining 18 were known SNPs.

Details of all variants found are given in Table 13-1. The one coding sequence variant was identified in 4 cases, rs12928078, which is a synonymous SNP (c.2121A $\rightarrow$G, P307P).

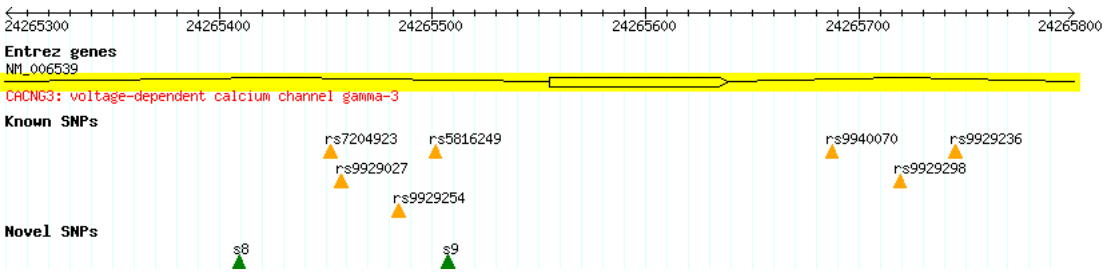
The minor allele frequency (MAF) of each variant in the sequenced cases was compared with that quoted in the NCBI database (if that information was available), as an initial assessment of a possible causal effect. The two SNPs showing the greatest difference in MAF between cases and controls were rs1494550 (0.340 vs 0.209) and rs12932291 (0.238 vs 0.129). SNP rs1494550 also showed the maximum MAF in the sequenced cases. These two SNPs were therefore selected for typing in the whole resource and included in the association analysis (Chapter 14).

Sequence conservation in three species for each variant is shown in Table 13-1. Eight SNPs occurred in sequence known to be conserved in all three species, increasing the likelihood that these variants may have functional significance.

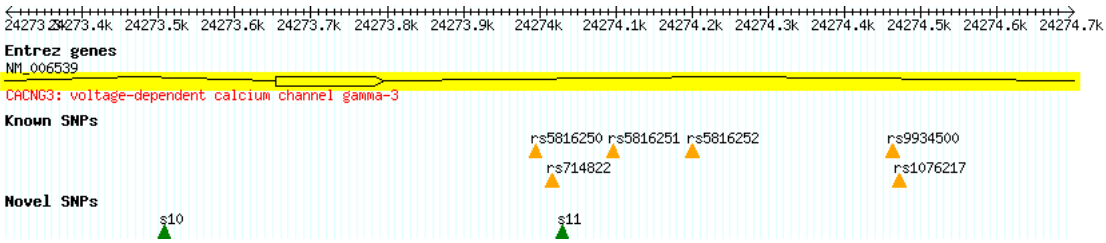
*a) Variants identified in region of exon 1*




*b) Variants identified in region of exon 2*




*c) Variants identified in region of exon 3*




*d) Variants identified in region of exon 4*

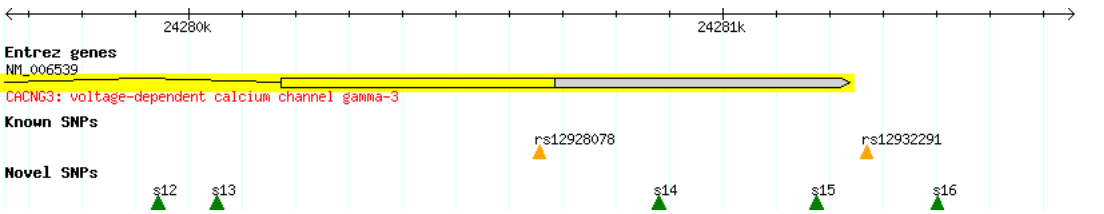



**Figure 13-1** *Maps illustrating locations of all variants in CACNG3 found by sequencing of 73 affected individuals.*

**Table 13-1** *Details of variants found during sequencing of CACNG3 exons and neighbouring intronic regions.*

| SNP | Position | Reference allele | Variant allele | Control Population MAF (NCBI SNP database) | Case Population MAF (73 sequenced cases) | Conserved in *Pan troglodytes* | Conserved in *Rattus norvegicus* | Conserved in *Mus musculus* |
|---|---|---|---|---|---|---|---|---|
| rs1494550 | 24174105 | T | C | 0.209 | 0.340 (0.287[2]) | Yes | Yes | Yes |
| s1 | 24174172 | G | C | | 0.009 | Yes | Yes | Yes |
| rs1494551 | 24174285 | A | G | 0.278 | 0.198 | Yes | No | No |
| s2 | 24174371 | C | T | | 0.019 | Yes | Unknown | Unknown |
| s3 | 24174934 | C | A | | 0.017 | Unknown | No | Unknown |
| s4 | 24174980 | T | delT | 0.074[1] | 0.058 (0.054[2]) | Yes | Yes | Unknown |
| rs5816246 | 24175122 | CA | delCA | | 0.185 | Unknown | Yes | Yes |
| s5 | 24175161 | T | A | | 0.141 | Unknown | Yes | No |
| s6 | 24175163 | A | T | | 0.043 | Unknown | Yes | Yes |
| s7 | 24175353 | C | T | | 0.021 | Yes | Yes | Yes |
| s8 | 24265409 | A | delA | | 0.016 | Yes | Unknown | Unknown |
| rs7204923 | 24265452 | C | G | | 0.008 | Yes | Unknown | Unknown |
| rs9929027 | 24265457 | C | T | 0.044 | 0.008 | Yes | Unknown | Unknown |
| rs9929254 | 24265484 | G | A | 0.055 | 0.008 | Yes | Unknown | Unknown |
| rs5816249 | 24265501 | G | delG | | 0.159 | Yes | Unknown | Unknown |
| s9 | 24265507 | A | G | | 0.008 | No | Unknown | Unknown |
| rs9940070 | 24265687 | T | C | 0.003 | 0.016 | No | No | No |

| SNP | Position | Reference allele | Variant allele | Control Population MAF (NCBI SNP database) | Case Population MAF (73 sequenced cases) | Conserved in *Pan troglodytes* | Conserved in *Rattus norvegicus* | Conserved in *Mus musculus* |
|---|---|---|---|---|---|---|---|---|
| rs9929298 | 24265719 | C | G | 0.053 | 0.016 | No | No | No |
| rs9929236 | 24265745 | A | G | 0.053 | 0.016 | No | Unknown | Unknown |
| s10 | 24273508 | C | T | | 0.008 | Yes | Yes | Yes |
| rs5816250 | 24273995 | A | delA | | 0.134 | Yes | Yes | Yes |
| rs714822 | 24274017 | C | G | 0.127 | 0.176 | No | No | No |
| s11 | 24274030 | T | C | | 0.010 | Yes | No | No |
| rs5816251 | 24274096 | AA | delAA | | 0.038 | Yes | Partial | No |
| rs5816252 | 24274201 | – | delC | | 0.278 | Yes | Yes | Yes |
| rs9934500 | 24274463 | A | C | | 0.206 | Yes | Unknown | Unknown |
| rs1076217 | 24274472 | T | C | | 0.250 | Yes | Unknown | Unknown |
| s12 | 24279944 | T | C | | 0.010 | Yes | No | No |
| s13 | 24280054 | T | C | | 0.011 | Yes | No | No |
| rs12928078 | 24280658 | G | A | 0.051 | 0.045 | Unknown | Yes | Yes |
| s14 | 24280881 | – | insT | | 0.056 | Unknown | Yes | Yes |
| s15 | 24281176 | C | T | | 0.008 | Yes | Yes | Yes |
| rs12932291 | 24281270 | C | G | 0.129 | 0.238 | Yes | Yes | Yes |
| s16 | 24281402 | G | A | | 0.244 | Yes | Unknown | No |

[1] Based on 94 ECACC (European Collection of Cell Cultures) controls

[2] Based on 299 genotyped unrelated affecteds

## 13.1 Bioinformatic analysis of variants

Identified variants were assessed for any potential functional affect by searching for predicted regulatory motifs, splice sites and exonic splicing enhancers (ESEs) as described in Chapter 10.

### 13.1.1 Exonic splicing enhancer motifs

ESE finder identified 14 exonic splicing enhancer binding motifs, encompassing a total of seven SNPs (including the synonymous exon 4 SNP rs12928078) which could potentially be altered by an identified variant (Table 13-2). These alterations can cause aberrant splicing and subsequently result in exon skipping. Eight putative ESE motifs, containing a total of six of the SNPs, were abolished by the variant SNP allele (ie the score, which is a measure of agreement between the putative ESE motif and the consensus sequences for a specific SR protein, was below the threshold value for that protein). The scores at six other ESE motifs, encompassing five SNPs, were altered by the variant allele (five increased, one decreased), but the scores remained above the threshold for both alleles of each SNP.

### 13.1.2 Splice site prediction

Many donor and acceptor sites were predicted by Genesplicer, three of which could potentially be affected by the variants identified through sequencing (Table 13-3). A novel acceptor site is created when the non-reference sequence form of rs1494550 is present. Another acceptor site predicted to exist becomes significantly stronger when the variant form of SNP s11 is present. The variant form of SNP s16 alters the position of a potential donor site.

**Table 13-2** *Analysis of exonic splicing enhancers potentially altered by variants identified in CACNG3. The position of the indicated variant is highlighted in bold.*

| SNP | Location in gene | Alleles (reference/ variant) | SR protein | Threshold | Motif | Score with reference allele | Score with variant allele |
|---|---|---|---|---|---|---|---|
| s3 | 5'UTR | C/A | SC35 | 2.383 | **C**GCCCCCA | 3.329 | 3.605 |
| s5 | 5'UTR | T/A | SF2/ASF | 1.956 | CACAC**T**C | 2.262 | deleted |
| | | | SRp40 | 2.670 | **T**CACACA | 3.637 | deleted |
| s6 | 5'UTR | A/T | SF2/ASF | 1.956 | C**A**CACAG | 3.285 | 2.162 |
| | | | SRp40 | 2.670 | TC**A**CACA | 3.637 | deleted |
| | | | SRp40 | 2.670 | **A**CACAGA | 2.893 | 3.943 |
| s7 | 5'UTR | C/T | SF2/ASF | 1.956 | C**C**CCCGG | 3.143 | 3.746 |
| | | | SF2/ASF | 1.956 | **C**CCCGGA | 3.266 | deleted |
| rs12928078 | Exon 4 | G/A | SF2/ASF | 1.956 | C**G**GCCAA | 2.324 | 2.772 |
| | | | SRp55 | 2.676 | TCC**G**GC | 3.018 | deleted |
| | | | SC35 | 2.383 | **G**GCCAACA | 3.608 | deleted |
| s14 | 3'UTR | -/insT | SRp40 | 2.670 | TTTCAA**G** | 4.397 | deleted |
| s15 | 3'UTR | C/T | SF2/ASF | 1.956 | GCCC**C**GA | 2.293 | deleted |
| | | | SC35 | 2.383 | GAGCCC**C**G | 3.139 | 3.783 |

**Table 13-3** *Splice sites predicted by Genesplicer, potentially affected by the identified variants. The position of the indicated variant is identified in bold. The splice-site dinucleotide is underlined. Strength of site according to the scoring of Pertea et al (Pertea, Lin et al. 2001).*

a) With all variants in the reference sequence form:

| *SNP* | *Start position* | *Sequence* | *End position* | *Score* | *Splice site type* | *Direction* |
|---|---|---|---|---|---|---|
| s11 | 24274030 | **t**tgaaagaga | 24274039 | 4.91 | acceptor | reverse |
| s16 | 24281402 | **a**tgtccaa...gctgg | 24281471 | 1.47 | donor | reverse |

b) With all variants in the non-reference sequence form:

| *SNP* | *Start position* | *Sequence* | *End position* | *Score* | *Splice site type* | *Direction* |
|---|---|---|---|---|---|---|
| rs1494550 | 24174105 | **c**ta | 24174107 | 4.64 | acceptor | reverse |
| s11 | 24274030 | **c**tgaaagaga | 24274039 | 8.70 | acceptor | reverse |
| s16 | 24281402 | **g**tgtccaaa | 24281410 | 1.53 | donor | forward |

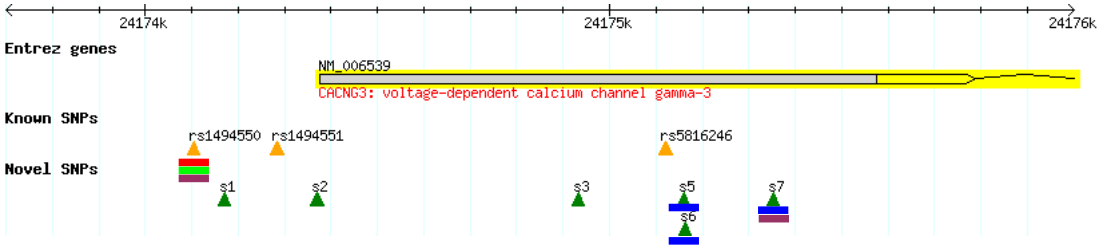
### 13.1.3 Other regulatory motifs

A search for regulatory motifs which could potentially be disrupted by any of the variants bases on the Transfac and Biobase databases (http://linux1.softberry.com/berry.phtml?topic=nsite&group=programs&subgroup=promoter) did not identify any clear candidates.

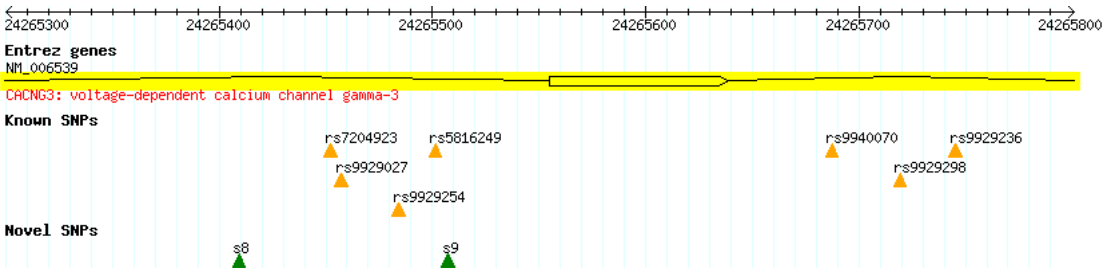
### 13.1.4 Summary of likely functional variants (Figure 13-2)

On the basis of the above bioinformatics analyses, the variants most likely to be functional are rs1494550 (splice site), s5 (ESE), s6 (ESE), s7 (ESE), s11 (splice site), rs12928078 (ESE), S14 (ESE), S15 (ESE). The variants rs1494550 and rs12932291 showed the greatest difference in MAF between cases and controls. These two SNPs, as well as s7 and s15, are shown to be conserved at the nucleotide level in the chimpanzee.

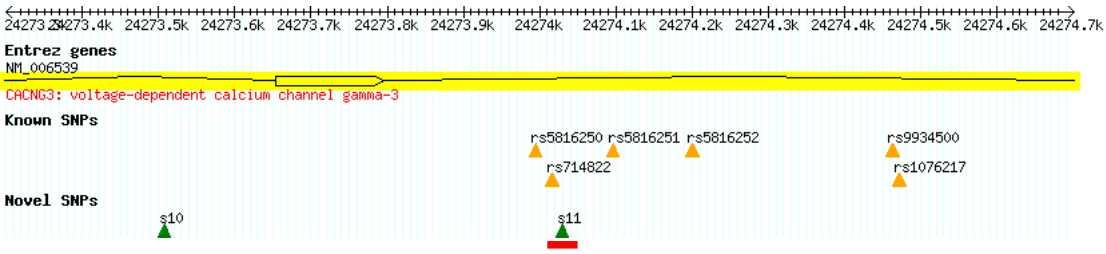
*a) Variants in region of exon 1*




*b) Variants in region of exon 2*




*c) Variants in region of exon 3*




*d) Variants in region of exon 4*

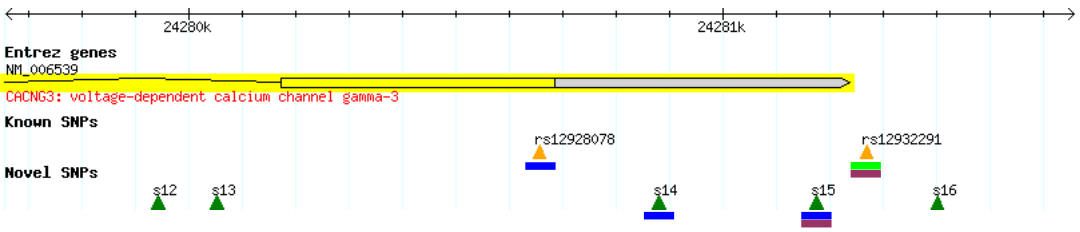



**Figure 13-2** *Maps illustrating locations of those variants in CACNG3 most likely to be functional (underlined variants).*

- Alters ESE
- Alters splice site
- Greatest MAF difference between cases and controls
- Conserved in chimpanzee

# *Chapter 14 SNP based association analysis of CACNG3*

As no clear functional candidates in *CACNG3* had been identified by sequencing of exonic DNA and surrounding intronic and promoter sequence in affected individuals, further SNP based analysis was performed to identify the LD block structure across the gene and to investigate association with the disease phenotypes in the entire resource of 82 extended pedigrees and 217 trios.

## 14.1 LD block structure of *CACNG3*

Analysis of LD based on the entire resource of 82 pedigrees and 217 trios identified 5 LD blocks (Figure 14-1). Block 1 encompassed exon 1 of *CACNG3*, block 4 exon 2, and block 5 exons 3 and 4. This was compared with the LD block structure predicted by the HapMap project genotyped SNPs (based on CEPH caucasian data for 233 SNPs). The HapMap data identified 21 blocks of LD across the same region (Figure 14-2).

To test whether a small subset of trios would generate a similar block structure to the entire resource, the SNP data from 31 trios, selected randomly, was analysed independently. As with the entire resource, 5 LD blocks were identified (Figure 14-3). There was some variation in the boundary SNPs of blocks 1 and 2, but blocks 3-5 were identical.


**Figure 14-1** *Block structure of CACNG3 locus based on entire patient resource as defined by Haploview using a minimum D' of 0.7.*

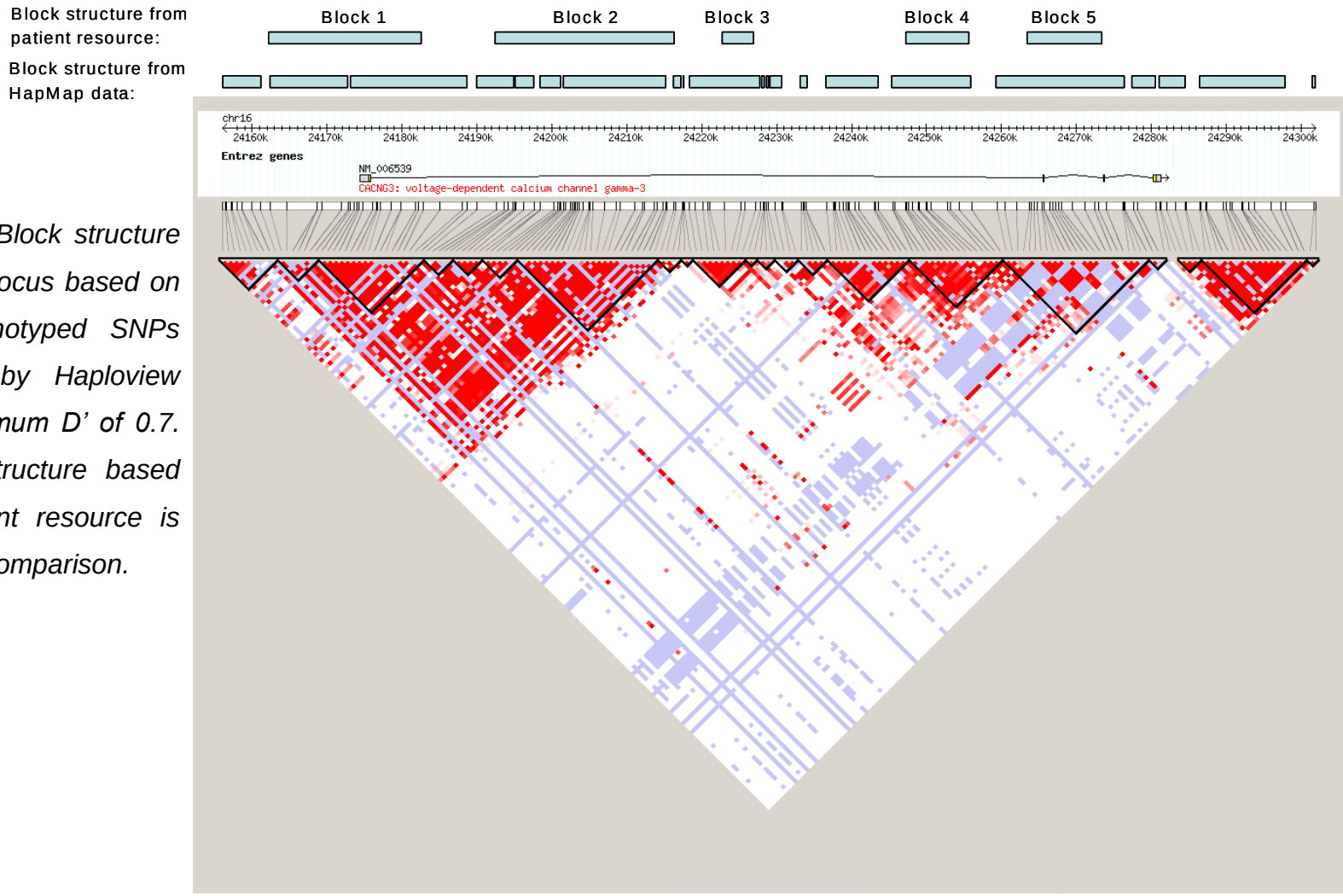



**Figure 14-2** *Block structure of CACNG3 locus based on HapMap genotyped SNPs as defined by Haploview using a minimum D' of 0.7. The block structure based on the patient resource is included for comparison.*

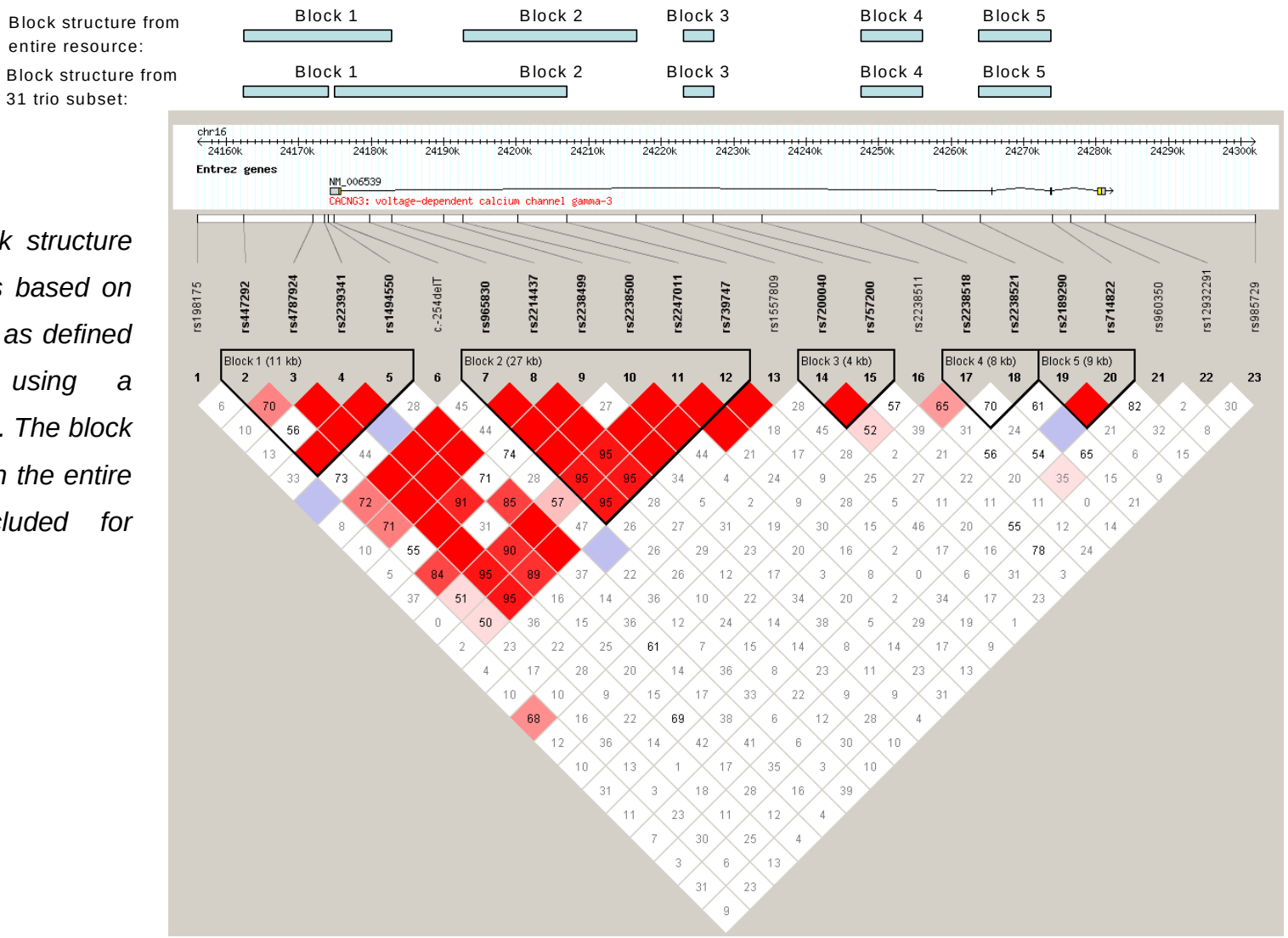



**Figure 14-3** *Block structure of CACNG3 locus based on subset of 31 trios as defined by Haploview using a minimum D' of 0.7. The block structure based on the entire resource is included for comparison.*

## 14.2 Tagging SNPs for *CACNG3*

Tagging SNP analysis using the two methodologies employed by Haploview produced similar results. Using the LD block structure and a haplotype frequency cut off of 1%, 19 tag SNPs were identified (Appendix 24a). The block haplotypes with frequency greater than 1% and the relevant tag SNPs are shown in Figure 14-4. The Tagger methodology identified 18 tag SNPs (Appendix 24a). The SNPs captured by the tag SNPs are shown in Appendix 24b.

For comparison, tag SNPs were also identified using the subset of 31 trios (Appendix 24a). Using the LD block structure and a haplotype frequency cut off of 1%, 20 tag SNPs were identified (incorporating the 19 tag SNPs identified in the whole resource). The Tagger methodology identified 19 tag SNPs (15 of which corresponded to those identified in the whole resource).

The HapMap SNP data across the same region was also used to identify tag SNPs. Of the 233 SNPs in total, 85 tag SNPs were identified using the LD block structure, and 76 SNPs identified by Tagger.

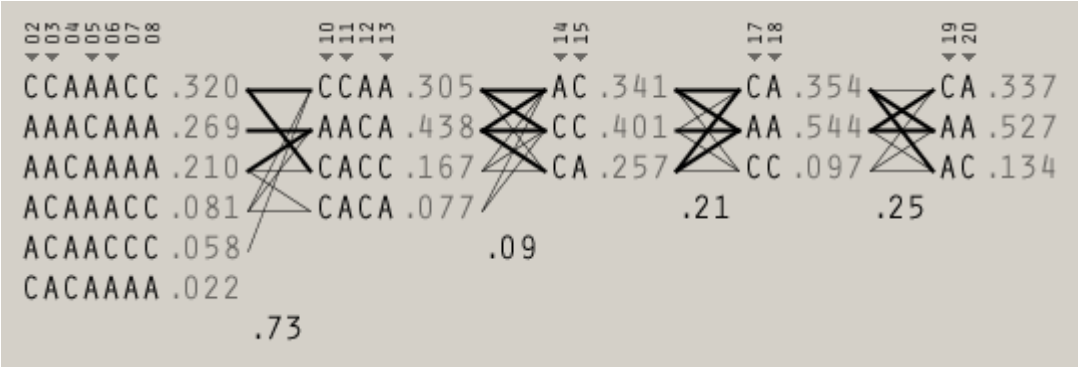



**Figure 14-4** *CACNG3 LD block haplotypes with frequency >1%, using the entire resource. Tag SNPs are indicated with arrows. The multiallelic D' is shown between blocks, representing the level of recombination between the two blocks.*

## 14.3 Association analysis of *CACNG3* – individual SNPs

### 14.3.1 Criteria for affectedness CAE

Two SNPs, SNP8 and SNP11, showed significant transmission disequilibrium ($p \leq 0.01$) by PDT with at least one of the test statistics (Table 14-1a). Both SNPs lie in intron 1; SNP8 in LD block 1 and SNP11 in LD block 2.

### 14.3.2 Criteria for affectedness CAE, JAE or AE

Three SNPs, SNP3, SNP7 and SNP8, showed significant transmission disequilibrium by PDT with at least one of the test statistics (Table 14-1b). All three SNPs are in LD block 1. SNP3 is located 2.3 kb upstream of *CACNG3* and SNP7 and SNP8 are located in intron 1. The full list of PDT results can be found in Appendix 25.

**Table 14-1** *SNPs showing statistically significant disease association ($p \leq 0.01$) in at least one PDT test statistic in the entire resource.*

a) Affectedness criteria CAE:

| | | | | *SUM PDT* | | *AVE PDT* | |
|---|---|---|---|---|---|---|---|
| *SNP* | *Allele* | *Transmitted* | *Not transmitted* | *Z (1df)* | *p-value* | *Z (1df)* | *p-value* |
| 8 | 2(A) | 266 | 222 | 2.64 | 0.008 | 2.47 | 0.013 |
| 11 | 2(G) | 179 | 139 | 2.90 | 0.004 | 2.49 | 0.013 |

b) Affectedness criteria CAE, JAE or AE:

| | | | | *SUM PDT* | | *AVE PDT* | |
|---|---|---|---|---|---|---|---|
| *SNP* | *Allele* | *Transmitted* | *Not transmitted* | *Z (1df)* | *p-value* | *Z (1df)* | *p-value* |
| 3 | 2(G) | 331 | 279 | 2.68 | 0.007 | 2.46 | 0.014 |
| 7 | 2(G) | 298 | 248 | 2.67 | 0.008 | 2.52 | 0.012 |
| 8 | 2(A) | 312 | 256 | 2.98 | 0.003 | 2.75 | 0.006 |

## 14.4 Haplotype-based association analysis of *CACNG3*

Block-based haplotype analysis was performed on the entire data set using the PDT. A 'sliding widow' approach was used on LD blocks 1 and 2 to identify associated haplotypes, as no single complete haplotype was sufficiently common to allow demonstration of disease association on the global level.

### 14.4.1 Affectedness criteria CAE

Using the sliding window approach, there are 9 haplotypes (2-4 SNPs) in block 1, consisting of combinations of SNPs 2-8, which demonstrate significant overtransmission and disease association with CAE ($p \leq 0.05$; Table 14-2a). The individual overtransmitted haplotypes within each window, when combined, form a larger haplotype composed of alleles 2211122. The frequency of this haplotype in the parental population is 26.4%.

The sliding window approach also produced some significant associations in LD block 2 (SNPs 10-13) (Table 14-2b). The combination of the individual overtransmitted haplotypes produced two haplotypes across the whole block, 2211 and 2221. The frequency of haplotype 2211 in the parental population was 23.8%, and the frequency of haplotype 2221 was 3.2%.

Full details of the significant PDT results for affection status CAE can be found in Appendix 26.

**Table 14-2** *SNP based sliding-window analysis of LD blocks 1 and 2 of CACNG3 showing windows which demonstrate significant (P≤0.05) global transmission disequilibrium in the entire resource when analysed using the PDT with affectedness criteria CAE. Only haplotypes showing significant (P≤0.05) overtransmission are shown*

a) LD block 1:

| *SNP* | | | | | | | *Frequency in parents* | | | *SUM PDT* | | *GLOBAL* | | *AVE PDT* | | *GLOBAL* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *2* | *3* | *4* | *5* | *6* | *7* | *8* | *(%)* | *Transmitted* | *Not Transmitted* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* |
| 2 | 2 | | | | | | 31.5 | 195 | 154 | 2.18 | 0.029 | $9.00_{(3)}$ | 0.029 | 2.69 | 0.007 | $9.43_{(3)}$ | 0.024 |
| | 2 | 1 | | | | | 46.2 | 250 | 208 | 2.54 | 0.011 | $8.44_{(3)}$ | 0.038 | 2.38 | 0.017 | $7.21_{(3)}$ | 0.066 |
| | | 1 | 1 | | | | 49.2 | 263 | 223 | 2.41 | 0.016 | $7.80_{(3)}$ | 0.050 | 2.24 | 0.025 | $6.62_{(3)}$ | 0.085 |
| | | 1 | 1 | 1 | | | 41.9 | 215 | 167 | 2.90 | 0.004 | $15.46_{(6)}$ | 0.017 | 2.83 | 0.005 | $15.29_{(6)}$ | 0.018 |
| | | | 1 | 1 | 2 | | 38.6 | 206 | 155 | 3.09 | 0.002 | $19.47_{(7)}$ | 0.007 | 3.08 | 0.002 | $19.71_{(7)}$ | 0.006 |
| | | | 1 | 1 | 2 | 2 | 37.9 | 198 | 145 | 3.26 | 0.001 | $26.20_{(12)}$ | 0.010 | 3.35 | 0.001 | $26.76_{(12)}$ | 0.008 |
| | | | | 1 | 2 | | 41.0 | 219 | 168 | 3.09 | 0.002 | $16.19_{(3)}$ | 0.001 | 3.08 | 0.002 | $16.25_{(3)}$ | 0.001 |
| | | | | 1 | 2 | 2 | 40.4 | 211 | 157 | 3.31 | 0.001 | $24.24_{(6)}$ | 0.001 | 3.43 | 0.001 | $25.12_{(6)}$ | 0.000 |
| | | | | | 2 | 2 | 46.9 | 259 | 210 | 2.95 | 0.003 | $15.15_{(3)}$ | 0.002 | 2.91 | 0.004 | $14.58_{(3)}$ | 0.002 |

b) LD block 2:

| *SNP* | | | | *Frequency* | | | *SUM PDT* | | *GLOBAL* | | *AVE PDT* | | *GLOBAL* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *10* | *11* | *12* | *13* | *in parents (%)* | *Transmitted* | *Not Transmitted* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* |
| 2 | 2 | | | 29.9 | 169 | 123 | 3.32 | 0.001 | $11.21_{(3)}$ | 0.011 | 3.09 | 0.002 | $10.72_{(3)}$ | 0.013 |
| | 2 | 1 | | 23.9 | 96 | 71 | 2.27 | 0.023 | $12.58_{(3)}$ | 0.006 | 2.23 | 0.026 | $12.37_{(3)}$ | 0.006 |
| 2 | 2 | 1 | | 23.4 | 93 | 68 | 2.30 | 0.021 | | | 2.30 | 0.021 | | |
| 2 | 2 | 2 | | 4.4 | 19 | 11 | 2.13 | 0.033 | $14.01_{(7)}$ | 0.051 | 2.65 | 0.008 | $16.36_{(7)}$ | 0.022 |
| | 2 | 1 | 1 | 24.4 | 95 | 69 | 2.35 | 0.019 | $9.45_{(6)}$ | 0.150 | 2.35 | 0.019 | $12.14_{(6)}$ | 0.059 |

### 14.4.2 Affectedness criteria CAE, JAE or AE

With the broader affectedness criteria of CAE, JAE or AE, the sliding window approach identified 13 haplotypes in block 1 (SNPs 2-8), which demonstrate significant overtransmission and disease association ($p \leq 0.05$; Table 14-3a). The larger haplotype formed by the individual overtransmitted haplotypes within each window again consists of the alleles 2211122, as was found with affectedness criteria CAE only.

The sliding window results in LD block 2 (SNPs 10-13) are shown in Table 14-3b. Seven haplotypes demonstrate a significant disease association ($p \leq 0.05$). The combination of the individual overtransmitted haplotypes produced three haplotypes across the whole block, 2211 and 2221 (as with affectedness criteria CAE) and 2222. The frequency of haplotype 2222 in the parental population was 1.7%.

Full details of the significant PDT results for affection status CAE, JAE or AE can be found in Appendix 27.

**Table 14-3** *SNP based sliding-window analysis of LD blocks 1 and 2 of CACNG3 showing windows which demonstrate significant (P≤0.05) global transmission disequilibrium in the entire resource when analysed using the PDT with affectedness criteria CAE, JAE or AE. Only haplotypes showing significant (P≤0.05) overtransmission are shown*

a) LD block 1

| *SNP* | | | | | | | *Frequency* | | | *SUM PDT* | | *GLOBAL* | | *AVE PDT* | | *GLOBAL* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *2* | *3* | *4* | *5* | *6* | *7* | *8* | *in parents (%)* | *Transmitted* | *Not Transmitted* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* |
| 2 | 2 | | | | | | 31.9 | 236 | 187 | 2.56 | 0.011 | $10.74_{(3)}$ | 0.013 | 2.67 | 0.008 | $9.83_{(3)}$ | 0.020 |
| 2 | 2 | 1 | | | | | 31.6 | 212 | 165 | 2.55 | 0.011 | $15.97_{(7)}$ | 0.025 | 2.67 | 0.008 | $15.10_{(7)}$ | 0.035 |
| | 2 | 1 | | | | | 46.1 | 303 | 252 | 2.70 | 0.007 | $10.19_{(3)}$ | 0.017 | 2.60 | 0.009 | $8.96_{(3)}$ | 0.030 |
| | 2 | 1 | 1 | 1 | | | 36.4 | 224 | 172 | 3.22 | 0.001 | $21.27_{(11)}$ | 0.031 | 3.15 | 0.002 | $21.12_{(11)}$ | 0.032 |
| | | 1 | 1 | | | | 48.8 | 313 | 269 | 2.42 | 0.016 | $8.38_{(3)}$ | 0.039 | 2.26 | 0.024 | $7.17_{(3)}$ | 0.067 |
| | | 1 | 1 | 1 | | | 41.9 | 255 | 204 | 3.08 | 0.002 | $16.21_{(6)}$ | 0.013 | 2.93 | 0.003 | $15.25_{(6)}$ | 0.018 |
| | | 1 | 1 | 1 | 2 | | 36.9 | 224 | 174 | 3.11 | 0.002 | $21.49_{(11)}$ | 0.029 | 2.93 | 0.003 | $20.97_{(11)}$ | 0.034 |
| | | 1 | 1 | 1 | 2 | 2 | 36.5 | 217 | 167 | 3.17 | 0.002 | $26.82_{(16)}$ | 0.044 | 3.03 | 0.003 | $25.93_{(16)}$ | 0.055 |
| | | | 1 | 1 | 2 | | 38.5 | 242 | 188 | 3.27 | 0.001 | $22.30_{(7)}$ | 0.002 | 3.17 | 0.002 | $22.71_{(7)}$ | 0.002 |
| | | | 1 | 1 | 2 | 2 | 38 | 234 | 178 | 3.44 | 0.001 | $29.11_{(12)}$ | 0.004 | 3.44 | 0.001 | $29.78_{(12)}$ | 0.003 |
| | | | | 1 | 2 | | 40.9 | 259 | 204 | 3.31 | 0.001 | $17.40_{(3)}$ | 0.001 | 3.17 | 0.002 | $16.99_{(3)}$ | 0.001 |
| | | | | 1 | 2 | 2 | 40.4 | 251 | 193 | 3.52 | 0.000 | $25.96_{(6)}$ | 0.000 | 3.52 | 0.000 | $26.28_{(6)}$ | 0.000 |
| | | | | | 2 | 2 | 46.9 | 313 | 253 | 3.19 | 0.001 | $17.64_{(3)}$ | 0.001 | 3.19 | 0.001 | $17.23_{(3)}$ | 0.001 |

b) LD block 2

| *SNP* | | | | *Frequency* | | | *SUM PDT* | | *GLOBAL* | | *AVE PDT* | | *GLOBAL* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *10* | *11* | *12* | *13* | *in parents (%)* | *Transmitted* | *Not Transmitted* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* |
| 2 | 2 | | | 29.6 | 199 | 151 | 2.72 | 0.007 | $7.50_{(3)}$ | 0.058 | 2.87 | 0.004 | $9.04_{(3)}$ | 0.029 |
| | 2 | 1 | | 23.5 | 103 | 75 | 2.45 | 0.014 | | | 2.35 | 0.019 | | |
| | 2 | 2 | | 5.4 | 24 | 17 | 2.11 | 0.035 | $15.27_{(3)}$ | 0.002 | 2.11 | 0.035 | $14.52_{(3)}$ | 0.002 |
| 2 | 2 | 1 | | 23.1 | 100 | 72 | 2.49 | 0.013 | | | 2.42 | 0.015 | | |
| 2 | 2 | 2 | | 4.4 | 21 | 12 | 2.12 | 0.034 | $15.07_{(7)}$ | 0.035 | 2.31 | 0.021 | $15.36_{(7)}$ | 0.032 |
| | 2 | 1 | 1 | 24.0 | 102 | 73 | 2.53 | 0.011 | | | 2.46 | 0.014 | | |
| | 2 | 2 | 2 | 1.6 | 8 | 4 | 1.28 | 0.201 | $11.59_{(6)}$ | 0.072 | 1.13 | 0.257 | $13.36_{(6)}$ | 0.038 |

# *Chapter 15 Linkage analysis of the family with absence epilepsy and episodic ataxia*

As mutations in *CACNA1A* have previously been identified in individuals with absence epilepsy and episodic ataxia, *CACNA1A* was tested by linkage analysis in the family described in section 5.9.

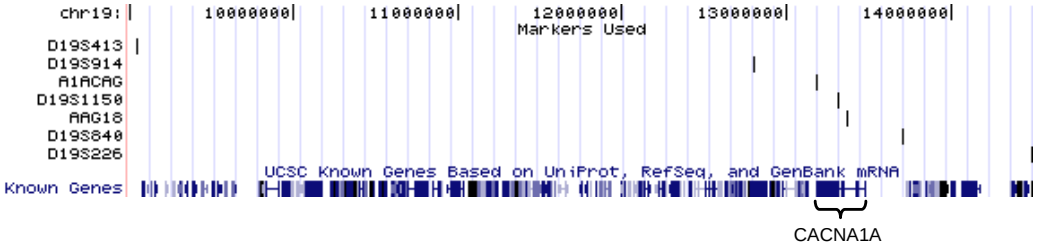
DNA was available from eleven family members (not individual 101), including the five with the AEA phenotype. The maximum simulated LOD score using SLINK was 2.4 at θ=0, under the assumption of autosomal dominant inheritance with 100% penetrance. *CACNA1A* was tested initially by linkage analysis using the intragenic marker *D19S1150* plus another six SSLPs. The seven markers span 5.4Mb encompassing *CACNA1A*, and include three *CACNA1A* intragenic polymorphisms: *AAG18*(Trettel, Mantuano et al. 2000) (intron 3), *D19S1150*(Ophoff, Terwindt et al. 1996) (intron 7) and *A1ACAG*(Ophoff, Terwindt et al. 1996) (CAG repeat in the 3' untranslated region) (Figure 15-1).

Five of the seven microsatellite markers (*D19S413, D19S914, D19S1150, D19S840, and D19S226*) were informative in this family. The five affected individuals shared a single common haplotype between *D19S413* and *D19S840* (Figure 15-2). This haplotype was not shared by any of the unaffected individuals including individual 319 who had a typical 3 Hz spike–wave EEG but no clinical phenotype and was neurologically normal. Assuming AD inheritance with penetrance=1, phenocopy rate=0 and disease allele frequency=0.005, the multipoint LOD score across this haplotype was 2.1 (Figure 15-3). The AEA phenotype in this pedigree was thus consistent with linkage to *CACNA1A*.

Subsequent DNA sequencing of each exon and the flanking intronic regions of *CACNA1A* was performed in all available family members at the Department of Molecular Neuroscience and Clinical and Experimental Epilepsy, Institute of Neurology, University College London.

This sequencing identified a heterozygous point mutation (G439A) in exon 3 of *CACNA1A*, which results in a substitution of lysine for glutamic acid at codon 147 (E147K). This previously unreported change segregated with the AEA disease phenotype. The mutation was not detected in unaffected individuals or in the patient with an abnormal EEG but no clinical phenotype (individual 319). All family members were sequenced. The mutation identified results in the loss of a *Taq*1 restriction site. This mutation was absent from a panel of 200 control chromosomes, and segregation of the mutant with the AEA disease was confirmed in the family. The mutation affects a highly conserved residue within the second transmembrane segment of domain I of CaV2.1$\alpha$1, the main pore-forming subunit of CaV2.1 channels.

a) Broad view showing all seven markers




b) Detailed view of intragenic markers

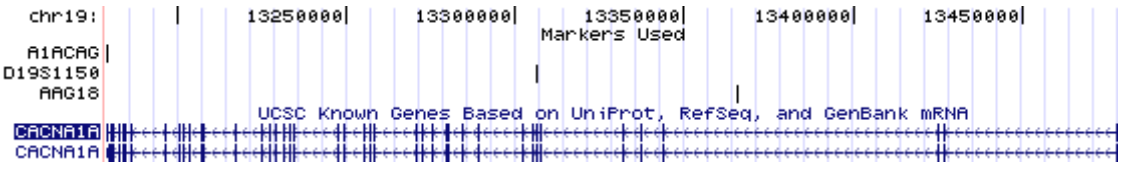



**Figure 15-1** *Position of seven SSLPs relative to CACNA1A on chromosome 19p13.3. Image modified from the UCSC Genome Browser (http://genome.cse.ucsc.edu/cgi-bin/hgGateway)*

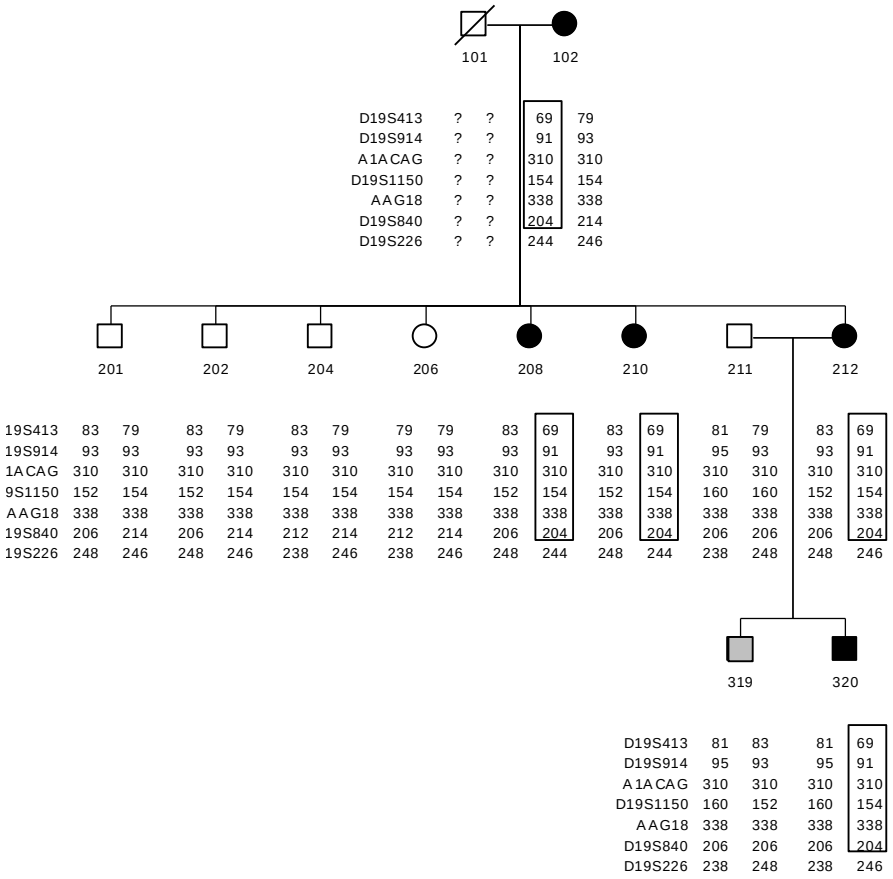



**Figure 15-2** *Haplotypes of AEA family members using seven microsatellite markers spanning 5.4Mb encompassing CACNA1A. The "disease" haplotype is outlined.*

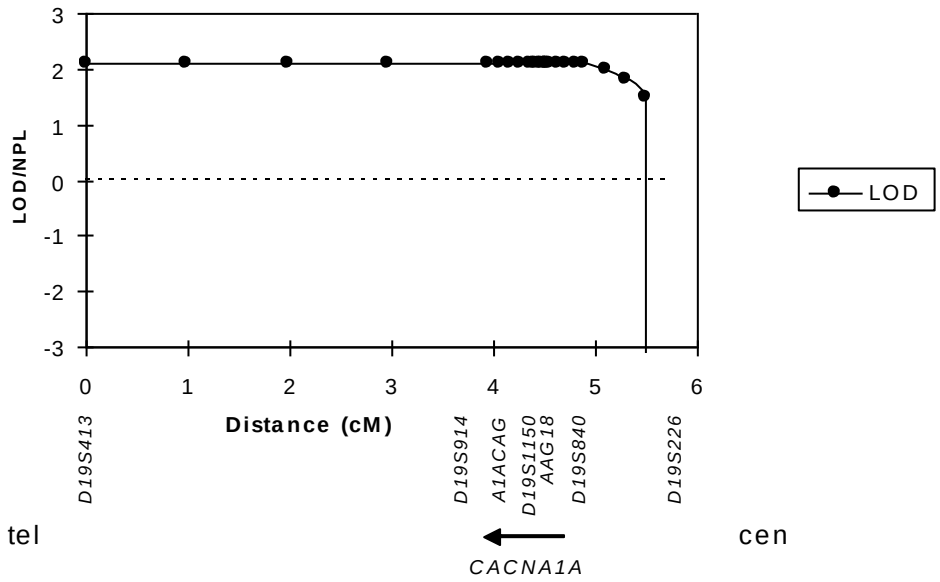



**Figure 15-3** *AEA family: Parametric linkage analysis, assuming AD inheritance, at CACNA1A locus on chromosome 19p13.3*

# PART 4 Discussion

# *Chapter 16 Genetic analysis of human absence epilepsy*

## 16.1 Summary

The aim of this project was to elucidate the molecular genetic basis of human absence epilepsy. The genetic analysis of complex traits is complicated by several issues which are now well recognized: the uncertainty of the genetic architecture of the condition, the selection of criteria which best delineate the phenotype, the lack of correlation between phenotype and genotype, the degree of locus and allelic heterogeneity and variable expressivity, all of which influence the optimum strategy for investigation. The approach employed here involved (i) ascertainment of a patient resource including affected sibling pairs, trios of affected individuals with both parents, and extended nuclear pedigrees and (ii) investigation of candidate genes using linkage analysis, SNP-based association analysis and mutation screening. Ion channel genes emerged as excellent candidates. Linkage analysis was performed using two phenotypic criteria, a specific diagnosis of Childhood Absence Epilepsy and a broader category of 'absence' which included CAE, JAE and individuals with many of the features of CAE and an onset of absence seizures below the age of 12 years, but who could not be classified as CAE due to certain atypical features. Two loci showed positive linkage: *CACNG3* on chromosome 16p12-13.3 and the $GABA_AR$ gene cluster *GABRB3*, *GABRA5* and *GABRG3* on chromosome 15q11-13. Multilocus linkage analysis was consistent with a digenic effect from these loci. Sequence analysis of *CACNG3* identified 34 sequence variants but no definite functional changes. SNP based association analysis of *CACNG3* showed significant transmission disequilibrium for SNPs in intron 1 (affectedness criteria CAE or Absence) and upstream of the gene (affectedness criteria Absence only). The LD block structure across the gene was identified, and the sliding window approach revealed haplotypes in LD block 1 (encompassing exon 1) and LD block 2 (encompassing part of intron

1) which were significantly associated with both CAE and Absence. This data provides supportive evidence for functional variants in both *CACNG3* and at least one of the *$GABA_AR$* genes on chromosome 15q. Further sequence and association analysis of *CACNG3* has subsequently been performed, providing additional evidence for *CACNG3* as a susceptibility locus in a subset of patients with CAE (Everett, Chioza et al. 2007).

## 16.2 Selection of phenotype

CAE is an idiopathic generalised epilepsy of childhood which is a well characterised and consistent phenotype in terms of age of onset, clinical features and EEG findings. There is strong evidence for a genetic aetiology, with a concordance in phenotype of 28% between probands with CAE and first and second degree relatives with epilepsy (Marini, Scheffer et al. 2004). Therefore the ascertainment of patients with the specific diagnosis of CAE for this study provides a homogeneous clinical phenotype with an expectation of a common genetic aetiology. However, there is also a strong genetic relationship between CAE and JAE, with 31% of affected relatives of a proband with JAE having CAE. This supports the existence of common susceptibility loci for absence epilepsy, including CAE and JAE. It was therefore decided to investigate, as well as pure CAE, the broader phenotype of 'absence epilepsy', incorporating CAE and JAE. Clustering of other IGE phenotypes is also seen in families with a proband with CAE, and rarely absence epilepsy may evolve into another IGE phenotype such as JME. This raises the question of susceptibility loci for an even broader phenotype of IGE, and other studies have attempted to identify these genes (Sander, Schulz et al. 2000; Sander, Windemuth et al. 2003).

## 16.3 Patient recruitment

A problem frequently encountered in genetic studies of complex diseases is recruitment of sufficient patients, particularly when the genetic architecture of the condition is not known. This project is a good example of international collaboration, with families recruited from both Northern Europe and Canada. As far as possible the families were limited to those of Northern European origin (allowing for the inevitable uncertainty and complexity of the geographical origin of many families), in order to provide a degree of genetic homogeneity. The potential importance of this has been exemplified by the report of putative mutations in the T-type calcium channel gene *CACNA1H* in childhood absence epilepsy patients of Han ethnicity recruited from North China (Chen, Lu et al. 2003). Our study found no evidence of linkage to *CACNA1H* and subsequent sequence analysis in 220 of these European patients failed to identify any of the Chinese variations (Chioza, Everett et al. 2006). A study of CAE patients in Australia also failed to replicate the findings of the Chinese study (Heron, Phillips et al. 2004). Further screening of *CACNA1H* in epilepsy patients, predominantly white, with diagnoses including CAE, did detect some potential functional variants, but all variants were also identified in unaffected individuals (Heron, Khosravani et al. 2007). Genetic heterogeneity between different populations is one explanation for the variability in replication of the Chinese study, and we have attempted to reduce this factor by restricting the geographical origin of our patient population.

Recruitment of patients for a genetic study entails important ethical considerations which have implications for patient confidentiality and informed consent. These include the process of identifying suitable patients for invitation to participate in the study; the collection of DNA samples by invasive techniques (e.g. venesection) in children too young to give consent; the implications of identifying putative susceptibility alleles in asymptomatic family members and, by inference, in non-participating relatives; and the potential use of DNA samples for future studies. These issues were addressed in the application for ethical approval to the North Thames Multi-

centre Research Ethics Committee, and, when relevant, were discussed with participating families during the consenting process.

## 16.4 Laboratory methodology

DNA has traditionally been extracted from venous blood, which provides high yields of good quality DNA. Despite this, after genotyping numerous markers and sequencing large sections of DNA, stores of DNA can become very low. Also, obtaining blood samples from children is often time consuming, technically difficult, distressing to the child, unacceptable to the parents and difficult to justify for the purpose of genetic research. The use of buccal swabs to obtain specimens avoids these difficulties and samples can be collected by post. The quantity of DNA obtained from cheek swabs is much lower than that obtained from blood. Whole genome amplification by multiple displacement amplification provided a quick, easy and reliable method of increasing DNA stores (Dean, Hosono et al. 2002). Therefore in the latter stages of patient recruitment, cheek swabs became the method of choice for DNA collection and this was much more acceptable to parents of young children.

The selection of polymorphic markers for testing of candidate genes by linkage analysis has been greatly facilitated by the completion of the human genome project and the publication of large-scale genome wide linkage maps (Dib, Faure et al. 1996; Kong, Gudbjartsson et al. 2002; Kong, Murphy et al. 2004). In this study, this has enabled the identification of markers flanking or intragenic to the candidate genes under investigation with data on both physical and genetic map distances. In only a minority of cases physical map distances were used to estimate linkage map distances. The availability of sequence information also allowed three novel repeat markers to be identified in the vicinity of two candidate genes (*CACNG3* and *CACNA1G*).

A wide variety of SNP genotyping methods are now available, as described in sections 4.3.4-4.3.9 and section 6.7. Melting curve analysis of SNPs (McSNP®) (6.7.1) was initially selected for this project. The combination of a tried and tested approach to allele discrimination, restriction enzyme digestion, with a novel method for detecting DNA fragments, melting curve analysis, promised a low-cost, accurate, high-throughput genotyping method (Akey, Sosnoski et al. 2001; Ye, Parra et al. 2002). However attempts to establish a McSNP facility in collaboration with the London School of Hygiene and Tropical Medicine was beset by a range of problems, as described in section 6.7.1. These were predominantly failures of equipment and software beyond our control, which exemplifies the difficulties in establishing a new technical facility from scratch when funds are limited. Ultimately it was decided to use established facilities providing TaqMan® (section 1.1.1), Pyrosequencing® (section 6.7.3) and Amplifluor® (section 4.3.8). Pyrosequencing was particularly suited for validation of novel SNPs in a panel of control DNAs.

## 16.5 Candidate gene selection

Candidate genes can be identified for many ‘complex’ traits, on both functional grounds, and by the segregation of many of these traits in a monogenic fashion in rare extended multigenerational pedigrees. The genes identified in such rare subtypes are also candidates for ‘susceptibility’ sequence variants in the ‘complex’ trait. However, although many biologically plausible candidate genes exist for any trait, it is only a small minority of these in which sequence variants will actually exist in the population at allele frequencies high enough to account for common traits. Also, many genes causing monogenic subtypes of ‘complex’ traits have been tested as susceptibility loci in the more common traits, but with little success.

The absence epilepsies should have a particular advantage as a phenotype for genetic analysis. The molecular neurophysiological basis of spike-wave discharges, the neurophysiological hallmark of these epilepsies, has been

extensively investigated in rodents and humans. There is now strong evidence that the generation of spike-wave seizures is determined, in part, by interactions between voltage-dependent calcium channels and GABA receptors, and that these genes are implicated in their aetiology (section 3.5.1). Of particular interest are the four voltage-dependant calcium channel subunit genes identified in mouse mutants with spike-wave seizures, *Cacna1a, Cacnb4, Cacng2 and Cacna2d2* (section 3.5.2). Of these, *CACNA1A* and *CACNB4* have previously been implicated in human epilepsies. $GABA_A$ receptors mutations have also been identified in both rodent and human spike-wave epilepsies, and are implicated in Angelman Syndrome (section 7.3). In contrast, the presence of functional $GABA_B$ receptors appears to be necessary for the generation of absence seizures in mice (section 7.4). Further supportive evidence for a role for several of these genes in spike-wave epilepsy is provided by previous linkage and association studies, expression patterns in brain and mechanisms of action of antiepileptic drugs, as discussed in section 3.5.

## 16.6 Approaches to genetic analysis

The relative power of either parametric or non-parametric linkage analysis or association studies in the investigation of traits with 'complex' inheritance is highly dependent on the underlying genetic architecture, which can only be estimated. Association analysis has greater power to detect loci of minor effect with ascertainable patient numbers provided a high degree of allelic diversity is not present and susceptibility alleles are present at frequencies conferring sufficient power. Linkage is not dependent on allelic homogeneity, but requires sibling pairs or nuclear pedigrees with an adequate proportion of linked families.

Thresholds for statistical significance depend upon the prior probability that one of the markers tested is linked to or associated with the trait, and the influence of multiple testing (Freimer and Sabatti 2004). The traditional stringent threshold for significant linkage (LOD score>3, $p<10^{-4}$) was

proposed when few markers were available and the prior probability that one of these markers was linked to the gene of interest was very small (Morton 1955). With genome wide scans using multiple markers, or candidate gene approaches testing several strong functional candidates, the prior probability that some marker is linked to the locus of interest becomes much greater. However statistical testing of multiple independent markers increases the likelihood of a false positive result, and requires correction for multiple comparisons. Statically approaches using either Bonferroni correction or Gaussian process approximations for linkage statistics have generated very similar LOD score thresholds (3-3.5) and shown that little additional adjustment is needed for very closely spaced markers (where testing of adjacent markers may not be independent as linkage data is effectively replicated). Thus the original criteria for significant linkage remain applicable to this study.

The significance threshold for association studies is more problematic, and there is no consensus. Some authors have suggested significant p-values of lower than ~$10^{-7}$ (Freimer and Sabatti 2005). However this assumes a prior probability of ~1/30000 that a candidate gene is associated with the trait under investigation (ie the gene has been picked at random), which is an extremely conservative, least favourable scenario. The prior probability of association of a trait to a single marker is much lower than the prior probability of linkage, as association extends over much shorter distances than linkage. This factor is reduced by testing only variants in strong candidate genes, rather than a genome-wide association study. The need to correct for multiple comparisons is also reduced by identifying the common SNP haplotypes and testing for association using these rather than individual SNPs. Testing many individual SNPs for association will increase the false positive rate due to multiple testing, unless the SNPs are all in perfect LD with each other. However, where LD is present, corrections for multiple testing would overcorrect for the false positive rate (Nyholt 2004). Methods have been suggested for calculating the effective number of independent tests and correcting for multiple testing (Meng, Zaykin et al. 2003; Nyholt

2004). However these methods have not been validated and remain the subject of debate (Nyholt 2005; Salyakina, Seaman et al. 2005).

Significance levels for two-locus LOD scores are even harder to assess, particularly when maximised over disease-model parameters. Significance again depends on assumptions about prior probability of linkage to the two loci. Inflation of two-locus LOD scores should be expected when compared to the single locus results. If maximized over two recombination fractions, Strauch et al have calculated that a single-locus LOD score of 3 corresponds to a two-locus LOD score of 3.5 (Strauch, Fimmers et al. 2000).

## 16.7 Results of preliminary candidate gene analysis

Thirty candidate gene loci were tested in the initial cohort of 33 absence families. The negative LOD scores obtained under locus homogeneity provided strong evidence that 26 gene loci do not account independently for the trait in a majority of the families (sixteen VDCC subunit genes, seven $GABA_AR$ subunit genes, two $GABA_B$ receptor genes and the *ECA1* locus).

*CACNG3*, the VDCC on 16p12-p13.1 and the cluster of $GABA_AR$ subunits (*GABRA5, GABRB3, GABRG3*) on chromosome 15q11-q13 could not be excluded from linkage in a proportion of families. Although both loci gave negative homogeneity LODs, the HLODS and NPL scores were positive. The data did not therefore exclude the possibility that sequence variants in *CACNG3* or in any of the $GABA_A$ receptor subunit genes in the 15q cluster may contribute to phenotype in a proportion of families. These loci were therefore tested in the larger collection of CAE families, AE families and parent-child trios.

## *16.8 CACNG3* analysis

Both parametric and non-parametric linkage analysis provided significant evidence for linkage, indicating *CACNG3* is a susceptibility loci for Absence in a subset of the 82 nuclear pedigrees analysed. Analysis of individual SSLPs using the TDT and PDT produced significant associations. Sequencing of coding regions did not identify any plausible causal sequence variants.

Variation in non-protein coding DNA sequences can disrupt normal gene function by several mechanisms. For example, gene expression is controlled by transcription regulatory elements which can be intronic or reside up- or downstream of the transcription unit (Kleinjan and van Heyningen 2005). Similarly, non-coding splicing regulatory elements are important in preventing exon skipping, intron retention or creation of aberrant splice sites (Baralle and Baralle 2005). Therefore several methods were employed to identify potential non-coding causative alleles. Variants in non-coding sequence with possible functional effects were identified, including two potential splice site variants and six ESE motif variants. However identification of functional variants in regulatory elements by sequence inspection is difficult and thus searching for pathogenic variations in these sequences requires an indirect approach.

The approach adopted for further analysis of *CACNG3* comprised (i) identifying commons SNPs spanning the gene, (ii) establishing the pattern of LD across *CACNG3* and (iii) performing association analysis with both individual SNPs and SNP haplotypes.

### 16.8.1 *CACNG3* LD block structure and tag SNPs

The NCBI SNP database allowed selection of SNPs spanning the gene. One additional novel SNP was included in the LD analysis. At the time of selection, all but four of the NCBI SNPs had minor allele frequency data available. The missing allele frequency data was obtained by typing the SNPs in a control population using pyrosequencing. Analysis of LD based on

the entire resource of pedigrees and trios identified 5 LD blocks. In comparison, the HapMap data using a much larger number of SNPs (233) identified 21 blocks of LD across the same region. Several small LD blocks have been missed by using a restricted number of SNPs, whilst other neighbouring LD blocks have been incorporated in a single larger block. Inevitably, typing a greater numbers of SNP reveals a more detailed pattern of LD across the region. However, for the purposes of association analysis, the aim is to identify large LD blocks for which a subset of SNPs will represent the majority of the genetic variation by tagging the common haplotypes. It has been shown that analysis using incomplete, more spaced out sets of SNPs can be almost as powerful as SNPs selected from complete panels, and that tests using exhaustive multimarker tests improves power to detect less common causal alleles but can actually reduce power when the causal SNP is common (de Bakker, Yelensky et al. 2005). The likelihood that a significant positive association will be missed by failing to identify detailed LD structure is small.

The LD block structure generated by a small subset of 31 trios was also compared with that generated by the entire resource. There was some variation in the boundary SNPs of two of the five LD blocks. However the tagging SNPs identified by either of the two methodologies (haplotype frequency and tagger) were similar. In future larger studies investigating other chromosomal regions, it would be a reasonable strategy to use a subset of trios to identify the LD block structure and tagging SNPs, and only to type the tagging SNPs in the entire resource. For the purposes of this study, as the selection of tag SNPs did not significantly reduce the total number of SNPs originally selected, it was decided to type all 23 SNPs in all individuals.

The HapMap SNP data generated a larger number of tag SNPs (85 by haplotype frequency, 76 by tagger), reflecting the more detailed pattern of LD. However, as discussed above, the additional tag SNPs are unlikely to reveal a significant positive association with a common causal allele that has been missed by the restricted set. The HapMap data was obviously not

generated from this study population. However it has now been demonstrated that tag SNPs identified from HapMap population data can be transferred to different study populations with little loss in power to detect an association (de Bakker, Burtt et al. 2006).

### 16.8.2 *CACNG3* association analysis

Association analysis of *CACNG3* was performed using individual SNPs and SNP haplotypes. This strategy relied on the assumption that any causal SNPs were of high enough allele frequency and were in sufficiently tight LD with the typed SNPs to be detected. The likelihood of a false negative result becomes significant if these assumptions are not valid (de Bakker, Yelensky et al. 2005). There also remains the possibility that a causal variant might lie outside the associated interval, which can be difficult to define.

Using the narrow phenotype of CAE, individual SNP analysis identified two SNPs in LD block 1, both in intron 1, which showed significant transmission disequilibrium. The broader 'absence' phenotype also identified two intron 1 SNPs with significant transmission disequilibrium, as well as one SNP upstream of exon 1. All three SNPs were also in LD block 1. The 'sliding window' approach identified nine haplotypes (using narrow CAE phenotype) and 13 haplotypes (using broader 'absence' phenotype) comprising SNPs 2-8 within LD block 1 that demonstrated transmission disequilibrium. These haplotypes together form an extended haplotype of alleles 2211122. The only SNP with potential functional implications is SNP 5 (rs1494550), in which the minor allele is predicted to create a splice acceptor site. However, individually this SNP does not show significant transmission disequilibrium.

Haplotypes were also identified by the sliding window approach in LD block 2 that demonstrated transmission disequilibrium. These did not form a single extended phenotype, either for the narrow CAE phenotype or the broader 'absence' phenotype, and none of the SNPs could be implicated on functional grounds.

These positive associations can only be considered as provisional due to the limitations discussed in section 16.6, including issues around multiple testing. Independent replication of these results is required.

### 16.8.3 *CACNG3* sequence analysis

Direct sequencing of exonic DNA and surrounding intronic and promoter sequence DNA of *CACNG3* in 73 individuals identified 34 sequence variants. Of these, eight variants showed possible functional significance on the basis of bioinformatics analysis, and four of these (rs1494550, s5, s6, s7).are located in LD block 1. Mean allele frequency data in a control population, from the NCBI SNP database, is only available for one of these, rs1494550. The difference in MAF of the reference allele between cases and controls (0.340 vs 0.209) is suggestive, although rs1494550 did not show significant transmission disequilibrium in the study population. If the difference between case and control MAF is real, the variant allele, predicted to create a splice acceptor site, is associated with an increased risk of CAE. However rs1494550 is nongenic and lies 270 bases upstream of the *CACNG3* 5'UTR. It is still possible that rs1494550 is exerting a regulatory effect on *CACNG3*, similar to that described in alpha-thalassaemia (De Gobbi, Viprakasit et al. 2006). Alternatively rs1494550 does not have a functional effect but is in LD with an unidentified causal variant.

The other potentially functional variants within LD block1, s5, s6 and s7, all lie within the *CACNG3* 5' UTR, and are predicted to alter splicing enhancer binding motifs. S7 is conserved in the chimpanzee. However all three variants were found at very low frequencies in the sequenced cases.

Further analysis of the region comprising LD block 1 has now been performed (Everett, Chioza et al. 2007). Re-sequencing of 35 kb of genomic sequence in 48 chromosomes identified a total of 72 sequence variants. One additional variant, rs2021512, has shown significant transmission disequilibrium when typed in the entire patient resource (p=0.005), with overtransmission of the reference allele. This variant creates a novel

acceptor site. However it is also nongenic and lies approximately 14kb upstream of *CACNG3*.

### 16.8.4 *CACNG3* – interpretation of results

Several explanations for the observed results can be postulated.

**i. The observed linkage is a false positive and *CACNG3* is not a susceptibility locus for CAE.** However, the employed criteria for significant linkage are fairly stringent.

**ii. The observed transmission disequilibrium is also a false positive.** The TDT and PDT were used to reduce this likelihood as the results are unaffected by population stratification. However, significance thresholds remain contentious, as discussed in section 16.6

**iii. The linkage and association results are real, but the causal variants have not been identified by the current sequence analysis as they lie outside the sequenced region.** LD has now been shown to exist over long ranges, sometimes hundreds of kilobases. Within these regions of long range LD, clusters of genetically indistinguishable SNPs (giSNPs) exist, which are in perfect LD with each other (Lawrence, Evans et al. 2005). Approximately 50% of SNPs appear to have at least one giSNP.

**iv. The linkage and association results are real, but the causal variants are too rare or heterogeneous to have been detected in the limited number of sequenced chromosomes.**

**v. Causal variants for CAE have been identified, but the functional consequences of these variants are yet to be demonstrated.**

## 16.9 $GABA_AR$ gene cluster on 15q

The parametric linkage analysis for the $GABA_AR$ cluster did not reach accepted criteria for significant linkage. However, the maximum HLOD of 2.31 at marker *155CA2*, the broader 'Absence' phenotype, could be considered 'suggestive' of linkage. The non-parametric linkage analysis, generated a maximum NPL score of 3.75 (p=0.00004), which is a significant result. This may be a reflection of the fact that a susceptibility locus exists in this region, but does not conform to the genetic model specified in the parametric linkage analysis. Absence epilepsy is inherited in a complex 'non-mendelian' manner, and therefore the application of a strictly mendelian model to any susceptibility locus is likely to be an approximation.

The maximum HLOD and NPL scores were generated at marker *155CA2*. The transmission disequilibrium test also identified one allele of *155CA2* which may be associated with a protective effect ($\chi^2$=7.36, p=0.007), although this applies to the narrower CAE phenotype. *155CA2* is intragenic to *GABRB3,* the gene contained within the Angelman Syndrome deletion region that has been shown to produces EEG abnormalities and seizures in mice. One study has identified a *GABRB3* promoter haplotype associated with childhood absence epilepsy (p=0.007075) (Urak, Feucht et al. 2006). A functional *GABRB3* exon 1a promoter polymorphism was identified which displayed a reduced transcriptional activity. However, a subsequent larger study with 250 CAE patients failed to replicate an association of the promoter polymorphism with CAE (Hempelmann, Cobilanschi et al. 2007). Heterozygous mutations in *GABRB3* have now been identified in four out of 48 families with CAE from Mexico and Honduras, and all mutations demonstrated hyperglycosylation and reduced whole cell GABA-evoked current density when expressed in vitro (Tanaka, Olsen et al. 2008). Further investigation of *GABRB3* as well as *GABRA5* and *GABRG3* will require sequencing across the whole genomic region and further association analysis using SNP haplotypes.

## 16.10 Multilocus analysis

The *CACNG3* locus on chromosome 16p12-p13.1 and the *$GABA_AR$* subunit cluster on chromosome 15q11-13 both showed evidence of linkage to the Absence and CAE phenotypes. Absence epilepsy is likely to be a multigenic trait, and it was therefore decided to perform a two-trait-locus study by testing both loci simultaneously with GENEHUNTER-TWOLOCUS(Strauch, Fimmers et al. 2000). The non-parametric scores increased for both phenotypic classifications, supporting the hypothesis that both chromosomal loci harbour susceptibility alleles for the Absence phenotype. A digenic dominant effect was also tested using three parametric models. The multiplicative model produced the greatest increase in the HLOD when compared to the single locus analysis, particularly for the broader Absence phenotype. This provides an insight into how the two loci may act together. The multiplicative model reflects that individuals with a susceptibility allele at both loci have a high probability of manifesting the trait. The results therefore support the hypothesis that susceptibility alleles at the loci on chromosomes 15 and 16 may both be required to develop the disease. In contrast, the heterogeneity model produced maximum HLODs that were lower than the single locus HLODs at *CACNG3*. This rejects the hypothesis that a mutant allele at a single locus (the chromosome 15 locus) alone is sufficient to cause the disease. These results suggest that both *CACNG3* on chromosome 16 and at least one of GABRB3, *GABRA5* and *GABRG3* on chromosome 15 are involved in the pathogenesis of absence epilepsy.

## 16.11 Limitations of this study

Although attempts were made to be as accurate as possible when phenotyping patients, in some cases limited data was available, particularly for those patients from outside the UK. When there was significant doubt about the diagnosis, patients were classified as unknown. In some cases, where all available data was consistent, it seemed reasonable to accept the

diagnosis of an experienced collaborator. Inevitably it remains possible that a small number of patients were wrongly classified.

The use of buccal swabs for DNA collection was clearly more acceptable and convenient for families than blood sampling. The reduced quantity and quality of DNA obtained, in addition to the use of whole genome amplification to increase DNA stores, may have introduced some errors into the gentotyping and DNA sequencing.

The linkage analysis suggested a role for both *CACNG3* and the $GABA_AR$ cluster on 15q in the pathogenesis of absence epilepsy. However it was not possible to pursue the 15q linkage with further analysis due to the limited time and resources available.

Perhaps the major limitation of this study is the lack of functional work with in-vitro systems. Again this was not possible due to time limitation, but this will be required in future to determine whether causal variants for CAE have been identified.

## 16.12 Future directions for research

As discussed above, in-vitro analysis of implicated *CACNG3* variants is needed for futher evaluation of their functional significance. This can be performed by expression of variant *CACNG3* cDNA in *xenopus* oocytes or HEK cells and assessing channel function by patch-clamp current measurement. Membrane expression of channels can also be assessed by fluoresecent labelling of cDNA.

Causal variants in *CACNG3* may have not been identified by the current sequencing. Extended sequencing to include intronic and other potential regulatory regions may reveal additional plausible functional variants

The 15q $GABA_AR$ warrants further investigation by SNP genotyping, association analysis and DNA sequencing.

It is certain that other genes are also susceptibility loci for absence epilepsy, and these may be identified by both candidate gene analysis and genome-wide linkage and association analysis. Improved high-throughput genotyping and analytical technologies, as well as advances in bioinformatics, will increase the likelihood of detecting true causal variants with only modest genetic effects. Positive findings must be replicated in independent patient groups.

Whatever genetic techniques are employed, access to additional large, well characterised patient cohorts is vital. There is therefore a need for ongoing patient recruitment and accurate phenotypying, which requires multi-centre collaborations and the active involvement of clinicians.

## 16.13 Conclusions

This work has provided genetic evidence that *CACNG3* and at least one of the three $GABA_A$ receptor genes of *GABRB3*, *GABRA5* and *GABRG3* are susceptibility loci for absence epilepsy, including CAE. These genes are strong functional candidates for absence epilepsy. The data supports a digenic effect from the chromosome 16 and 15 loci, suggesting that susceptibility alleles at both loci may be required to manifest the trait. Common variants at the *CACNG3* locus showing transmission disequilibrium have been identified, although no variants as yet display functional significance. Replication studies in similar patient groups would help to confirm that these loci contribute to the absence phenotype. Definitive evidence to confirm or exclude these loci will require sequencing across extended genomic regions encompassing *CACNG3* and the $GABA_AR$ cluster in large numbers of patients, in addition to functional analysis of potential causative variants.

# *Chapter 17 Genetic analysis of a family with absence epilepsy and episodic ataxia*

A syndrome of absence epilepsy and episodic ataxia was first characterised in 2001. Jouvenceau et al reported an 11 year old boy with primary generalised epilepsy, episodic and progressive ataxia, and mild learning difficulties (Jouvenceau, Eunson et al. 2001). From the age of 3 years he developed nocturnal generalised tonic-clonic seizures and daytime absence seizures. The generalised tonic-clonic seizures stopped after age 8 years but he continued to have absence seizures. An interictal EEG showed generalised polyspike and wave complexes. From age 8 years he also developed attacks of gait unsteadiness, dysarthria, and diplopia, lasting up to three hours. There was no family history of epilepsy or episodic ataxia. Examination between attacks showed cerebellar eye signs gait and limb ataxia. Sequencing of the coding region of *CACNA1A*, the gene encoding the voltage-gated P/Q type calcium channel, identified a heterozygous point mutation (C5733T). This mutation introduces a premature stop codon which was shown to impair calcium channel function and exhibit a dominant negative effect.

Mutations in *CACNA1A* are known to cause a range of neurological disorders in humans (familial hemiplegic migraine (Ophoff, Terwindt et al. 1996), episodic ataxia type 2 (Ophoff, Terwindt et al. 1996), spinocerebellar ataxia type 6 (Zhuchenko, Bailey et al. 1997)). The tottering mouse mutant, which exhibits spike-wave seizures similar to those occurring in human absence epilepsy, is caused by mutations in *Cacna1a* (Fletcher, Lutz et al. 1996).

*CACNA1A* was therefore an excellent candidate gene in the family identified by Dr Jaffe, in which absence epilepsy segregates with episodic ataxia type 2 (EA2) in an autosomal dominant fashion through three generations (Section 5.9).

The linkage analysis and pattern of segregation of markers spanning *CACNA1A* were strongly suggestive that *CACNA1A* was the causative gene, and this was confirmed by the sequence analysis. The one family member

with spike-wave discharges on EEG but no clinical seizures or ataxia did not carry the “disease” haplotype or the *CACNA1A* mutation. Repeat blood samples were obtained from this patient to ensure there was no sample error. It can be argued that the spike-wave discharges in this individual, as well as all the affected family members, are caused by mutations in another gene. However, if this were the case, the putative unidentified mutation must be insufficient to cause clinical seizures and/or ataxia. Also, for the clinical phenotype of EAE to have been caused by a digenic effect from the *CACNA1A* mutation and this “second” mutation, the segregation pattern for both variants in this family would have to be identical, diverging only at individual 319. Unless the two loci were in tight linkage (itself highly unlikely), the likelihood of this identical segregation occurring by chance is very small, as there are 512 ($2^9$) different ways that each allele inherited from individual 102 can segregate in the rest of the family.

This is the first reported family in which a *CACNA1A* mutation that impairs calcium channel function cosegregates with typical absence seizures and 3Hz spike-wave discharges on EEG. The ataxia exhibited by affected members was very variable in both severity and episodic nature. In some cases, cerebellar ataxia was precipitated by AEDs such as phenytoin at therapeutic doses. Other individuals developed intermittent ataxia typical of EA2.

The previously reported patient with absence epilepsy and episodic ataxia attributed to a *CACNA1A* mutation had a relatively severe phenotype with refractory seizures and progressive ataxia (Jouvenceau, Eunson et al. 2001). The mutation identified in this case was a premature stop codon leading to complete loss of channel function and a dominant negative effect. The present family members exhibit a much milder phenotype with regard to both epilepsy and ataxia. Imbrici et al have tested the missense mutation identified in this family by functional expression studies in in *Xenopus* oocytes using human *CACNA1A* cDNA (Imbrici, Jaffe et al. 2004). E147K alters a transmembrane segment and has a relatively subtle effect on channel function. The partial reduction in calcium channel function appears to be

caused by impaired trafficking to the cell membrane, and overexpression of accessory subunits can restore function. It has been postulated that the ataxia may be a direct consequence of the reduced calcium channel current density, whilst the absence seizures may be caused by a compensatory increase in currents mediated by other calcium channels. The latter is supported by the finding of increased low threshold calcium currents mediated by t-type calcium channels in mice with spike-wave seizures and mutations in *Cacna1a* (Zhang, Mori et al. 2002). As discussed in section 7.2.6, nonsynonymous SNPs in *CACNA1H*, which encodes a t-type calcium channel, have been found in patients with CAE (Chen, Lu et al. 2003; Heron, Khosravani et al. 2007). Functional characterisation of some of these SNPs in rat channels demonstrated a gain of function, with increased calcium influx during physiological activation and an increased propensity to channel opening (Khosravani, Altier et al. 2004).

This family has therefore provided additional supportive evidence for a role for *CACNA1A* mutations in some cases of human absence epilepsy. Whether *CACNA1A* contributes to a significant proportion of CAE cases remains uncertain. A significant association has been demonstrated between a *CACNA1A* intragenic SNP and patients with IGE (Chioza, Wilkie et al. 2001). Although *CACNA1A* did not show positive linkage in our preliminary analysis of 33 families, this may reflect the limitations of this methodology in detecting one of perhaps several genes incorporating susceptibility alleles for human absence epilepsy.

## References


*Adams, M. D., S. E. Celniker, R. A. Holt, et al. (2000). "The genome sequence of Drosophila melanogaster." Science **287**(5461): 2185-95.*

*Akey, J. M., D. Sosnoski, E. Parra, et al. (2001). "Melting curve analysis of SNPs (McSNP): a gel-free and inexpensive approach for SNP genotyping." Biotechniques **30**(2): 358-62, 364, 366-7.*

*Annegers, J. F., W. A. Rocca and W. A. Hauser (1996). "Causes of epilepsy: contributions of the Rochester epidemiology project." Mayo Clin Proc **71**(6): 570-5.*

*Appleton, R. E. (1995). Eyelid myoclonia with absences. Typical absences and related epileptic syndromes. J. S. Duncan and C. P. Panayiotopoulos. London, Churchill Communications Europe.*

*Aridon, P., C. Marini, C. Di Resta, et al. (2006). "Increased sensitivity of the neuronal nicotinic receptor alpha 2 subunit causes familial epilepsy with nocturnal wandering and ictal fear." Am J Hum Genet **79**(2): 342-50.*

*Audenaert, D., E. Schwartz, K. G. Claeys, et al. (2006). "A novel GABRG2 mutation associated with febrile seizures." Neurology **67**(4): 687-90.*

*Avanzini, G. and S. Franceschetti (2003). "Cellular biology of epileptogenesis." Lancet Neurol **2**(1): 33-42.*

*Bahi-Buisson, N., A. Kaminska, N. Boddaert, et al. (2008). "The three stages of epilepsy in patients with CDKL5 mutations." Epilepsia **49**(6): 1027-37.*

*Bai, D., M. E. Alonso, M. T. Medina, et al. (2002). "Juvenile myoclonic epilepsy: linkage to chromosome 6p12 in Mexico families." Am J Med Genet **113**(3): 268-74.*

*Baralle, D. and M. Baralle (2005). "Splicing in action: assessing disease causing sequence changes." J Med Genet **42**(10): 737-48.*

*Barclay, J., N. Balaguero, M. Mione, et al. (2001). "Ducky mouse phenotype of epilepsy and ataxia is associated with mutations in the Cacna2d2 gene and decreased calcium channel current in cerebellar Purkinje cells." J Neurosci **21**(16): 6095-104.*

*Barrett, J. C., B. Fry, J. Maller, et al. (2005). "Haploview: analysis and visualization of LD and haplotype maps." Bioinformatics **21**(2): 263-5.*

*Baulac, S., G. Huberfeld, I. Gourfinkel-An, et al. (2001). "First genetic evidence of GABA(A) receptor dysfunction in epilepsy: a mutation in the gamma2-subunit gene." Nat Genet **28**(1): 46-8.*

*Beaussart, M. (1972). "Benign epilepsy of children with Rolandic (centro-temporal) paroxysmal foci. A clinical entity. Study of 221 cases." Epilepsia **13**(6): 795-911.*

*Beck-Mannagetta, G. and D. Janz (1991). "Syndrome-related genetics in generalised epilepsy." Epilepsy Research **Suppl. 4**: 105-111.*

*Berkovic, S. F., S. E. Heron, L. Giordano, et al. (2004). "Benign familial neonatal-infantile seizures: characterization of a new sodium channelopathy." Ann Neurol **55**(4): 550-7.*

*Berkovic, S. F., R. A. Howell, D. A. Hay, et al. (1998). "Epilepsies in twins: genetics of the major epilepsy syndromes." Ann Neurol **43**(4): 435-45.*

*Berkovic, S. F., P. Izzillo, J. M. McMahon, et al. (2004). "LGI1 mutations in temporal lobe epilepsies." Neurology **62**(7): 1115-9.*

*Bettler, B., K. Kaupmann, J. Mosbacher, et al. (2004). "Molecular structure and physiological functions of GABA(B) receptors." Physiol Rev **84**(3): 835-67.*

*Bienvenu, T., K. Poirier, G. Friocourt, et al. (2002). "ARX, a novel Prd-class-homeobox gene highly expressed in the telencephalon, is mutated in X-linked mental retardation." Hum Mol Genet **11**(8): 981-91.*

*Black, J. L., 3rd (2003). "The voltage-gated calcium channel gamma subunits: a review of the literature." J Bioenerg Biomembr **35**(6): 649-60.*

*Blom, S. and J. Heijbel (1975). "Benign epilepsy of children with centro-temporal EEG foci. Discharge rate during sleep." Epilepsia **16**(1): 133-40.*

*Blume, W. T., H. O. Luders, E. Mizrahi, et al. (2001). "Glossary of descriptive terminology for ictal semiology: report of the ILAE task force on classification and terminology." Epilepsia **42**(9): 1212-8.*

*Botstein, D. and N. Risch (2003). "Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease." Nat Genet **33 Suppl**: 228-37.*

*Botstein, D., R. L. White, M. Skolnick, et al. (1980). "Construction of a genetic linkage map in man using restriction fragment length polymorphisms." Am J Hum Genet **32**(3): 314-31.*

*Bray, P. F. and W. C. Wiser (1965). "Hereditary characteristics of familial temporal-central focal epilepsy." Paediatrics **36**: 207-211.*

*Bray, P. F. and W. C. Wiser (1969). "A unifying concept of idiopathic epilepsy." Postgrad Med **46**(1): 82-7.*

*Breustedt, J., K. E. Vogt, R. J. Miller, et al. (2003). "Alpha1E-containing Ca2+ channels are involved in synaptic plasticity." Proc Natl Acad Sci U S A **100**(21): 12450-5.*

*Brill, J., R. Klocke, D. Paul, et al. (2004). "entla, a novel epileptic and ataxic Cacna2d2 mutant of the mouse." J Biol Chem **279**(8): 7322-30.*

*Bruyere, H., S. Lewis, S. Wood, et al. (1999). "Confirmation of linkage in X-linked infantile spasms (West syndrome) and refinement of the disease locus to Xp21.3-Xp22.1." Clin Genet **55**(3): 173-81.*

*Bulman, D. E. (1997). "Phenotype variation and newcomers in ion channel disorders." Hum Mol Genet **6**(10): 1679-85.*

*Buoni, S., S. Grosso, G. Di Cosmo, et al. (1998). "Segregation analysis in typical absence epilepsy." J Child Neurol **13**(2): 89-93.*

*Burgess, D. L., J. M. Jones, M. H. Meisler, et al. (1997). "Mutation of the $Ca^{2+}$ channel $\beta$ subunit gene Cchb4 is associated with ataxia and seizures in the lethargic (lh) mouse." Cell **88**(3): 385-392.*

*Calado, A., F. M. Tome, B. Brais, et al. (2000). "Nuclear inclusions in oculopharyngeal muscular dystrophy consist of poly(A) binding protein 2 aggregates which sequester poly(A) RNA." Hum Mol Genet **9**(15): 2321-8.*

*Caraballo, R., S. Pavek, A. Lemainque, et al. (2001). "Linkage of benign familial infantile convulsions to chromosome 16p12- q12 suggests allelism to the infantile convulsions and choreoathetosis syndrome." Am J Hum Genet **68**(3): 788-94.*

*Carlton, J. M., S. V. Angiuoli, B. B. Suh, et al. (2002). "Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii." Nature **419**(6906): 512-9.*

*Cartegni, L., J. Wang, Z. Zhu, et al. (2003). "ESEfinder: A web resource to identify exonic splicing enhancers." Nucleic Acids Res **31**(13): 3568-71.*

*Cavazzuti, G. B. (1980). "Epidemiology of different types of epilepsy in school age children of Modena, Italy." Epilepsia **21**(1): 57-62.*

*Charlier, C., N. A. Singh, S. G. Ryan, et al. (1998). "A pore mutation in a novel KQT-like potassium channel gene in an idiopathic epilepsy family." Nat Genet **18**(1): 53-5.*

*Chen, L., D. M. Chetkovich, R. S. Petralia, et al. (2000). "Stargazin regulates synaptic targeting of AMPA receptors by two distinct mechanisms." Nature **408**(6815): 936-43.*

*Chen, Y., J. Lu, H. Pan, et al. (2003). "Association between genetic variation of CACNA1H and childhood absence epilepsy." Ann Neurol **54**(2): 239-43.*

*Chioza, B., K. Everett, H. Aschauer, et al. (2006). "Evaluation of CACNA1H in European patients with childhood absence epilepsy." Epilepsy Res **69**(2): 177-81.*

*Chioza, B., H. Wilkie, L. Nashef, et al. (2001). "Association between the alpha(1a) calcium channel gene CACNA1A and idiopathic generalized epilepsy." Neurology **56**(9): 1245-6.*

*Claes, L., J. Del-Favero, B. Ceulemans, et al. (2001). "De novo mutations in the sodium-channel gene SCN1A cause severe myoclonic epilepsy of infancy." Am J Hum Genet **68**(6): 1327-32.*

*Claes, S., K. Devriendt, L. Lagae, et al. (1997). "The X-linked infantile spasms syndrome (MIM 308350) maps to Xp11.4-Xpter in two pedigrees." Ann Neurol **42**(3): 360-4.*

*Collins, F. S., E. D. Green, A. E. Guttmacher, et al. (2003). "A vision for the future of genomics research." Nature **422**(6934): 835-47.*

*Concolino, D., M. A. Iembo, E. Rossi, et al. (2002). "Familial pericentric inversion of chromosome 5 in a family with benign neonatal convulsions." J Med Genet **39**(3): 214-6.*

*C. Elegans Sequencing Consortium. (1998). "Genome sequence of the nematode C. elegans: a platform for investigating biology." Science **282**(5396): 2012-8.*

*Cossette, P., L. Liu, K. Brisebois, et al. (2002). "Mutation of GABRA1 in an autosomal dominant form of juvenile myoclonic epilepsy." Nat Genet **31**(2): 184-9.*

*Coulter, D. A., J. R. Huguenard and D. A. Prince (1989). "Characterization of ethosuximide reduction of low-threshold calcium current in thalamic neurons." Ann Neurol **25**(6): 582-93.*

*Coulter, D. A., J. R. Huguenard and D. A. Prince (1990). "Differential effects of petit mal anticonvulsants and convulsants on thalamic neurones: calcium current reduction." Br J Pharmacol **100**(4): 800-6.*

*Cox, G. A., C. M. Lutz, C. L. Yang, et al. (1997). "Sodium/hydrogen exchanger gene defect in slow-wave epilepsy mutant mice [published erratum appears in Cell 1997 Dec 12;91(6):861]." Cell **91**(1): 139-48.*

*Crunelli, V. and N. Leresche (2002). "Childhood absence epilepsy: genes, channels, neurons and networks." Nat Rev Neurosci **3**(5): 371-82.*

*Currier, R. D., K. A. Kooi and L. J. Saidman (1963). "Prognosis of "Pure" Petit Mal; a Follow-up Study." Neurology **13**: 959-67.*

*Daly, M. J., J. D. Rioux, S. F. Schaffner, et al. (2001). "High-resolution haplotype structure in the human genome." Nat Genet **29**(2): 229-32.*

*de Bakker, P. I., N. P. Burtt, R. R. Graham, et al. (2006). "Transferability of tag SNPs in genetic association studies in multiple populations." Nat Genet **38**(11): 1298-303.*

*de Bakker, P. I., R. Yelensky, I. Pe'er, et al. (2005). "Efficiency and power in genetic association studies." Nat Genet **37**(11): 1217-23.*

*De Fusco, M., A. Becchetti, A. Patrignani, et al. (2000). "The nicotinic receptor beta 2 subunit is mutant in nocturnal frontal lobe epilepsy." Nat Genet **26**(3): 275-6.*

*De Gobbi, M., V. Viprakasit, J. R. Hughes, et al. (2006). "A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter." Science **312**(5777): 1215-7.*

*Dean, F. B., S. Hosono, L. Fang, et al. (2002). "Comprehensive human genome amplification using multiple displacement amplification." Proc Natl Acad Sci U S A **99**(8): 5261-6.*

*Dedek, K., B. Kunath, C. Kananura, et al. (2001). "Myokymia and neonatal epilepsy caused by a mutation in the voltage sensor of the KCNQ2 K+ channel." Proc Natl Acad Sci U S A **98**(21): 12272-7.*

*Degen, R. and H. E. Degen (1990). "Some genetic aspects of rolandic epilepsy: waking and sleeping EEGs in siblings." Epilepsia **31**: 795-801.*

*Delgado-Escueta, A. V., D. Greenberg and K. Weissbecker (1990). "Gene mapping in the idiopathic generalised epilepsies." Epilepsia **31 (Suppl 3)**: S19-S29.*

*DeLorey, T. M., A. Handforth, S. G. Anagnostaras, et al. (1998). "Mice lacking the beta3 subunit of the GABAA receptor have the epilepsy phenotype and many of the behavioral characteristics of Angelman syndrome." J Neurosci **18**(20): 8505-14.*

*Depienne, C., O. Trouillard, C. Saint-Martin, et al. (2009). "Spectrum of SCN1A gene mutations associated with Dravet syndrome: analysis of 333 patients." J Med Genet **46**(3): 183-91.*

*Dib, C., S. Faure, C. Fizames, et al. (1996). "A comprehensive genetic map of the human genome based on 5,264 microsatellites." Nature **380**: 152-154.*

*Dibbens, L. M., H. J. Feng, M. C. Richards, et al. (2004). "GABRD encoding a protein for extra- or peri-synaptic GABAA receptors is a susceptibility locus for generalized epilepsies." Hum Mol Genet **13**(13): 1315-9.*

*Dolphin, A. C. (2003). "Beta subunits of voltage-gated calcium channels." J Bioenerg Biomembr **35**(6): 599-620.*

*Doose, H., E. Volzke and D. Scheffner (1965). "Course forms of infantile epilepsies with spike waves absences." Arch Psychiatr Nervenkr **207**(4): 394-415.*

*Dravet, C. (1978). "Les épilepsies graves de l'enfant." Vie Med **8**: 543-548.*

*Dulac, O., J. Feingold, P. Plouin, et al. (1993). "Genetic predisposition to West syndrome." Epilepsia **34**(4): 732-7.*

*Duncan, J. S. and C. P. Panayiotopoulos (1992). Juvenile absence epilepsy: an alternative view. Epileptic Syndromes in infancy, childhood and adolescence (2nd Edition). J. Roger, M. Bureau, C. Dravetet al. London. Paris., John Libbey & Co. Ltd.: 307-312.*

*Durner, M., T. Sander, D. A. Greenberg, et al. (1991). "Localisation of idiopathic generalised epilepsy on chromosome 6p in families of juvenile myoclonic epilepsy patients." Neurology **41**(10): 1651-1655.*

*Echenne, B., V. Humbertclaude, F. Rivier, et al. (1994). "Benign infantile epilepsy with autosomal dominant inheritance." Brain Dev **16**(2): 108-11.*

*Elmslie, F. V., M. Rees, M. P. Williamson, et al. (1997). "Genetic mapping of a major susceptibility locus for juvenile myoclonic epilepsy on chromosome 15q." Hum Mol Genet **6**(8): 1329-34.*

*Elmslie, F. V., M. P. Williamson, M. Rees, et al. (1996). "Linkage analysis of juvenile myoclonic epilepsy and microsatellite loci spanning 61 cM of human chromosome 6p in 19 nuclear pedigrees provides no evidence for a susceptibility locus in this region." American Journal of Human Genetics **59**: 653-663.*

*Engel, J. (2001). "A proposed diagnostic scheme for people with epileptic seizures and with epilepsy: report of the ILAE task force on classification and terminology." Epilepsia **42**(6).*

*Engel, J. and T. A. Pedley (1997). Epilepsy : a comprehensive textbook. Philadelphia, Lippincott-Raven.*

*Escayg, A., M. De Waard, D. D. Lee, et al. (2000). "Coding and noncoding variation of the human calcium-channel beta4-subunit gene CACNB4 in patients with idiopathic generalized epilepsy and episodic ataxia." Am J Hum Genet **66**(5): 1531-9.*

*Escayg, A., B. T. MacDonald, M. H. Meisler, et al. (2000). "Mutations of SCN1A, encoding a neuronal sodium channel, in two families with GEFS+2." Nat Genet **24**(4): 343-5.*

*Esteban, J. A., M. Salas and L. Blanco (1993). "Fidelity of phi 29 DNA polymerase. Comparison between protein-primed initiation and DNA polymerization." J Biol Chem **268**(4): 2719-26.*

*Everett, K., B. Chioza, J. Aicardi, et al. (2007). "Linkage and mutational analysis of CLCN2 in childhood absence epilepsy." Epilepsy Res **75**(2-3): 145-53.*

*Everett, K. V., B. Chioza, J. Aicardi, et al. (2007). "Linkage and association analysis of CACNG3 in childhood absence epilepsy." Eur J Hum Genet **15**(4): 463-72.*

*Feinberg, A. P. and W. R. Leahy (1977). "Infantile spasms: case report of sex-linked inheritance." Dev Med Child Neurol **19**(4): 524-6.*

*Feucht, M., K. Fuchs, E. Pichlbauer, et al. (1999). "Possible association between childhood absence epilepsy and the gene encoding GABRB3." Biol Psychiatry **46**(7): 997-1002.*

*Fletcher, C. F., C. M. Lutz, N. O'Sullivan, et al. (1996). Defects in the alpha-1a voltage-sensitive calcium channel in tottering mice cause ataxia, motor seizures and absence epilepsy. Second Biennial Workshop in Mouse Molecular Neurogenetics, The Jackson Laboratory, Bar Harbor, Maine, USA. 18-22 Sept. 1996.*

*Fletcher, C. F., C. M. Lutz, T. N. O'Sullivan, et al. (1996). "Absence epilepsy in tottering mutant mice is associated with calcium channel defects." Cell **87**: 607-617.*

*Fong, G. C., P. U. Shah, M. N. Gee, et al. (1998). "Childhood absence epilepsy with tonic-clonic seizures and electroencephalogram 3-4-Hz spike and multispike-slow wave complexes: linkage to chromosome 8q24." Am J Hum Genet **63**(4): 1117-29.*

*Freimer, N. and C. Sabatti (2004). "The use of pedigree, sib-pair and association studies of common diseases for genetic mapping and epidemiology." Nat Genet **36**(10): 1045-51.*

*Freimer, N. B. and C. Sabatti (2005). "Guidelines for association studies in Human Molecular Genetics." Hum Mol Genet **14**(17): 2481-3.*

*Fujiwara, T., T. Sugawara, E. Mazaki-Miyazaki, et al. (2003). "Mutations of sodium channel alpha subunit type 1 (SCN1A) in intractable childhood epilepsies with frequent generalized tonic-clonic seizures." Brain **126**(Pt 3): 531-46.*

*Fukata, Y., H. Adesnik, T. Iwanaga, et al. (2006). "Epilepsy-related ligand/receptor complex LGI1 and ADAM22 regulate synaptic transmission." Science **313**(5794): 1792-5.*

*Gardiner, R. M. (1995). Genetics of human typical absence syndromes. Typical absences and related epileptic syndromes. J. S. Duncan and C. P. Panayiotopoulos. London, Churchill Communications Europe.*

*Gardner, M. J., N. Hall, E. Fung, et al. (2002). "Genome sequence of the human malaria parasite Plasmodium falciparum." Nature **419**(6906): 498-511.*

*Gastaut, H. (1970). "Clinical and electroencephalographical classification of epileptic seizures." Epilepsia **11**(1): 102-13.*

*Giaretta, D., M. Avoli and P. Gloor (1987). "Intracellular recordings in pericruciate neurons during spike and wave discharges of feline generalized penicillin epilepsy." Brain Res **405**(1): 68-79.*

*Gibbs, F. A., H. Davis and W. G. Lennox (1935). "The EEG in epilepsy and in conditions of impaired consciousness." Arch Neurol Psychiatry(34): 1134-48.*

*Gloor, P. (1978). "Generalized epilepsy with bilateral synchronous spike and wave discharge. New findings concerning its physiological mechanisms." Electroencephalogr Clin Neurophysiol Suppl(34): 245-9.*

*Greenberg, D. A. and A. V. Delgado-Escueta (1993). "The chromosome 6p epilepsy locus: exploring mode of inheritance and heterogeneity through linkage analysis." Epilepsia **34**(Suppl 3): S12-8.*

*Greenberg, D. A., A. V. Delgado-Escueta, H. Widelitz, et al. (1989). "Strengthened evidence for linkage of juvenile myoclonic epilepsy to HLA and BF." Cytogenetics and Cell Genetics **51**: 1008.*

*Greenberg, D. A., A. V. Delgado-Escueta, H. Widelitz, et al. (1988). "Juvenile myoclonic epilepsy may be linked to the BF and HLA loci on human chromosome 6." American Journal of Medical Genetics **31**(1): 185-192.*

*Greenberg, D. A., M. Durner, M. Keddache, et al. (2000). "Reproducibility and complications in gene searches: linkage on chromosome 6,*

*heterogeneity, association, and maternal inheritance in juvenile myoclonic epilepsy." Am J Hum Genet **66**(2): 508-16.*

*Guerrini, R., P. Bonanni, N. Nardocci, et al. (1999). "Autosomal recessive rolandic epilepsy with paroxysmal exercise-induced dystonia and writer's cramp: delineation of the syndrome and gene mapping to chromosome 16p12-11.2." Ann Neurol **45**(3): 344-52.*

*Guerrini, R., F. Moro, M. Kato, et al. (2007). "Expansion of the first PolyA tract of ARX causes infantile spasms and status dystonicus." Neurology **69**(5): 427-33.*

*Guipponi, M., F. Rivier, F. Vigevano, et al. (1997). "Linkage mapping of benign familial infantile convulsions (BFIC) to*

*chromosome 19q." Hum Mol Genet **6**(3): 473-477.*

*Hainsworth, A. H., N. C. McNaughton, A. Pereverzev, et al. (2003). "Actions of sipatrigine, 202W92 and lamotrigine on R-type and T-type Ca2+ channel currents." Eur J Pharmacol **467**(1-3): 77-80.*

*Hanna, I. M., A. Jouvenceau, L. H. Eunson, et al. (2001). "Human Epilepsy - A Possible Role of the Human Voltage-gated P/Q Type Calcium Channel." JNS **187**: S280.*

*Harkin, L. A., D. N. Bowser, L. M. Dibbens, et al. (2002). "Truncation of the GABA(A)-receptor gamma2 subunit in a family with generalized epilepsy with febrile seizures plus." Am J Hum Genet **70**(2): 530-6.*

*Harkin, L. A., J. M. McMahon, X. Iona, et al. (2007). "The spectrum of SCN1A-related infantile epileptic encephalopathies." Brain **130**(Pt 3): 843-52.*

*Heijbel, J., S. Blom and M. Rasmuson (1975). "Benign epilepsy of childhood with centrotemporal EEG foci: a genetic study." Epilepsia **16**: 285-293.*

*Hempelmann, A., J. Cobilanschi, A. Heils, et al. (2007). "Lack of evidence of an allelic association of a functional GABRB3 exon 1a promoter polymorphism with idiopathic generalized epilepsy." Epilepsy Res **74**(1): 28-32.*

*Heron, S. E., K. Cox, B. E. Grinton, et al. (2007). "Deletions or duplications in KCNQ2 can cause benign familial neonatal seizures." J Med Genet **44**(12): 791-6.*

*Heron, S. E., K. M. Crossland, E. Andermann, et al. (2002). "Sodium-channel defects in benign familial neonatal-infantile seizures." Lancet **360**(9336): 851-2.*

*Heron, S. E., H. Khosravani, D. Varela, et al. (2007). "Extended spectrum of idiopathic generalized epilepsies associated with CACNA1H functional variants." Ann Neurol **62**(6): 560-8.*

*Heron, S. E., H. A. Phillips, J. C. Mulley, et al. (2004). "Genetic variation of CACNA1H in idiopathic generalized epilepsy." Ann Neurol **55**(4): 595-6.*

*Hirose, S., H. Iwata, H. Akiyoshi, et al. (1999). "A novel mutation of CHRNA4 responsible for autosomal dominant nocturnal frontal lobe epilepsy." Neurology **53**(8): 1749-53.*

*Hirschhorn, J. N. and M. J. Daly (2005). "Genome-wide association studies for common diseases and complex traits." Nat Rev Genet **6**(2): 95-108.*

*Holmes, G. L. (1993). "Benign focal epilepsies of childhood." Epilepsia **34**(Suppl 3): S49-61.*

*Holmes, G. L., M. McKeever and M. Adamson (1987). "Absence seizures in children: clinical and electroencephalographic features." Ann Neurol **21**(3): 268-73.*

*Holt, R. A., G. M. Subramanian, A. Halpern, et al. (2002). "The genome sequence of the malaria mosquito Anopheles gambiae." Science **298**(5591): 129-49.*

*Hosford, D. A., S. Clark, Z. Cao, et al. (1992). "The role of GABAB receptor activation in absence seizures of lethargic (lh/lh) mice." Science **257**(5068): 398-401.*

*Hugot, J. P., M. Chamaillard, H. Zouali, et al. (2001). "Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease." Nature **411**(6837): 599-603.*

*Huguenard, J. R. (1999). "Neuronal circuitry of thalamocortical epilepsy and mechanisms of antiabsence drug action." Adv Neurol **79**: 991-9.*

*Iannetti, P., A. Spalice, P. F. De Luca, et al. (2001). "Ictal single photon emission computed tomography in absence seizures: apparent implication of different neuronal mechanisms." J Child Neurol **16**(5): 339-44.*

*ILAE (1981). "Proposal for revised clinical and electroencephalographic classification of epileptic seizures. From the Commission on Classification and Terminology of the International League Against Epilepsy." Epilepsia **22**(4): 489-501.*

*ILAE (1985). "Proposal for classification of epilepsies and epileptic syndromes. Commission on Classification and Terminology of the International League Against Epilepsy." Epilepsia **26**(3): 268-78.*

*ILAE (1989). "Proposal for revised classification of epilepsies and epileptic syndromes. Commission on Classification and Terminology of the International League Against Epilepsy." Epilepsia **30**(4): 389-99.*

*ILAE (2009). Report of the Commission on Classification andTerminology: Update and Recommendations.*

*Imbrici, P., S. L. Jaffe, L. H. Eunson, et al. (2004). "Dysfunction of the brain calcium channel CaV2.1 in absence epilepsy and episodic ataxia." Brain **127**(Pt 12): 2682-92.*

*Janz, D. (1969). Die Epilepsien. Stuttgart, Thieme.*

*Janz, D. (1969). Die Epilepsien. Spezielle Pathologie und Therapie., Thieme, Stuttgart.*

*Janz, D. and W. Christian (1957). "Impulsiv-Petit Mal." Journal of Neurology **176**: 346-386.*

*Jeavons, P. M. (1977). "Nosological problems of myoclonic epilepsies in childhood and adolescence." Dev Med Child Neurol **19**(1): 3-8.*

*Jen, J., G. W. Kim and R. W. Baloh (2004). "Clinical spectrum of episodic ataxia type 2." Neurology **62**(1): 17-22.*

*Johnson, W. G., S. L. Kugler, E. S. Stenroos, et al. (1996). "Pedigree analysis in families with febrile seizures." Am J Med Genet **61**(4): 345-52.*

*Jouvenceau, A., L. H. Eunson, A. Spauschus, et al. (2001). "Human epilepsy associated with dysfunction of the brain P/Q-type calcium channel." Lancet **358**(9284): 801-7.*

*Kalachikov, S., O. Evgrafov, B. Ross, et al. (2002). "Mutations in LGI1 cause autosomal-dominant partial epilepsy with auditory features." Nat Genet **28**: 28.*

*Kamatani, N., A. Sekine, T. Kitamoto, et al. (2004). "Large-scale single-nucleotide polymorphism (SNP) and haplotype analyses, using dense SNP Maps, of 199 drug-related genes in 752 subjects: the analysis of the association between uncommon SNPs within haplotype blocks and the haplotypes constructed with haplotype-tagging SNPs." Am J Hum Genet **75**(2): 190-203.*

*Kananura, C., K. Haug, T. Sander, et al. (2002). "A splice-site mutation in GABRG2 associated with childhood absence epilepsy and febrile convulsions." Arch Neurol **59**(7): 1137-41.*

*Kantheti, P., X. Qiao, M. E. Diaz, et al. (1998). "Mutation in AP-3 delta in the mocha mouse links endosomal transport to storage deficiency in platelets, melanosomes, and synaptic vesicles." Neuron **21**(1): 111-22.*

*Kaplan, R. E. and D. J. Lacey (1983). "Benign familial neonatal-infantile seizures." Am J Med Genet **16**(4): 595-9.*

*Karle, J., D. P. Woldbye, L. Elster, et al. (1998). "Antisense oligonucleotide to GABA(A) receptor gamma2 subunit induces limbic status epilepticus." J Neurosci Res **54**(6): 863-9.*

*Khosravani, H., C. Altier, B. Simms, et al. (2004). "Gating effects of mutations in the Cav3.2 T-type calcium channel associated with childhood absence epilepsy." J Biol Chem **279**(11): 9681-4.*

*Kim, D., I. Song, S. Keum, et al. (2001). "Lack of the burst firing of thalamocortical relay neurons and resistance to absence seizures in mice lacking alpha(1G) T-type Ca(2+) channels." Neuron **31**(1): 35-45.*

*Kim, D., I. Song, S. Keum, et al. (2001). "Lack of the burst firing of thalamocortical relay neurons and resistance to absence seizures in mice lacking alpha(1G) T-type Ca(2+) channels." Neuron **31**(1): 35-45.*

*Kleinjan, D. A. and V. van Heyningen (2005). "Long-range control of gene expression: emerging mechanisms and disruption in disease." Am J Hum Genet **76**(1): 8-32.*

*Kleppe, K., E. Ohtsuka, R. Kleppe, et al. (1971). "Studies on polynucleotides. XCVI. Repair replications of short synthetic DNA's as catalyzed by DNA polymerases." J Mol Biol **56**(2): 341-61.*

*Komuro, H. and P. Rakic (1992). "Selective role of N-type calcium channels in neuronal migration." Science **257**(5071): 806-9.*

*Kong, A., D. F. Gudbjartsson, J. Sainz, et al. (2002). "A high-resolution recombination map of the human genome." Nat Genet **31**(3): 241-7.*

*Kong, X., K. Murphy, T. Raj, et al. (2004). "A combined linkage-physical map of the human genome." Am J Hum Genet **75**(6): 1143-8.*

*Kosambi, D. D. (1944). "The estimation of map distance from recombination values." Annals of Eugenics **12**(3): 172-175.*

*Kostopoulos, G. K. (2000). "Spike-and-wave discharges of absence seizures as a transformation of sleep spindles: the continuing development of a hypothesis." Clin Neurophysiol **111 Suppl 2**: S27-38.*

*Kruglyak, L., M. J. Daly, M. P. Reeve-Daly, et al. (1996). "Parametric and nonparametric linkage analysis: a unified multipoint approach." Am J Hum Genet **58**(6): 1347-63.*

*Kwok, P. Y. (2001). "Methods for genotyping single nucleotide polymorphisms." Annu Rev Genomics Hum Genet **2**: 235-58.*

*Lathrop, G. M. and J. M. Lalouel (1984). "Easy calculations of lod scores and genetic risks on small computers." Am J Hum Genet **36**(2): 460-5.*

*Lathrop, G. M., J. M. Lalouel, C. Julier, et al. (1984). "Strategies for multilocus linkage analysis in humans." Proc Natl Acad Sci U S A **81**(11): 3443-6.*

*Lawrence, R., D. M. Evans, A. P. Morris, et al. (2005). "Genetically indistinguishable SNPs and their influence on inferring the location of disease-associated variants." Genome Res **15**(11): 1503-10.*

*Lee, W. L., P. S. Low and U. Rajan (1993). "Benign familial infantile epilepsy." J Pediatr **123**(4): 588-90.*

*Lee, W. L., A. Tay, H. T. Ong, et al. (1998). "Association of infantile convulsions with paroxysmal dyskinesias (ICCA syndrome): confirmation of linkage to human chromosome 16p12-q12 in a Chinese family." Hum Genet **103**(5): 608-12.*

*Lennox, W. G. (1945). "The petit mal epilepsies - their treatment with tridione." JAMA(129): 1069-74.*

*Lennox, W. G. and M. A. Lennox (1960). Epilepsy and related disorders. Boston, Little, Brown.*

*Leonardi, M. and T. B. Ustun (2002). "The global burden of epilepsy." Epilepsia **43 Suppl 6**: 21-5.*

*Leppert, M., V. E. Anderson, T. Quattlebaum, et al. (1989). "Benign familial neonatal convulsions linked to genetic markers on chromosome 20." Nature **337**(6208): 647-648.*

*Leresche, N., H. R. Parri, G. Erdemli, et al. (1998). "On the action of the anti-absence drug ethosuximide in the rat and cat thalamus." J Neurosci **18**(13): 4842-53.*

*Lerman, P. and S. Kivity (1975). "Benign focal epilepsy of childhood. A follow-up study of 100 recovered patients." Arch Neurol **32**(4): 261-4.*

*Letts, V. A., R. Felix, G. H. Biddlecome, et al. (1998). "The mouse stargazer gene encodes a neuronal Ca2+-channel gamma subunit." nature genetics **19**: 340-347.*

*Lewis, T. B., R. J. Leach, K. Ward, et al. (1993). "Genetic heterogeneity in benign familial neonatal convulsions: identification of a new locus on chromosome 8q." American Journal of Human Genetics. **53**: 670-675.*

*Liu, A. W., A. V. Delgado-Escueta, M. N. Gee, et al. (1996). "Juvenile myoclonic epilepsy in chromosome 6p12-p11: locus heterogeneity and recombinations." Am J Med Genet **63**(3): 438-46.*

*Liu, A. W., A. V. Delgado-Escueta, J. M. Serratosa, et al. (1995). "Juvenile Myoclonic Epilepsy locus in chromosome 6p21.2-p11:linkage to convulsions and electroencephalography trait." American Journal of Human Genetics **57**: 368-381.*

*Livingston, S., I. Torres, L. L. Pauli, et al. (1965). "Petit mal epilepsy. Results of a prolonged follow-up study of 117 patients." Jama **194**(3): 227-32.*

*Loiseau, J., P. Loiseau, M. Guyot, et al. (1990). "Survey of seizure disorders in the French southwest. I. Incidence of epileptic syndromes." Epilepsia **31**(4): 391-6.*

*Loiseau, P. (1992). Childhood absence epilepsy. Epilepsy Syndromes in Infancy, Childhood and Adolescence (2nd Edition). J. Roger, M. Bureau, C. H. Dravetet al. London, John Libbey: 135-150.*

*Loiseau, P., B. Duche and J. M. Pedespan (1995). "Absence epilepsies." Epilepsia **36**(12): 1182-6.*

*Macdonald, R. L. and K. M. Kelly (1994). "Mechanisms of action of currently prescribed and newly developed antiepileptic drugs." Epilepsia **35**(Suppl 4): S41-50.*

*Malafosse, A., M. Leboyer, O. Dulac, et al. (1992). "Confirmation of linkage of benign familial neonatal convulsions to D20S19 and D20S20." Human Genetics **89**: 54-58.*

*Mantegazza, M., A. Gambardella, R. Rusconi, et al. (2005). "Identification of an Nav1.1 sodium channel (SCN1A) loss-of-function mutation associated with familial simple febrile seizures." Proc Natl Acad Sci U S A **102**(50): 18177-82.*

*Mari, F., S. Azimonti, I. Bertani, et al. (2005). "CDKL5 belongs to the same molecular pathway of MeCP2 and it is responsible for the early-onset seizure variant of Rett syndrome." Hum Mol Genet **14**(14): 1935-46.*

*Marini, C., I. E. Scheffer, K. M. Crossland, et al. (2004). "Genetic architecture of idiopathic generalized epilepsy: clinical genetic analysis of 55 multiplex families." Epilepsia **45**(5): 467-78.*

*Martin, E. R., M. P. Bass and N. L. Kaplan (2001). "Correcting for a potential bias in the pedigree disequilibrium test." Am J Hum Genet **68**(4): 1065-7.*

*Martin, E. R., S. A. Monks, L. L. Warren, et al. (2000). "A test for linkage and association in general pedigrees: the pedigree disequilibrium test." Am J Hum Genet **67**(1): 146-54.*

*Martin, S. C., S. J. Russek and D. H. Farb (1999). "Molecular identification of the human GABABR2: cell surface expression and coupling to adenylyl cyclase in the absence of GABABR1." Mol Cell Neurosci **13**(3): 180-91.*

*Meisler, M. H., J. Kearney, R. Ottman, et al. (2001). "Identification of epilepsy genes in human and mouse." Annu Rev Genet **35**: 567-88.*

*Meng, Z., D. V. Zaykin, C. F. Xu, et al. (2003). "Selection of genetic markers for association analyses, using linkage disequilibrium and haplotypes." Am J Hum Genet **73**(1): 115-30.*

*Morton, N. E. (1955). "Sequential tests for the detection of linkage." American journal of human genetics **7**: 277-318.*

*Moss, F. J., A. C. Dolphin and J. J. Clare (2003). "Human neuronal stargazin-like proteins, gamma2, gamma3 and gamma4; an investigation of their specific localization in human brain and their influence on CaV2.1 voltage-dependent calcium channels expressed in Xenopus oocytes." BMC Neurosci **4**(1): 23.*

*Murbartian, J., J. M. Arias and E. Perez-Reyes (2004). "Functional impact of alternative splicing of human T-type Cav3.3 calcium channels." J Neurophysiol **92**(6): 3399-407.*

*Nakayama, J., Y. H. Fu, A. M. Clark, et al. (2002). "A nonsense mutation of the MASS1 gene in a family with febrile and afebrile seizures." Ann Neurol **52**(5): 654-7.*

*Nayrac, P. and M. Beaussart (1958). "Les pointe-ondes prérolandique: expression EEG très particulière." Rev. Neurol **99**(201-206).*

*Neubauer, B. A., B. Fiedler, B. Himmelein, et al. (1998). "Centrotemporal spikes in families with rolandic epilepsy: linkage to chromosome 15q14." Neurology **51**(6): 1608-12.*

*Neyroud, N., F. Tesson, I. Denjoy, et al. (1997). "A novel mutation in the potassium channel gene KVLQT1 causes the Jervell and Lange-Nielsen cardioauditory syndrome." Nat Genet **15**(2): 186-9.*

*Nyholt, D. R. (2004). "A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other." Am J Hum Genet **74**(4): 765-9.*

*Nyholt, D. R. (2005). "Evaluation of Nyholt's procedure for multiple testing correction - author's reply." Hum Hered **60**(1): 61-2.*

*Ogura, Y., D. K. Bonen, N. Inohara, et al. (2001). "A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease." Nature **411**(6837): 603-6.*

*Ohmori, I., M. Ouchida, Y. Ohtsuka, et al. (2002). "Significant correlation of the SCN1A mutations and severe myoclonic epilepsy in infancy." Biochem Biophys Res Commun **295**(1): 17-23.*

*Ophoff, R. A., G. M. Terwindt, M. N. Vergouwe, et al. (1996). "Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the $Ca^{2+}$ channel gene CACNL1A4." Cell **87**: 543-552.*

*Ott, J. (1989). "Computer-simulation methods in human linkage analysis." Proc Natl Acad Sci U S A **86**(11): 4175-8.*

*Ottman, R., N. Risch, W. A. Hauser, et al. (1995). "Localization of a gene for partial epilepsy to chromosome 10q." Nature Genetics **10**: 56-60.*

*Ottman, R., M. R. Winawer, S. Kalachikov, et al. (2004). "LGI1 mutations in autosomal dominant partial epilepsy with auditory features." Neurology **62**(7): 1120-6.*

*Pal, D. K., O. V. Evgrafov, P. Tabares, et al. (2003). "BRD2 (RING3) is a probable major susceptibility gene for common juvenile myoclonic epilepsy." Am J Hum Genet **73**(2): 261-70.*

*Panayiotopoulos, C. P. (1993). "Benign childhood partial epilepsies: benign childhood seizure susceptibility syndromes." J Neurol Neurosurg Psychiatry **56**(1): 2-5.*

*Panayiotopoulos, C. P., C. D. Ferrie, G. S, et al. (1995). Perioral myoclonia with absences. Typical absences and related epileptic syndromes. J. S. Duncan and C. P. Panayiotopoulos. London, Churchill Communications Europe.*

*Panayiotopoulos, C. P. and T. Obeid (1989). "Juvenile myoclonic epilepsy: an autosomal recessive disease." Annals of Neurology **25**: 440-443.*

*Pavone, A. and E. Niedermeyer (2000). "Absence seizures and the frontal lobe." Clin Electroencephalogr **31**(3): 153-6.*

*Pennartz, C. M., M. T. de Jeu, N. P. Bos, et al. (2002). "Diurnal modulation of pacemaker potentials and calcium current in the mammalian circadian clock." Nature **416**(6878): 286-90.*

*Pertea, M., X. Lin and S. L. Salzberg (2001). "GeneSplicer: a new computational method for splice site prediction." Nucleic Acids Res **29**(5): 1185-90.*

*Phillips, H. A., I. Favre, M. Kirkpatrick, et al. (2001). "CHRNB2 is the second acetylcholine receptor subunit associated with autosomal dominant nocturnal frontal lobe epilepsy." Am J Hum Genet **68**(1): 225-31.*

*Phillips, H. A., C. Marini, I. E. Scheffer, et al. (2000). "A de novo mutation in sporadic nocturnal frontal lobe epilepsy." Ann Neurol **48**(2): 264-7.*

*Phillips, H. A., I. E. Scheffer, S. F. Berkovic, et al. (1995). "Localization of a gene for autosomal dominant nocturnal frontal lobe epilepsy to chromosome 20q13.2." Nature Genetics **10**: 117-118.*

*Phillips, H. A., I. E. Scheffer, K. M. Crossland, et al. (1998). "Autosomal dominant nocturnal frontal-lobe epilepsy: genetic heterogeneity and evidence for a second locus at 15q24." Am J Hum Genet **63**(4): 1108-16.*

*Pinault, D., N. Leresche, S. Charpier, et al. (1998). "Intracellular recordings in thalamic neurones during spontaneous spike and wave discharges in rats with absence epilepsy." J Physiol **509**(Pt 2): 449-56.*

*Porter, R. (1997). The greatest benefit to mankind. London, Harper Collins.*

*Poulter, M. O., L. A. Brown, S. Tynan, et al. (1999). "Differential expression of alpha1, alpha2, alpha3, and alpha5 GABAA receptor subunits in seizure-prone and seizure-resistant rat models of temporal lobe epilepsy." J Neurosci **19**(11): 4654-61.*

*Poza, J. J., A. Saenz, A. Martinez-Gil, et al. (1999). "Autosomal dominant lateral temporal epilepsy: clinical and genetic study of a large Basque pedigree linked to chromosome 10q." Ann Neurol **45**(2): 182-8.*

*Prevett, M. C., J. S. Duncan, T. Jones, et al. (1995). "Demonstration of thalamic activation during typical absence seizures using H2(15)O and PET." Neurology **45**(7): 1396-402.*

*Pritchard, J. K., M. Stephens, N. A. Rosenberg, et al. (2000). "Association mapping in structured populations." Am J Hum Genet **67**(1): 170-81.*

*Pumain, R., J. Louvel, M. Gastard, et al. (1992). "Responses to N-methyl-D-aspartate are enhanced in rats with petit mal-like seizures." J Neural Transm Suppl **35**: 97-108.*

*Reich, D. E. and E. S. Lander (2001). "On the allelic spectrum of human disease." Trends Genet **17**(9): 502-10.*

*Rett, A. and R. Teubel (1964). "Neugeborenen Krampfe im Rahmen einer epileptisch belasten Familie." Wiener Klinische Wochenschrift **76**: 609-613.*

*Rich, S. S., J. F. Annegers, W. A. Hauser, et al. (1987). "Complex segregation analysis of febrile convulsions." Am J Hum Genet **41**(2): 249-57.*

*Richards, D. A., J. P. Manning, D. Barnes, et al. (2003). "Targeting thalamic nuclei is not sufficient for the full anti-absence action of ethosuximide in a rat model of absence epilepsy." Epilepsy Res **54**(2-3): 97-107.*

*Rozen, S. and H. Skaletsky (2000). "Primer3 on the WWW for general users and for biologist programmers." Methods Mol Biol **132**: 365-86.*

*Rugtveit, J. (1986). "X-linked mental retardation and infantile spasms in two brothers." Dev Med Child Neurol **28**(4): 544-6.*

*Ryan, S. G., M. Wiznitzer, C. Hollman, et al. (1991). "Benign familial neonatal convulsions: evidence for clinical and genetic heterogeneity." Annals of Neurology **29**(5): 469-473.*

*Saenz, A., J. Galan, C. Caloustian, et al. (1999). "Autosomal dominant nocturnal frontal lobe epilepsy in a Spanish family with a Ser252Phe mutation in the CHRNA4 gene." Arch Neurol **56**(8): 1004-9.*

*Saiki, R. K., S. Scharf, F. Faloona, et al. (1985). "Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia." Science **230**(4732): 1350-4.*

*Salyakina, D., S. R. Seaman, B. L. Browning, et al. (2005). "Evaluation of Nyholt's procedure for multiple testing correction." Hum Hered **60**(1): 19-25; discussion 61-2.*

*Sander, T., H. Schulz, K. Saar, et al. (2000). "Genome search for susceptibility loci of common idiopathic generalised epilepsies." Hum Mol Genet **9**(10): 1465-72.*

*Sander, T., H. Schulz, A. M. Vieira-Saeker, et al. (1999). "Evaluation of a putative major susceptibility locus for juvenile myoclonic epilepsy on chromosome 15q14." Am J Med Genet **88**(2): 182-7.*

*Sander, T., C. Windemuth, H. Schulz, et al. (2003). "Exploration of a putative susceptibility locus for idiopathic generalized epilepsy on chromosome 8p12." Epilepsia **44**(1): 32-9.*

*Sayer, R. J., A. M. Brown, P. C. Schwindt, et al. (1993). "Calcium currents in acutely isolated human neocortical neurons." J Neurophysiol **69**(5): 1596-606.*

*Scheel, H., S. Tomiuk and K. Hofmann (2002). "A common protein interaction domain links two recently identified epilepsy genes." Hum Mol Genet **11**(15): 1757-62.*

*Scheffer, I. E. and S. F. Berkovic (1997). "Generalized epilepsy with febrile seizures plus. A genetic disorder with heterogeneous clinical phenotypes." Brain **120**(Pt 3): 479-90.*

*Scheffer, I. E., K. P. Bhatia, I. Lopes-Cendes, et al. (1994). "Autosomal dominant frontal epilepsy misdiagnosed as sleep disorder." Lancet **343**: 515-517.*

*Scheffer, I. E., H. A. Phillips, C. E. O'Brien, et al. (1998). "Familial partial epilepsy with variable foci: a new partial epilepsy syndrome with suggestion of linkage to chromosome 2." Ann Neurol **44**(6): 890-9.*

*Scheffer, I. E., R. H. Wallace, F. L. Phillips, et al. (2002). "X-linked myoclonic epilepsy with spasticity and intellectual disability: mutation in the homeobox gene ARX." Neurology **59**(3): 348-56.*

*Schuler, V., C. Luscher, C. Blanchet, et al. (2001). "Epilepsy, hyperalgesia, impaired memory, and loss of pre- and postsynaptic GABA(B) responses in mice lacking GABA(B(1))." Neuron **31**(1): 47-58.*

*Singer, V. L., L. J. Jones, S. T. Yue, et al. (1997). "Characterization of PicoGreen reagent and development of a fluorescence-based solution assay for double-stranded DNA quantitation." Anal Biochem **249**(2): 228-38.*

*Singh, R., R. J. Gardner, et al. (2002). "Chromosomal abnormalities and epilepsy: a review for clinicians and gene hunters." Epilepsia **43**(2): 127-40.*

*Singh, N. A., C. Charlier, D. Stauffer, et al. (1998). "A novel potassium channel gene, KCNQ2, is mutated in an inhertied epilepsy of newborns." Nature Genetics **18**: 25-29.*

*Singh, R., E. Andermann, W. P. Whitehouse, et al. (2001). "Severe myoclonic epilepsy of infancy: extended spectrum of GEFS+?" Epilepsia **42**(7): 837-44.*

*Snead, O. C., 3rd (1996). "Antiabsence seizure activity of specific GABAB and gamma-Hydroxybutyric acid receptor antagonists." Pharmacol Biochem Behav **53**(1): 73-9.*

*Solovyev, V. V. and I. A. Shahmuradov (2003). "PromH: Promoters identification using orthologous genomic sequences." Nucleic Acids Res **31**(13): 3540-5.*

*Somerville, R. P., O. Chernova, S. Liu, et al. (2000). "Identification of the promoter, genomic structure, and mouse ortholog of LGI1." Mamm Genome **11**(8): 622-7.*

*Spauschus, A., L. Eunson, M. G. Hanna, et al. (1999). "Functional characterization of a novel mutation in KCNA1 in episodic ataxia type 1 associated with epilepsy." Ann N Y Acad Sci **868**: 442-6.*

*Spielman, R. S., R. E. McGinnis and W. J. Ewens (1993). "Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM)." Am J Hum Genet **52**(3): 506-16.*

*Staak, R. and H. C. Pape (2001). "Contribution of GABA(A) and GABA(B) receptors to thalamic neuronal activity during spontaneous absence seizures in rats." J Neurosci **21**(4): 1378-84.*

*Steinlein, O., A. Magnusson, J. Stoodt, et al. (1997). "An insertion mutation of the CHRNA4 gene in a family with autosomal dominant nocturnal frontal lobe epilepsy." Human Molecular Genetics **6**(6): 943-947.*

*Steinlein, O. K., J. C. Mulley, P. Propping, et al. (1995). "A missense mutation in the neuronal nicotinic acetylcholine receptor alpha 4 subunit is associated with autosomal dominant nocturnal frontal lobe epilepsy." Nat Genet **11**(2): 201-3.*

*Steinlein, O. K., J. Stoodt, J. Mulley, et al. (2000). "Independent occurrence of the CHRNA4 Ser248Phe mutation in a Norwegian family with nocturnal frontal lobe epilepsy." Epilepsia **41**(5): 529-35.*

*Steriade, M. and D. Contreras (1998). "Spike-wave complexes and fast components of cortically generated seizures. I. Role of neocortex and thalamus." J Neurophysiol **80**(3): 1439-55.*

*Stogmann, E., P. Lichtner, C. Baumgartner, et al. (2006). "Idiopathic generalized epilepsy phenotypes associated with different EFHC1 mutations." Neurology **67**(11): 2029-31.*

*Strauch, K., R. Fimmers, T. Kurz, et al. (2000). "Parametric and nonparametric multipoint linkage analysis with imprinting and two-locus-trait models: application to mite sensitization." Am J Hum Genet **66**(6): 1945-57.*

*Stromme, P., M. E. Mangelsdorf, M. A. Shaw, et al. (2002). "Mutations in the human ortholog of Aristaless cause X-linked mental retardation and epilepsy." Nat Genet **30**(4): 441-5.*

*Stromme, P., K. Sundet, C. Mork, et al. (1999). "X linked mental retardation and infantile spasms in a family: new clinical data and linkage to Xp11.4-Xp22.11." J Med Genet **36**(5): 374-8.*

*Sugawara, T., E. Mazaki-Miyazaki, K. Fukushima, et al. (2002). "Frequent mutations of SCN1A in severe myoclonic epilepsy in infancy." Neurology **58**(7): 1122-4.*

*Sugawara, T., Y. Tsurubuchi, K. L. Agarwala, et al. (2001). "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**(11): 6384-9.*

*Sugimoto, Y., R. Morita, K. Amano, et al. (2000). "Childhood absence epilepsy in 8q24: refinement of candidate region and construction of physical map." Genomics **68**(3): 264-72.*

*Sundqvist, A. (1990). "Juvenile myoclonic epilepsy: events before diagnosis." Journal of Epilepsy **3**: 189-192.*

*Suzuki, T., A. V. Delgado-Escueta, K. Aguan, et al. (2004). "Mutations in EFHC1 cause juvenile myoclonic epilepsy." Nat Genet **36**(8): 842-9.*

*Szepetowski, P., J. Rochette, P. Berquin, et al. (1997). "Familial infantile convulsions and paroxysmal choreoathetosis: a new neurological syndrome linked to the pericentromeric region of human chromosome 16." Am J Hum Genet **61**(4): 889-98.*

*Talley, E. M., L. L. Cribbs, J. H. Lee, et al. (1999). "Differential distribution of three members of a gene family encoding low voltage-activated (T-type) calcium channels." J Neurosci **19**(6): 1895-911.*

*Tanaka, M., I. Castroviejo, M. Medina, et al. (2000). "Linkage analysis between subsyndromes of childhood absence epilepsy and the $GABA_A$ receptor beta3 subunit on chromosome 15q11.2-12." Epilepsia **41**(S7): 250.*

*Tanaka, M., R. W. Olsen, M. T. Medina, et al. (2008). "Hyperglycosylation and reduced GABA currents of mutated GABRB3 polypeptide in remitting childhood absence epilepsy." Am J Hum Genet **82**(6): 1249-61.*

*Taske, N. L., M. P. Williamson, A. Makoff, et al. (2002). "Evaluation of the positional candidate gene CHRNA7 at the juvenile myoclonic epilepsy locus (EJM2) on chromosome 15q13-14." Epilepsy Res **49**(2): 157-72.*

*Tassinari, C. A., S. Lyagoubi, V. Santos, et al. (1969). "[Study on spike and wave discharges in man. II. Clinical and electroencephalographic aspects of myoclonic absences]." Rev Neurol (Paris) **121**(3): 379-83.*

*Tassinari, C. A., R. Michelucci, R. G, et al. (1995). Myoclonic absence epilepsy. Typical absences and related epileptic syndromes. J. S. Duncan and C. P. Panayiotopoulos. London, Churchill Communications Europe.*

*Tranah, G. J., P. J. Lescault, D. J. Hunter, et al. (2003). "Multiple displacement amplification prior to single nucleotide polymorphism genotyping in epidemiologic studies." Biotechnol Lett **25**(13): 1031-6.*

*Trettel, F., E. Mantuano, V. Calabresi, et al. (2000). "A fine physical map of the CACNA1A gene region on 19p13.1-p13.2 chromosome." Gene **241**(1): 45-50.*

*Tsuboi, T. and W. Christian (1973). "On the genetics of primary generalised epilepsy with sporadic myoclonus of impulsive petit-mal type." Humangenetik **19**: 155-182.*

*Urak, L., M. Feucht, N. Fathi, et al. (2006). "A GABRB3 promoter haplotype associated with childhood absence epilepsy impairs transcriptional activity." Hum Mol Genet **15**(16): 2533-41.*

*Vadlamudi, L., A. S. Harvey, M. M. Connellan, et al. (2004). "Is benign rolandic epilepsy genetically determined?" Ann Neurol **56**(1): 129-32.*

*Vergnes, M., A. Boehrer, S. Reibel, et al. (2000). "Selective susceptibility to inhibitors of GABA synthesis and antagonists of GABA(A) receptor in rats with genetic absence epilepsy." Exp Neurol **161**(2): 714-23.*

*Vigevano, F., L. Fusco, M. Di Capua, et al. (1992). "Benign infantile familial convulsions." Eur J Pediatr **151**(8): 608-12.*

*Vitko, I., Y. Chen, J. M. Arias, et al. (2005). "Functional characterization and neuronal modeling of the effects of childhood absence epilepsy variants of CACNA1H, a T-type calcium channel." J Neurosci **25**(19): 4844-55.*

*Wagstaff, J., J. H. Knoll, J. Fleming, et al. (1991). "Localization of the gene encoding the GABAA receptor beta 3 subunit to the Angelman/Prader-Willi region of human chromosome 15." Am J Hum Genet **49**(2): 330-7.*

*Wallace, R., D. Wang, R. Singh, et al. (1998). "Febrile seizures and generalized epilepsy associated with a mutation in the Na+-channel beta1 subunit gene SCN1B." Nat Genet **19**(4): 366-370.*

*Wallace, R. H., C. Marini, S. Petrou, et al. (2001). "Mutant GABAA receptor gamma2-subunit in childhood absence epilepsy and febrile seizures." Nat Genet **28**(1): 49-52.*

*Wallace, S. (1988). Genetic factors. The Child with Febrile Seizures. S. Wallace. London, John Wright**:** 24.*

*Wang, H. S., Z. Pan, W. Shi, et al. (1998). "KCNQ2 and KCNQ3 potassium channel subunits: molecular correlates of the M-channel." Science **282**(5395): 1890-3.*

*Wang, Q., M. E. Curran, I. Splawski, et al. (1996). "Positional cloning of a novel potassium channel gene: KVLQT1 mutations cause cardiac arrhythmias." Nat Genet **12**(1): 17-23.*

*Waterston, R. H., K. Lindblad-Toh, E. Birney, et al. (2002). "Initial sequencing and comparative analysis of the mouse genome." Nature **420**(6915): 520-62.*

*Weaving, L. S., J. Christodoulou, S. L. Williamson, et al. (2004). "Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation." Am J Hum Genet **75**(6): 1079-93.*

*Weber, J. L. and P. E. May (1989). "Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction." Am J Hum Genet **44**(3): 388-96.*

*Weeks, D. E., J. Ott and G. M. Lathrop (1990). "SLINK: a general simulation program for linkage analysis." American Journal of Human Genetics **47**: A204.*

*Weissbecker, K. A., M. Durner, D. Janz, et al. (1991). "Confirmation of linkage between juvenile myoclonic epilepsy locus and the HLA region on chromosome 6." American Journal of Medical Genetics **38**(1): 32-36.*

*Whitehouse, W., U. Diebold, M. Rees, et al. (1993). "Exclusion of linkage of genetic focal sharp waves to the HLA region on chromosome 6p in families with benign partial epilepsy with centrotemporal spikes." Neuropaediatrics **24**(4): 208-210.*

*Winawer, M. R., D. Rabinowitz, T. A. Pedley, et al. (2003). "Genetic influences on myoclonic and absence seizures." Neurology **61**(11): 1576-81.*

*Wright, A. F. and N. D. Hastie (2001). "Complex genetic diseases: controversy over the Croesus code." Genome Biol **2**(8).*

*Xiong, L., M. Labuda, D. S. Li, et al. (1999). "Mapping of a gene determining familial partial epilepsy with variable foci to chromosome 22q11-q12." Am J Hum Genet **65**(6): 1698-710.*

*Ye, J., E. J. Parra, D. M. Sosnoski, et al. (2002). "Melting curve SNP (McSNP) genotyping: a useful approach for diallelic genotyping in forensic science." J Forensic Sci **47**(3): 593-600.*

*Yunker, A. M. and M. W. McEnery (2003). "Low-voltage-activated ("T-Type") calcium channels in review." J Bioenerg Biomembr **35**(6): 533-75.*

*Zhang, Y., M. Mori, D. L. Burgess, et al. (2002). "Mutations in high-voltage-activated calcium channel genes stimulate low-voltage-activated currents in mouse thalamic relay neurons." J Neurosci **22**(15): 6362-71.*

*Zhuchenko, O., J. Bailey, P. Bonnen, et al. (1997). "Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the alpha 1A-voltage-dependent calcium channel." Nat Genet **15**(1): 62-9.*

# Appendices

## Appendix 1 Invitation letters to clinicians

Appendix 1a Mailshot to Consultant Paediatricians in North and South Thames Regions

Royal Free and University College Medical School
**UNIVERSITY COLLEGE LONDON**

DEPARTMENT OF PAEDIATRICS

Gower Street Campus
The Rayne Institute
5 University Street
London WC1E 6JJ

Direct Line Telephone 0207 679 6131
Fax 0207 679 6103

Dear (name inserted)

**Re: Genetic analysis of childhood absence epilepsy (Petit Mal)**

As you may know, we have a programme of research directed towards identifying the genes causing certain childhood epilepsies.

At present, work is focused on absence epilepsies, ie:

**Childhood absence epilepsy (CAE)**
**Juvenile absence epilepsy (JAE)**

We are therefore ascertaining patients with the above diagnoses, even if there is no family history. Pairs of affected siblings are most informative, but single cases can provide vital information using recently developed techniques. All that will be required is some clinical information and either blood or saliva for DNA extraction. Full ethical permission has been granted.

If you have patients who might be suitable for study, would like to collaborate, and wish to receive more information about the project, please complete the form below and return by fax or post to the above address.

Yours sincerely

R.M. Gardiner
Professor of Paediatrics

Dr R. Robinson
Paediatric Research Fellow

---

I would like to receive more information about the research project on genetic analysis of childhood absence epilepsy.

The approximate number of suitable patients is: ☐

Signed:

Appendix 1b Letter of invitation to consultant paediatricians with patients identified from EEG records.

Royal Free and University College Medical School
**UNIVERSITY COLLEGE LONDON**

DEPARTMENT OF PAEDIATRICS

Gower Street Campus
The Rayne Institute
5 University Street
London WC1E 6JJ

Direct Line Telephone 0207 679 6131
Fax 0207 679 6103

Dear Dr (name inserted)

**Re: The genetic analysis of childhood absence epilepsy**

As you may know, we have a programme of research directed towards identifying the genes causing certain childhood epilepsies. We are currently collecting patients with a diagnosis of childhood absence epilepsy.

The EEG database at «EEG_Department» .has thrown up patients with probable absence epilepsy under your care namely:

«Patient_1» «DOB»
«Patient_2» «DOB1»
«Patient_3» «DOB2»
«Patient_4» «DOB3»
«Patient_5» «DOB4»
«Patient_6» «DOB5»

With your permission, we could write to the families directly with information about the study. If the families are suitable and agree to participate, all that would be required is a sample of DNA from each patient and their immediate family. A simple cheek swab sent through the post can obtain this.

Multi-centre ethics approval has been granted by the North Thames MREC.

We enclose a copy of the parent information sheet. Please let us know if you feel these patients would be suitable and whether you would be happy for us to contact the families directly.

Yours sincerely,

Dr Robert Robinson
Clinical Research Fellow in Paediatrics

Pauline Boland
Clinical Research Nurse

## Appendix 2 UK Neurophysiology Departments collaborating in study

| Neurophysiology Department | Contact |
|---|---|
| Addenbrooke's Hospital, Cambridge CB2 2QQ | Dr Simon Boniface |
| Birmingham Children's Hospital, Birmingham B4 6NL | Dr Seri |
| Bristol Royal Hospital for Children, Bristol B52 8BJ | Phillip Blackwell |
| Central Middlesex Hospital, London NW10 7NS | Dr Jamal |
| Chelsea and Westminster Hospital, London, SW10 9NH | Dr Ian Mak |
| Gloucestershire Royal Hospital | Dr Oware |
| Great Ormond Street Hospital, London WC1N 3JH | Dr Stewart Boyd |
| Guy's Hospital, London SE1 9RT | Dr Michael O'Brien |
| Hammersmith Hospital, London W12 ONN | Dr Khalil |
| King's College Hospital, London SE5 9RS | Professor C Binnie |
| Leicester Royal Infirmary, Leicester LE1 5WW | Karen Widley |
| Northampton General Hospital | Dr Bissessar |
| Park Hospital, Oxford OX3 7LQ | Dr Zenobia Zaiwalla |
| Queens Medical Centre, Nottingham NG7 2UH | Sylvia Remmington |
| Royal Devon & Exeter Hospital | Dr Elias Ragi |
| Royal London Hospital, London E1 1BB | Dr Franz Brunhuber |
| Royal Surrey County Hospital, Surrey GU2 5XX | Sharon Boxall |
| Southampton University Hospital, Southampton, SO16 3YD | Dr Vander Starr |
| St Hellier Hospital, Surrey, SM5 1AA | Karen Widley |
| St Mary's Hospital, London W2 1NY | Dr Steven White |
| St Mary's Hospital, Portsmouth PO3 6AD | Dr Merton |
| St Thomas' Hospital, London SE1 7EH | Dr C P Panyiotopoulos |
| University College London Hospital, London WC1E 3DB | Dr David Holder |
| York District Hospital, York YO31 8HE | Dr Sibte Hasan |

## Appendix 3 BPNA Newsletter January 2002 – Call for Patients

### Call for Patients – Childhood Absence Epilepsy (CAE)

Molecular Genetic Analysis of Childhood Absence Epilepsy
Robert Robinson, Nichole Taske, Michele Rees, Mark Gardiner
Department of Paediatrics
Royal Free & University College Medical School, University College London
WCIE 6JJ

The study is aiming for 300 patients – new families with a proband with CAE are being recruited on an ongoing basis. These families include:
- single affected children and their unaffected parents (trios)
- affected sibling pairs
- any pedigrees with two or more affected members

The inclusion criteria are:
- Absence seizures (of any type except myoclonic absences)
- Onset between 2 and 12 years
- Normal development (mild learning difficulties are allowed)
- Ictal EEG showing bilateral, synchronous, symmetrical discharges of 2.5 – 4 Hz spike-wave or polyspike wave complexes on a normal background

Once families agree to participate, we can obtain most of the clinical information directly from them by phone and post. Samples for DNA extraction are required from affected individuals and their parents. We can often arrange collection of blood samples at their own G.P. In some cases cheek swabs are adequate, which subjects can take themselves and send back using pre-prepared collection kits.

Copies of patient information sheets, consent forms, study protocol and ethical approval can be provided on request.

Please contact
Robert Robinson
Action Research Training Fellow
Email – robert.robinson@ucl.ac.uk
Department of Paediatrics
University College Hospital London
The Rayne Institute
5 University Street
London WC1 6JJ

Tel – 020 7679 6131
Fax – 020 7679 6103

## Appendix 4 Project summary and call for patients circulated at BPNA Annual Meeting 2001

# Molecular genetic analysis of childhood absence epilepsy

**Robert Robinson**, Department of Paediatrics, Royal Free & University College Medical School, University College London, WC1E 6JJ, UK.
E-mail: robert.robinson@ucl.ac.uk
Telephone: 44 (0)20 7679 6131 Fax: 44 (0)20 7679 6103

***Genetics of human epilepsy.*** Epilepsy is one of the most common childhood neurological disorders and genetic factors contribute to aetiology in about 40% of cases. Over 100 single-gene mendelian disorders include epilepsy as one component of what is usually a more complex neurological phenotype. Individually these diseases are rare and the majority of common idiopathic generalised epilepsies (IGEs) display a complex, non-mendelian pattern of inheritance.

***Childhood absence epilepsy*** (CAE) is an inherited IGE. Most studies demonstrate a complex form of inheritance, but one recent segregation analysis was consistent with an autosomal dominant mode of inheritance with reduced penetrance.

***Molecular genetics of inherited epilepsies.*** Idiopathic mendelian epilepsies in human and mouse have been identified as channelopathies. Mutations causing human epilepsy have been found in the gene encoding the α4 subunit of the neuronal nicotinic acetylcholine receptor (*CHRNA4*), two voltage gated $K^+$ channel genes (*KCNQ2*, *KCNQ3*), and two sodium channel genes (*SCN1B*, *SCN1A*). Mutations in voltage-dependent calcium channel genes have been identified in both human and murine absence epilepsies.

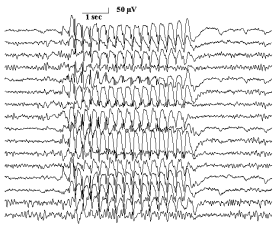
Our ***strategy*** is to ascertain families with childhood absence epilepsy, and to investigate the role of candidate genes in these families. As well as traditional linkage analysis, we are using intrafamilial association analysis which allows us to utilise the large number of families ***with only one affected individual.***

***Diagnostic criteria***

- Absence seizures (of any type except myoclonic absences)
- Onset between 2 and 12 years
- Normal development
- Ictal EEG showing bilateral, synchronous, symmetrical discharges of 2.5-4 Hz spike-wave or polyspike-wave complexes on a normal background

***DNA collection***

DNA is obtained from all affected family members, both parents and, where possible, unaffected siblings.
DNA is extracted from either blood samples or cheek swabs, both of which can be transported by mail.
Multi-centre research ethics approval has been obtained.




3 Hz spike and wave EEG in childhood absence epilepsy

***Molecular basis of absence seizures***. Aberrant thalamocortical rhythms underly the generalised 3Hz spike and wave discharges typical of absence seizures. Consciousness requires tonic firing of thalamocortical rhythms and abolition of rhythmic burst firing. Several classes of gene are involved in the generation of thalamocortical rhythms and absence seizures, and are thus candidate genes for CAE. These include voltage-dependent calcium channels, $GABA_A$ and $GABA_B$ receptors, potassium channels, sodium channels and hyperpolarisation-activated cyclic-nucleotide gated channels.

***Call for Families***

The power of a genetic study of this type is largely dependent on the size of the patient resource. We have now recruited 87 families with at least one child with CAE from centres around the UK and abroad, but are aiming for up to 300 families. We are recruiting new families with a proband with CAE on an ongoing basis.
These families include:

- trios (a single affected child with both unaffected parents)
- affected sibling pairs
- pedigrees with two or more affected members

**Do you treat children with childhood absence epilepsy?**
**Would you be interested in collaborating?**
**Please contact Robert Robinson (details above).**

## Appendix 5 Collaborators outside the UK

| Collaborator | Department |
|---|---|
| **Europe:** | |
| Professor Jean Aicardi | Hopital Robert Debré, Paris, France |
| Dr M Feucht, Professor H Aschauer | University Hospital for Child and Adolescent Neuropsychiatry, Vienna, Austria |
| Dr PMC Tijink-Callenbach, Professor OF Brouwer | Leiden University Medical Centre, Leiden, The Netherlands |
| Dr A Covanis | The Children's Hospital Agia Sophia, Goudi, Athens, Greece |
| Professor Olivier Dulac | Hopital Necker Enfants Malades, Paris, France |
| Professor Orvar Eeg-Olofsson | Uppsala University, Sweden |
| Dr Auli Siren, Dr Mogens Laue Friis | University Hospital of Tampere, Tampere, Finland |
| Dr Françoise Goutières | Hôpital des Enfants Malades, Paris, France |
| Professor Renzo Guerrini | University of Pisa, Italy |
| Dr Armin Heils | University Clinic of Epileptology, Bonn, Germany |
| Dr Marianne Juel Kjeldsen | Odense University Hospital, Odense, Denmark |
| Professor Anna-Elina Lehesjoki | Folkhälsan Institute of Genetics & Dept of Medical Genetics, Haartman Institute, University of Helsinki, Finland |
| Dr Thomas Sander | University Hospital Charité, Virchow Clinic, Berlin, Germany |
| Dr Anders Sundquist | Soder Hospital, Stockholm, Sweden |
| **Canada:** | |
| Professor JM Dooley | Dalhousie University and IWK Health Centre, Halifax, Nova Scotia |
| Dr Elaine Wirrell | Alberta Childrens Hospital, Calgary, Canada |

## Appendix 6 Letter of invitation to potential participants

Royal Free and University College Medical School
**UNIVERSITY COLLEGE LONDON**

DEPARTMENT OF PAEDIATRICS

Gower Street Campus
The Rayne Institute
5 University Street
London WC1E 6JJ

Direct Line Telephone 0207 679 6131
Fax 0207 679 6103

«Title» «AASurname» and «Child»
«Address1»
«Address2»
«Address3»
«Post_code»

Dear «Title» «AASurname» and «Child», <<Date>>

**Re: The genetic analysis of absence epilepsy**

«Consultant» suggested we write to you about our study, and we very much hope you will agree to participate. At University College London, research is currently progressing in the molecular genetics of epilepsy. We are interested in the genetics of individuals who have, or have had, mainly absence seizures, and are ascertaining children with absence epilepsy.

40% of all epilepsies are genetic, even when no other family members have epilepsy. Our aim is to identify which genes in the body are different between those affected and those not affected with absence epilepsy. Identifying abnormal genes may allow us to develop more effective treatments and improve diagnostic and prognostic methods.

We have enclosed some more information but will gladly discuss matters further and answer any questions you may have. If you agree to take part, all that is required are some cheek swabs from «Child», and if possible mum, dad and any brothers or sisters, all of which can be done at home. No extra hospital visits are required.

All information collected is confidential. We will keep you informed about the general results of the research project. If a genetic test is developed which is of proven benefit, we will provide guidance on how testing can be obtained in a clinical setting with appropriate counselling. If at any stage individuals decide that they no longer want to participate in the trial all their details will be removed immediately and there will be no interference with their medical management
.
If you are interested in the project then please return the slip below, or simply contact us by phone or e-mail.

Yours sincerely,

Pauline Boland
Clinical Research Nurse

Dr Robert Robinson
Clinical Research Fellow in Paediatrics

---

We would like to know more about the research project on genetic analysis of childhood absence epilepsy

Telephone No: ……..
email: …….
Contact Times……
«Title» «AASurname» and «Child».

## Appendix 7 Family information sheets

### Appendix 7a Information sheet – young children

Royal Free and University College Medical School
UNIVERSITY COLLEGE LONDON

DEPARTMENT OF PAEDIATRICS

Gower Street Campus
The Rayne Institute
5 University Street
London WC1E 6JJ

Direct Line Telephone 0207 679 6131
Fax 0207 679 6103

**INFORMATION SHEET FOR CHILDREN UP TO 12 YEARS PARTICIPATING IN EPILEPSY STUDY**

**We are conducting a study of epilepsy in families, and invite you to take part.**

Your genes make up a special code that is in every cell in your body, which you get partly from your mother and partly from your father. This code stores all the information that makes you "you".

The project is looking at the genes of people with epilepsy, and the genes of their brothers and sisters, and parents. This is to discover if there is part of the code that makes some people more likely to develop epilepsy.

This is important as it may help us to understand more about epilepsy, and hopefully develop better drugs to treat people with epilepsy.

If you agree to take part in the study, you, your mum and dad, and your brothers and sisters will be contacted by a doctor. The doctor will ask some questions about your health (how you feel), and take a cheek swab or a small blood sample from each of you. A special cream will mean that you hardly feel a thing if a blood sample is taken.

We will look at a special part of your sample, which contains your genes (the special code). We will compare it to the codes of people in your family, and codes from other families who are also taking part in the study.

This will allow us to find out if there is a special part of the code that carries information about epilepsy.

The results of the study may be published in a medical journal. Your name would never be given out to anyone unless you said it was OK.

**You do not have to take part in this study if you do not want to. If you decide to take part you can stop at any time, without having to give a reason. Whether you decide to take part or not will not affect the way you are looked after normally by your doctors.**

**FOR FURTHER INFORMATION CONTACT: Dr Robert Robinson** at the above address.

An ethics committee looks at all studies involving people before they can go ahead, to make sure that everything is fair and done for a good reason. This study was approved by the North Thames Multi-centre Research Ethics Committee.
The ethics committee may need to check your medical records to make sure all our information is correct. However no details about you will be passed on to anyone else.

Head of Department *Professor RM Gardiner MD, FRCP*

## Appendix 7b Information sheet – teenagers

Royal Free and University College Medical School
UNIVERSITY COLLEGE LONDON

DEPARTMENT OF PAEDIATRICS

Gower Street Campus
The Rayne Institute
5 University Street
London WC1E 6JJ

Direct Line Telephone 0207 679 6131
Fax 0207 679 6103

**INFORMATION SHEET FOR TEENAGERS PARTICIPATING IN EPILEPSY STUDY**

**We are conducting a study of epilepsy in families, and invite you to take part.**

Your genes make up a special code that is in every cell in your body, which you get partly from your mother and partly from your father. This code stores all the information that makes you "you".

The project is looking at the genes of people with epilepsy, and the genes of their brothers and sisters, and parents. This is to discover if there is part of the code that makes some people more likely to develop epilepsy.

This is important as it may help us to understand more about epilepsy, and hopefully develop better drugs to treat people with epilepsy.

If you agree to take part in the study, you, your mum and dad, and your brothers and sisters will be seen by a doctor. The doctor will ask you some questions about your health (how you feel), and take some cheek swabs or a blood sample (5-20ml) from each of you.

We will look at a special part of your sample, which contains the DNA which makes up your genes. We will compare it to the genes of people in your family, and genes from other families who are also taking part in the study.

This will allow us to find out if there is a special gene that carries information about epilepsy.

The results of the study may be published in a medical journal. Your name would never be identified from published results without your agreeing to this.

**You do not have to take part in this study if you do not want to. If you decide to take part you may withdraw at any time, without having to give a reason. Your decision whether to take part or not will not affect your care and management in any way.**

**FOR FURTHER INFORMATION CONTACT: Dr Robert Robinson** at the above address.

An ethics committee reviews all proposals for research involving people before they can proceed. This proposal was approved by the North Thames Multi-centre Research Ethics Committee.
The ethics committee and regulatory authorities may need to see your medical records to make sure all our information is correct. However your medical details will not be seen by anyone else.

Head of Department *Professor RM Gardiner MD, FRCP*

## Appendix 7c Information sheet – adults

Royal Free and University College Medical School
UNIVERSITY COLLEGE LONDON

DEPARTMENT OF PAEDIATRICS

Gower Street Campus
The Rayne Institute
5 University Street
London WC1E 6JJ

Direct Line Telephone 0207 679 6131
Fax 0207 679 6103

**INFORMATION SHEET FOR ADULTS PARTICIPATING IN EPILEPSY STUDY**

**We are conducting a study of epilepsy in families, and invite you to take part.**

**WHY IS THE STUDY BEING DONE**? The aim of our study is to find genes that make certain individuals susceptible to having seizures. Once genes are identified we can determine how they work and discover exactly what goes wrong when a seizure occurs.

**HOW IS THE STUDY TO BE DONE?** The first step in identifying genes is to find a large number of families in which there is one or more family member with epilepsy. We ask family members for descriptions of their seizures, at what age they started, how long they last, how often they occur, and whether they had any serious illnesses, head injuries or birth difficulties. If we then think that we have enough information to be able to include them in our study, we will ask for some cheek swabs or a blood sample (5-20ml) from as many family members as possible, both affected and unaffected by epilepsy. If required, the blood sample will be taken at a routine outpatient appointment or by your GP (we will not contact your GP without your permission). From the samples we will extract DNA, the material of which genes are made. Many thousands of genes exist, and these are grouped together to form chromosomes. Humans have 23 pairs of chromosomes (46 in total). We will analyse the DNA to see if we can first work out on which chromosomes genes causing epilepsy are to be found, and then to work out exactly what type of genes they are. The results of the study may be published in a medical journal. Your name would never be identified from published results without your consent.

**WHAT ARE THE POSSIBLE BENEFITS OF THE STUDY?** By identifying genes that cause epilepsy, we can work out what goes wrong to cause a seizure and why some people have seizures and others do not. Then, hopefully, we would be able to develop better drugs to treat epilepsy, and even prevent epilepsy developing at all. However all these developments are a long way off, and it may be many years before we really understand what causes epilepsy.

**WHAT ARE THE RISKS OF THE STUDY?** There are no real risks associated with participation in the study.

**You do not have to take part in this study if you do not want to. If you decide to take part you may withdraw at any time, without having to give a reason. Your decision whether to take part or not will not affect your care and management in any way.**

**FOR FURTHER INFORMATION CONTACT: Dr Robert Robinson** at the above address.

An ethics committee reviews all proposals for research using human subjects before they can proceed. This proposal was approved by the North Thames Multi-centre Research Ethics Committee.
The ethics committee and regulatory authorities may require access to your medical records for purposes of verification. However this will not violate the confidentiality of any medical information.

Head of Department *Professor RM Gardiner MD, FRCP*

## Appendix 8 Consent form

Royal Free and University College Medical School
UNIVERSITY COLLEGE LONDON

DEPARTMENT OF PAEDIATRICS

Gower Street Campus
The Rayne Institute
5 University Street
London WC1E 6JJ

Direct Line Telephone 0207 679 6131
Fax 0207 679 6103

**CONFIDENTIAL**

**CONSENT FORM FOR FAMILIES PARTICIPATING IN EPILEPSY STUDY**

**Please answer all the questions below before agreeing to take part in the study.**

| | |
|---|---|
| Have you read the information sheet about this study? | YES/NO |
| Have you had an opportunity to ask questions and discuss this study? | YES/NO |
| Have you received enough information about this study? | YES/NO |
| Have you received satisfactory answers to all your questions? | YES/NO |

Which doctor/nurse have you spoken to about this study?……

Do you understand that
you are free to withdraw from this study at any time
without giving a reason for withdrawing
without affecting you future medical care? YES/NO

Do you agree to take part in this study? YES/NO

Signature of
volunteer……..

Name (please
print)……Date of Birth…….

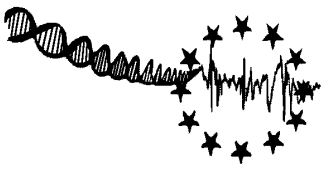
Relationship to child (if under 16 years
old)……

Signature of
Investigator……...

Name (please
print)……

Head of Department *Professor RM Gardiner MD, FRCP*

# Genetic Analysis of Absence Epilepsy



# Family Clinical Data Set

## Instructions for completing the family clinical data sets

One clinical data set should be completed for each family.

Each set comprises:

**Sheet one:** **Family name/code. Referring physicians' details.**

**Sheet two:** **Family overview**
- please enter the local code for *each different family member*

**Sheet three:** **Pedigree diagram**
- instructions for drawing pedigrees below

**Clinical documentation sets for 3 individuals:**

Please complete one clinical documentation set for:
a) Each affected family member.
b) Each asymptomatic family member on whom an EEG has been performed.

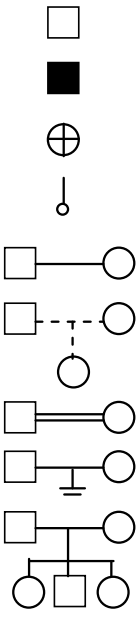
Each set comprises:

**Sheet one:** **Clinical information**
- please give as full clinical information as possible

**Sheet two:** **Seizure descriptions**
- boxes are provided for 3 seizure types (absences and two others)

**Sheet three:** **EEG data**
- please include most informative EEGs

**Symbols for drawing pedigrees:**
- underneath each affected person indicate the diagnosis if known.



Normal male

Affected male

Stillbirth

Abortion

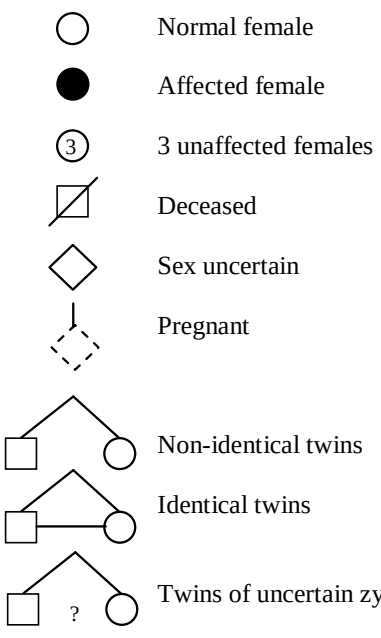
Marriage

Illegitimate offspring

Consanguinous marriage

No offspring

Marriage with 3 children




Normal female

Affected female

3 unaffected females

Deceased

Sex uncertain

Pregnant

Non-identical twins

Identical twins

Twins of uncertain zygosity

**Genetic Analysis of Absence Epilepsy**

**Family Clinical Data Set**

Family name:

(may be left blank)

Local code:

**REFERRING PHYSICIAN 1**

Name:

Address:

Telephone:

Fax:

Email:

**REFERRING PHYSICIAN 2**

Name:

Address:

Telephone:

Fax:

Email:

## Family Overview

SUBJECTS

| Name | Local code | Date of birth D M Y | Sex | Date of last Observation D M Y | EEG |
|---|---|---|---|---|---|
| | | | ♂ ♀ | | Y N |
| | | | ♂ ♀ | | Y N |
| | | | ♂ ♀ | | Y N |
| | | | ♂ ♀ | | Y N |
| | | | ♂ ♀ | | Y N |
| | | | ♂ ♀ | | Y N |
| | | | ♂ ♀ | | Y N |
| | | | ♂ ♀ | | Y N |
| | | | ♂ ♀ | | Y N |
| | | | ♂ ♀ | | Y N |

EPILEPSIES

- Childhood Absence Epilepsy (Petit Mal)
- Juvenile Absence Epilepsy
- OTHER

**Pedigree Structure**

Enter epilepsy diagnostic and local codes against individuals

## CLINICAL DOCUMENTATION SET

**Subject name:** | **Local Code:**

**Mother's Code:** | **Father's Code:**

Epilepsy: CAE | JAE | Other

### Clinical Information

History: **Normal?** **YES** | **NO** **Give details below:**

Examination: **Normal?** **YES** | **NO** **Give details below:**

Development: **Normal?** **YES** | **NO** **Give details below:**

Imaging: **NO** | **YES** **Normal?** **YES** | **NO** **Give details below:**

Anti-epileptic drugs used and response to treatment:

Local Subject Code:

## Seizures

**Absence Seizures:** Typical? ☐ Atypical? ☐ Eyelid myoclonia? ☐

Age of onset: ☐ Age of last seizure: ☐ Total number: 1 ☐ 2-6 ☐ >6 ☐

Average duration: ☐ mins ☐ secs Frequency: ☐ per yr ☐ mth ☐ wk ☐ day ☐ hr ☐

Timing: Anytime ☐ In sleep ☐ Awake ☐ On waking ☐ Evening resting ☐

Provoking Factors:

Description:

**Other Seizures:** Type: ☐

Age of onset: ☐ Age of last seizure: ☐ Total number: 1 ☐ 2-6 ☐ >6 ☐

Average duration: ☐ mins ☐ secs Frequency: ☐ per yr ☐ mth ☐ wk ☐ day ☐ hr ☐

Timing: Anytime ☐ In sleep ☐ Awake ☐ On waking ☐ Evening resting ☐

Provoking Factors:

Description:

**Other Seizures:** Type: ☐

Age of onset: ☐ Age of last seizure: ☐ Total number: 1 ☐ 2-6 ☐ >6 ☐

Average duration: ☐ mins ☐ secs Frequency: ☐ per yr ☐ mth ☐ wk ☐ day ☐ hr ☐

Timing: Anytime ☐ In sleep ☐ Awake ☐ On waking ☐ Evening resting ☐

Provoking Factors:

Description:

Local Subject Code:

## EEG data

Date of EEG: Age at recording: ☐ yrs ☐ mths

Medication: N ☐ Y ☐ Specify_____

Patient state: Awake ☐ Sleep-deprived ☐ Sleep ☐ Inter-ictal ☐ Ictal ☐

Background: Normal ☐ Abnormal ☐ Give details:

_____

Abnormalities: Spontaneous ☐ Provoked by: hyperventilation ☐ photic stimulation ☐

Generalised spike and wave discharges ☐ Frequency ☐

Other:_____

Video Telemetry:_____

Conclusion: _____

_____

_____

Date of EEG: Age at recording: ☐ yrs ☐ mths

Medication: N ☐ Y ☐ Specify_____

Patient state: Awake ☐ Sleep-deprived ☐ Sleep ☐ Inter-ictal ☐ Ictal ☐

Background: Normal ☐ Abnormal ☐ Give details:

_____

Abnormalities: Spontaneous ☐ Provoked by: hyperventilation ☐ photic stimulation ☐

Generalised spike and wave discharges ☐ Frequency ☐

Other:_____

Video Telemetry:_____

Conclusion: _____

_____

_____

**Please Return the completed forms to :**

**Dr Robert Robinson**
**Department of Paediatrics**
**Royal Free and University College London Medical School**
**Gower Street Campus**
**5 University Street**
**London**
**WC1E 6JJ**
**UK**

**Tel: 44 (0)20 7679 6131**
**Fax: 44 (0)20 7679 6103**
**Email: robert.robinson@ucl.ac.uk**

## Appendix 10 Instructions for entering data into clinical database

**Entering data into Access database**

Open database in Microsoft Access.
Password #######
Ensure 'Forms' is selected in left hand menu (Objects)
Enter data in each of the three forms
Within each form, to scroll through previous records use arrows at bottom of screen
To delete a record, in 'edit' menu click 'select record', then 'delete'

**1. Family Investigator**

- One record per family
- Family ID is a 3-figure number (numbers to be allocated)
- Epilepsies – this represents all the epilepsies/seizures in the family, each represented by a 2-letter code (eg CAFCIG). Select from drop-down list or, if correct combination of epilepsies is not present in list, type directly into box.
    - CA: CAE
    - JA: JAE
    - JM: JME
    - IG: IGE
    - EM: Eyelid myoclonis with absence
    - FC: Febrile convulsions
    - PS: Photosensitive seizures
    - OE: Other epilepsy

**2. Family Overview**

- One record for every individual in the family (***with*** or ***without*** seizures)
- Full ID is a 6-figure number – a combination of the family ID and the individual ID. The usual individual IDs are: father – 101, mother – 102, children 201, 202, 203 etc.
- Epilepsies – this represents all epilepsies/seizures for that individual.
- Relationship: P – proband, PP – parent of proband, SP – sib of proband. More complicated pedigrees must be drawn by hand!

**3. Clinical Data**

- One record for each individual ***with*** epilepsy/seizures
- It may be necessary to maximise the clinical data window to access the scroll bar on the right
- Full ID – as in Family Overview
- Clinical History – use Clinical Comments box to give details of any abnormalities
- Seizures: Up to five different seizure types (three in detail)
- Seizure types:
    - GSA: generalised seizure, absence
    - GSM: generalised seizure, myoclonic
    - GTCS: generalised tonic-clonic seizure
    - FC: febrile convulsion
    - GSAmy: generalised seizure, absence + myoclonus
    - OS: other seizure
- Use Seizure Comments box for any additional details
- EEG: Use EEG Comments box for additional details

## Appendix 11 Sample collection

Appendix 11a Instructions for blood sample collection and transportation

**Royal Free and University College Medical School**
**UNIVERSITY COLLEGE LONDON**

DEPARTMENT OF PAEDIATRICS

Gower Street Campus
The Rayne Institute
5 University Street
London WC1E 6JJ

Direct Line Telephone 0207 679 6131
Fax 0207 679 6103

### Collection of blood samples for DNA analysis

| | |
|---|---|
| Volume: | 5-20 mls<br>(but even 1-2 mls is sufficient on small infants) |
| Anti-coagulant: | Ca – EDTA<br>Tubes used for full blood count estimations are fine and usually have a purple top (in the UK). |
| Storage: | Room temperature.<br>It is not necessary to freeze or put at 4°C |
| Labeling: | Ensure bottles labeled including full name and date of birth. |
| Transfer: | Samples need to arrive in the UK not more than three days from the time of sampling. Samples should be shipped to the following address: |

**Dr Robert Robinson**
**Department of Paediatrics**
**Royal Free and University College Medical School**
**The Rayne Institute**
**5 University Street**
**London**
**WC1E 6JJ**
**UK**

**Tel: 44 (0)20 7679 6131**
**Fax: 44 (0)20 7679 6103**
**Email: robert.robinson@ucl.ac.uk**

Consent forms should be included with the samples. Shipping costs will be reimbursed.
Thank you for your time and effort.

Head of Department *Professor RM Gardiner MD, FRCP*

Appendix 11b Instructions for cheek swab sample collection and transportation

**Royal Free and University College Medical School**
UNIVERSITY COLLEGE LONDON

DEPARTMENT OF PAEDIATRICS

Gower Street Campus
The Rayne Institute
5 University Street
London WC1E 6JJ

Direct Line Telephone 0207 679 6131
Fax 0207 679 6103

## Collection of DNA samples by mouth swab method

Dr Robert Robinson
Department of Paediatrics
Rayne Institute
5 University Street
London WC1E 6JJ
UK
Tel 44 (0)20 7679 6131
Fax 44 (0)20 7679 6103
e-mail robert.robinson@ucl.ac.uk

**Padded envelope contains (one set for each person)**

- Tube containing 5 swabs
- Tube marked P on top containing preservative solution, labelled with name
- Cotton wool or tissue to wrap tube in
- Gloves

**Procedure**

1. Each family member providing cheek swabs should take their swabs on the same day, or within a day or two of each other, so that all the swabs can be returned in the same envelope.
2. Collect sample first thing in the morning before drinking, eating or doing teeth.
3. Please wear gloves provided whenever handling swabs to avoid contamination.
4. Each person should use all 5 swabs on one morning.
5. Remove each swab from tube and rub along both cheeks inside mouth for about 20 seconds.
6. Place used swab in **CORRECTLY NAMED TUBE** containing preservative solution. **MAKE SURE THAT THE COTTON WOOL END IS INSERTED DOWNWARDS INTO THE LIQUID.**
7. All 5 swabs should fit in the tube with the preservative.
8. Be sure to do the lid up tightly once all five swabs are in the tube.
9. Make sure the swabs are covered in the solution.
10. The empty tubes can be discarded.
11. Wrap outside of swab tube in cotton wool/tissue provided and sellotape together.
12. Send back to Dr Robert Robinson in padded envelope provided with consent form.

Thank you for your time and effort.

**Appendix 12 Absence pedigree clinical data**

CAE: childhood absence epilepsy; JAE: juvenile absence epilepsy; AE: absence epilepsy; GSA: generalised absence seizure; GSAmy: generalised absence seizure with myoclonus; FS: febrile seizure; GSWD: generalised spike-wave discharges; GPSWD: generalised polyspike-wave discharges

Country of Origin: FR=France (Généthon);
UK=United Kingdom;
FI=Finland;
DK=Denmark;
DE=Germany;
IT=Italy;
SE=Sweden;
NL=The Netherlands;
GR=Greece;
AT=Austria.

| *Fam ID* | *Full ID* | *Sex* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality* | *Frequency (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 5204 | F | FR | CAE | GSA | 4 | | GSWD | 3 | ?mild eyelid myoclonus |
| 5 | 5207 | M | FR | CAE | GSA | 5 | | GSWD | 3 | |
| 9 | 9201 | F | UK | AE | GSAmy | 4 | | GPSWD | | |
| 9 | 9202 | F | UK | AE | GSA<br>GTCS | 2 | | GSWD, PS | | |

| *Fam ID* | *Full ID* | *Sex* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality* | *Frequency (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|---|
| 26 | 26202 | M | UK | CAE | GSA FS | 6 | | GSWD | 3 | Mild learning difficulties |
| 26 | 26204 | F | UK | CAE | GSA FS | 6 | | GSWD | 3 | Frequent paroxysms of bilaterally synchronous 3 sec spike&slow wave activity associated with absence |
| 27 | 27202 | F | UK | AE | GSA GTCS | 4 | | GSWD, PS | | Some generalised theta activity..Photic stimulation produced bursts of generalised spike&wave complexes |
| 27 | 27301 | F | UK | CAE | GSA GTCS | 7 | | GSWD, PS | 3 | GTCS. Loss of consciousness, generalised shaking of all limbs. No incontinence or tongue biting |
| 27 | 27302 | F | UK | CAE | GSA GTCS | 7 | | GSWD | 3 | |
| 92 | 92102 | F | UK | IGE | GTCS | | | | | |
| 92 | 92201 | F | UK | CAE | GSA | 2-12 | | GSWD | | |
| 92 | 92202 | F | UK | CAE | GSA | 2-12 | | GSWD | | |
| 96 | 96301 | F | UK | AE | GSA | 7 | | GPSWD | | Numerous generalised polyspike & slow waves sometimes associated with eyelid flutter and impaired awareness |
| 96 | 96302 | M | UK | AE | GSA | 6 | | GPSWD | 3-5 | Bursts of spike, polyspike activity. Twitch of the mouth evident |
| 137 | 137101 | M | UK | IGE | GTCS, MS | 18 | | GSWD | | |
| 137 | 137202 | F | UK | IGE | GTCS, MS | 12 | | GSWD | | |
| 137 | 137301 | F | UK | AE | GSAmy | 1 | | GPSWD | | |

| *Fam ID* | *Full ID* | *Sex* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality* | *Frequency (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|---|
| 137 | 137302 | M | UK | AE | GSAmy | 1 | | GPSWD | | |
| 173 | 173201 | F | UK | CAE | GSA | 6 | | SW | 3-3.5 | |
| 173 | 173202 | F | UK | CAE | GSA | 5 | | PSW | 3 | |
| 202 | 202101 | M | FI | CAE | GSA GTCS | 5 | | GSWD | | |
| 202 | 202201 | F | FI | CAE | GSA FS | 3 | | GSWD | | |
| 202 | 202202 | F | FI | CAE | GSA FS | 3 | | GSWD | | |
| 202 | 202203 | F | FI | CAE | GSA FS | 6 | | GSWD | | |
| 317 | 317201 | F | FI | CAE | GSA GTCS | 5 | | GSWD | 3 | |
| 317 | 317202 | M | FI | CAE | GSA | 5 | | GSWD | 3 - 2.5 | |
| 321 | 321101 | M | DK | CAE | GSA GTCS | 9 | | GSWD | 3 | |
| 321 | 321201 | F | DK | CAE | GSA | 6 | | GSWD | | |
| 321 | 321202 | F | DK | CAE | GSA | 3 | | GSWD | | |
| 321 | 321204 | F | DK | CAE | GSA FS | 3 | | GSWD | 3 | |
| 342 | 342201 | F | UK | AE | GSAmy | 11 | | GPSWD | | |

| *Fam ID* | *Full ID* | *Sex* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality* | *Frequency (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|---|
| 342 | 342202 | F | UK | AE | GSAmy | <13 | | GPSWD | | |
| 346 | 346201 | F | UK | CAE | GSA | 2 | | GSWD | 3 | |
| 346 | 346202 | F | UK | CAE | GSA | 2-12 | | GSWD | 3 | mild visuo-perceptual problems |
| 348 | 348203 | M | UK | CAE | GSA | 3 | | GSWD | | |
| 348 | 348206 | M | UK | CAE | GSA | 3 | | GSWD | 3 | absences causing some interference with school work |
| 353 | 353101 | M | UK | CAE | GSA | 10 | | GSWD | | |
| 353 | 353201 | M | UK | CAE | GSA | 6 | Y | GSWD | 3 | aged 7/12 suspected sepsis, septic screen NAD |
| 353 | 353202 | M | UK | CAE | GSA | 9 | | GSWD | 3 | |
| 358 | 358102 | F | UK | CAE | GSA | 4 | | GSWD | | |
| 358 | 358201 | F | UK | CAE | GSA | 5 | | GSWD | 3 | Definite abnormality of spike & wave activity |
| 372 | 372102 | F | DE | CAE | GSA GTCS | 6 | | GSWD | 3 | |
| 372 | 372201 | F | DE | CAE | GSA | 4 | | GSWD | 3 | |
| 372 | 372202 | F | DE | CAE | GSA GTCS | 5 | | GSWD | 3 | |
| 381 | 381101 | M | DE | CAE | GSA GTCS | 2-12 | | GSWD | | |
| 381 | 381201 | F | DE | CAE | GSA | 5 | | GSWD | | |
| 381 | 381202 | M | DE | CAE | GSA | 5 | | GSWD | | |
| 381 | 381203 | M | DE | CAE | GSA | 6 | | GSWD | | |
| 382 | 382102 | F | DE | | | | | GSWD | 3 | No definite seizures |
| 382 | 382201 | M | DE | CAE | GSA | 7 | | GSWD | | |

| *Fam ID* | *Full ID* | *Sex* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality* | *Frequency (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|---|
| 382 | 382202 | F | DE | CAE | GSA | 6 | | GSWD | 3 | |
| 393 | 393201 | F | DE | CAE | GSA | 8 | | GSWD | | |
| 393 | 393202 | M | DE | JAE | GSA GTCS | 15 | | GPSWD | | |
| 394 | 394301 | M | DE | CAE | GSA | 7 | | GSWD | | |
| 394 | 394302 | M | DE | CAE | GSA | 8 | | GSWD | | |
| 395 | 395102 | F | DE | | | | | GSWD | | No definite seizures |
| 395 | 395201 | M | DE | CAE | GSA | 4 | | GSWD | | |
| 395 | 395203 | M | DE | CAE | GSA | 3 | | GSWD | | |
| 396 | 396101 | M | DE | CAE | GSA | | | GSWD | | |
| 396 | 396201 | F | DE | AE | GSA | 3 | | GPSWD | | |
| 396 | 396202 | M | DE | CAE | GSA | 6 | | GSWD | | |
| 397 | 397102 | F | DE | CAE | GSA | 5 | | GSWD | | |
| 397 | 397201 | F | DE | CAE | GSA GTCS | 3 | | GSWD | | |
| 397 | 397202 | F | DE | CAE | GSA | 3 | | GSWD | | |
| 400 | 400201 | M | DE | AE | GSA GTCS | 5 | | GPSWD | | |
| 400 | 400202 | F | DE | CAE | GSA GTCS | 5 | | GSWD | | |
| 401 | 401101 | F | DE | IGE | GTCS | 20 | | | | |
| 401 | 401201 | M | DE | CAE | GSA | 4 | | GSWD | | |

| *Fam ID* | *Full ID* | *Sex* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality* | *Frequency (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | GTCS | | | | | |
| 401 | 401202 | F | DE | CAE | GSA GTCS | 3 | | GSWD | | |
| 402 | 402201 | F | DE | CAE | GSA GTCS | 12 | | GSWD | | |
| 402 | 402202 | F | DE | CAE | GSA | 11 | | GSWD | | |
| 403 | 403201 | F | DE | CAE | GSA GTCS | 3 | | GSWD | | |
| 403 | 403202 | M | DE | CAE | GSA FS | 3 | | GSWD | | |
| 404 | 404101 | M | DE | IGE | GTCS | 20 | | | | |
| 404 | 404202 | F | DE | CAE | GSA GTCS | 6 | | GSWD | | |
| 404 | 404302 | F | DE | CAE | GSA | 7 | | GSWD | | |
| 404 | 404303 | M | DE | CAE | GSA | 7 | | GSWD | | |
| 404 | 404304 | M | DE | CAE | GSA | 5 | | GSWD | | |
| 405 | 405201 | M | DE | CAE | GSA | 6 | | GSWD | | |
| 405 | 405202 | M | DE | CAE | GSA | 3 | | GSWD | | |
| 407 | 407102 | F | UK | CAE | GSA | 5 | | GSWD | | Separation anxiety aged 4 years. ?jerks limbs started after first pregnancy, stopped spontaneously |
| 407 | 407201 | F | UK | CAE | GSA | 3 | | GSWD | 3 | Absence attacks associated with 3 second spike and wave. 3+ years behavioural problems |
| 407 | 407202 | F | UK | CAE | GSA | 5 | | GSWD | 3 | |

| *Fam ID* | *Full ID* | *Sex* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality* | *Frequency (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|---|
| 421 | 421201 | F | IT | CAE | GSA | 2-12 | | GSWD | | |
| 421 | 421202 | F | IT | CAE | GSA | 2-12 | | GSWD | | |
| 422 | 422103 | M | IT | IGE | GTCS | 9 | | | | |
| 422 | 422201 | F | IT | CAE | GSA | 6 | | GSWD | | |
| 422 | 422202 | M | IT | CAE | GSA | 13 | | SW | | |
| 426 | 426103 | F | UK | AE | GSA | <13 | | GPSWD | | Paternal aunt thought to have absence attacks although never confirmed on EEG |
| 426 | 426201 | F | UK | CAE | GSA | 6 | | GSWD | 3 | Hyperventilation provides 3 episodes of generalised spike & wave, associated with clinical absences. Maternal Toxoplasmosis during pregnancy. Seizures did affect schooling but no extra help required |
| 429 | 429101 | M | FR | AE | GSA | <13 | | GPSWD | | |
| 429 | 429203 | F | FR | AE | GSA | <13 | | GPSWD | | |
| 430 | 430202 | F | FR | AE | GSA | <13 | | GPSWD | | |
| 430 | 430301 | M | FR | AE | GSA | <13 | | GPSWD | | |
| 431 | 431201 | F | FR | JAE | GSA FS | 13+ | | GPSWD | | |
| 431 | 431202 | F | FR | JAE | GSA FS | 13+ | | GPSWD | | |
| 432 | 432201 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 432 | 432202 | F | FR | CAE | GSA | 2-12 | | GSWD | | |

| *Fam ID* | *Full ID* | *Sex* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality* | *Frequency (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|---|
| 433 | 433202 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 433 | 433301 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 434 | 434201 | M | FR | CAE | GSA | 2-12 | | GSWD | | |
| 434 | 434202 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 435 | 435201 | M | FR | CAE | GSA | 2-12 | | GSWD | | |
| 435 | 435202 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 436 | 436201 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 436 | 436202 | M | FR | CAE | GSA | 2-12 | | GSWD | | |
| 437 | 437201 | M | FR | CAE | GSA | 2-12 | | GSWD | | |
| 437 | 437202 | M | FR | CAE | GSA GTCS | 2-12 | | GSWD | | |
| 438 | 438201 | F | FR | JAE | GSA | 13+ | | GPSWD | | |
| | | | | | GTCS | | | | | |
| 438 | 438301 | F | FR | JAE | GSA | 13+ | | GPSWD | | |
| | | | | | GTCS | | | | | |
| 438 | 438302 | M | FR | JAE | GSA | 13+ | | GPSWD | | |
| | | | | | GTCS | | | | | |
| 439 | 439201 | F | FR | JAE | GSA | | | GPSWD | | |
| 439 | 439202 | F | FR | JAE | GSA | | | GPSWD | | |
| | | | | | GTCS | | | | | |
| 440 | 440202 | F | FR | CAE | GSA | | | SW | | |

| *Fam ID* | *Full ID* | *Sex* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality* | *Frequency (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|---|
| 440 | 440203 | M | FR | JAE | GSA | | | SW | | |
| 441 | 441202 | F | FR | CAE | GSA | | | SW | | |
| 441 | 441203 | M | FR | JAE | GSA | | | SW | | |
| 442 | 442201 | F | FR | CAE | GSA GTCS | 2-12 | | GSWD | | |
| 442 | 442202 | M | FR | CAE | GSA GTCS | 2-12 | | GSWD | | |
| 443 | 443202 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 443 | 443203 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 444 | 444201 | M | FR | AE | GSA | <13 | | GPSWD | | |
| 444 | 444202 | M | FR | AE | GSA | <13 | | GPSWD | | |
| 445 | 445202 | M | FR | CAE | GSA FS | 2-12 | | GSWD | | |
| 445 | 445203 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 446 | 446201 | M | FR | CAE | GSA | 2-12 | | GSWD | | |
| 446 | 446202 | M | FR | CAE | GSA | 2-12 | | GSWD | | |
| 447 | 447201 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 447 | 447202 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 448 | 448201 | M | FR | AE | GSA FS | <13 | | GPSWD | | |
| 448 | 448202 | M | FR | AE | GSA | <13 | | GPSWD | | |
| 449 | 449201 | F | FR | JAE | GSA | 13+ | | GPSWD | | |

| *Fam ID* | *Full ID* | *Sex* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality* | *Frequency (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | GTCS | | | | | |
| 449 | 449202 | F | FR | JAE | GSA | 13+ | | GPSWD | | |
| 450 | 450201 | M | FR | CAE | GSA | 2-12 | | GSWD | | |
| 450 | 450202 | M | FR | CAE | GSA | 2-12 | | GSWD | | |
| 451 | 451202 | F | FR | JAE | GSA | 13+ | | | | |
| | | | | | FS | | | | | |
| 451 | 451203 | F | FR | JAE | GSA | 13+ | | | | |
| 452 | 452995 | M | FR | AE | GSA | <13 | | GPSWD | | |
| 452 | 452996 | F | FR | AE | GSA | <13 | | GPSWD | | |
| 461 | 461201 | F | SE | CAE | GSA | 2-12 | | GSWD | | |
| 461 | 461202 | M | SE | CAE | GSA | 2-12 | | GSWD | | |
| 466 | 466201 | F | SE | CAE | GSA | 2-12 | | GSWD | | |
| 466 | 466202 | F | SE | CAE | GSA | 2-12 | | GSWD | | |
| 525 | 525201 | F | NL | AE | GSA | 10 | | GPSWD | | |
| 525 | 525202 | F | NL | AE | GSA | 12 | | GSWD | | |
| | | | | | GTCS | | | | | |
| 525 | 525203 | F | NL | CAE | GSA | 6 | | GSWD | | |
| 527 | 527201 | F | NL | CAE | GSA | 7 | Y | GSWD | | |
| 527 | 527202 | F | NL | CAE | GSA | 3 | Y | GSWD | 3-4 | |
| 528 | 528101 | M | NL | IGE | GSA | 11 | | | | |
| | | | | | GTCS | | | | | |

| *Fam ID* | *Full ID* | *Sex* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality* | *Frequency (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|---|
| 528 | 528204 | F | NL | AE | GSA | 5 | | GPSWD | | |
| 528 | 528301 | M | NL | IGE | GSA<br>GTCS | 2 | | Non-specific abnormalities | | |
| 528 | 528302 | M | NL | AE | GSA | 1 | | GSWD | | |
| 529 | 529202 | F | NL | CAE | GSA | 7 | | GSWD | | |
| 529 | 529301 | M | NL | CAE | GSA | 5 | | GSWD | 2-3 | Absences - primary generalised epilepsy |
| 530 | 530201 | M | NL | CAE | GSA | 5:5 | | GSWD | 3 | |
| 530 | 530202 | M | NL | CAE | GSA | 4 | Y | GSWD | 2.5-3 | Frontal sharp waves and spikes |
| 577 | 577201 | M | GR | CAE | GSA | 10 | | GSWD | | |
| 577 | 577202 | F | GR | CAE | GSA | 7 | | GSWD | | |
| 586 | 586202 | F | NL | CAE | GSA | 5 | | GSWD | | GTCS often preceded by headaches |
| 586 | 586301 | F | NL | AE | GSAmy | 5 | | GPSWD | | |
| 586 | 586302 | M | NL | CAE | GSA | 7.5 | | GSWD | | |
| 871 | 871201 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 871 | 871202 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 872 | 872201 | M | FR | CAE | GSA | 2-12 | | GSWD | | |
| 872 | 872203 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 873 | 873101 | M | FR | JAE | GSA | 13+ | | GPSWD | | |
| 873 | 873201 | F | FR | JAE | GSA | 13+ | | GPSWD | | |
| 873 | 873202 | F | FR | JAE | GSA | 13+ | | GPSWD | | |
| 874 | 874102 | F | FR | CAE | GSA | 2-12 | | GSWD | | |

| *Fam ID* | *Full ID* | *Sex* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality* | *Frequency (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|---|
| 874 | 874201 | M | FR | CAE | GSA | 2-12 | | GSWD | | |
| 874 | 874202 | F | FR | IGE | GTCS | 2-12 | | | | |
| 875 | 875201 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 875 | 875301 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 875 | 875302 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 876 | 876201 | M | FR | JAE | GSA | 13+ | | GPSWD | | |
| 876 | 876202 | F | FR | JAE | GSA | 13+ | | GPSWD | | |
| 877 | 877201 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 877 | 877202 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 878 | 878201 | M | FR | CAE | GSA | 2-12 | | GSWD | | |
| 878 | 878202 | M | FR | CAE | GSA | 2-12 | | GSWD | | |
| 879 | 879102 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| | | | | | FS | | | | | |
| 879 | 879103 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| | | | | | FS | | | | | |
| 879 | 879201 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 880 | 880201 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 880 | 880202 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| | | | | | GTCS | | | | | |
| | | | | | FS | | | | | |
| 881 | 881201 | M | FR | JAE | GSA | 13+ | | GPSWD | | |

| *Fam ID* | *Full ID* | *Sex* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality* | *Frequency (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | FS | | | | | |
| 881 | 881202 | F | FR | JAE | GSA | 13+ | | GPSWD | | |
| 881 | 881203 | M | FR | JAE | GSA | 13+ | | GPSWD | | |
| | | | | | GTCS | | | | | |
| 882 | 882201 | M | FR | JAE | GSA | 13+ | | GPSWD | | |
| 882 | 882202 | M | FR | JAE | GSA | 13+ | | GPSWD | | |
| 883 | 883201 | F | FR | JAE | GSA | 13+ | | GPSWD | | |
| 883 | 883202 | F | FR | JAE | GSA | 13+ | | GPSWD | | |
| 884 | 884201 | F | FR | CAE | GSA | 13+ | | GPSWD | | |
| | | | | | GTCS | | | | | |
| 884 | 884202 | F | FR | JAE | GSA | 13+ | | GPSWD | | |
| 885 | 885102 | F | FR | CAE | GSA | 2-12 | | GSWD | | |
| 885 | 885201 | F | FR | CAE | GSA | 2-12 | | GSWD | | |

## Appendix 13 Absence trios clinical data

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| 126 | 126302 | UK | CAE | GSA GTCS | 8 | | GSWD | | Hyperventilation caused build up of slow waves/spikes. GTCS, making a funny noise.Rigid and shaking all over.speech difficulties post& shaky on feet. |
| 128 | 128101 | SE | JAE | GSA GTCS | 13 | | GSWD | 3 | 3HZ bilateral generalised spike waves slight L dominance associated with clinical absence |
| 128 | 128203 | SE | JAE | GSA GTCS | 10 | | GSWD | 3 | one episode when fell repeatedly without loss of consciousness when playing over period of 15 mins single day |
| 132 | 132101 | SE | JAE | GSAmy | | | GSWD | | Generalised paroxysmal theta & delta activity with sharp waves |
| 132 | 132203 | SE | JAE | GSA GTCS | 11 | | GSWD | 3-4 | Episodes of irregular 3-4Hz GSW enhanced by hyperventilation |
| 133 | 133202 | SE | CAE JME | AS | 5 | | GSWD | 3 | Generalised spike waves. Slow waves posterior |
| 234 | 234208 | UK | JME | GSA MS GTCS | 12 | | Excess of B activity | | Excess of generalised B activity, otherwise NAD. AS always follows MS . GTCS no warnings, often bites tongue |
| 234 | 234313 | UK | JAE | GSA GTCS | 18 | | | | Asthma from aged 6-7 On no RX. Feels nauseated, eyes roll upwards following GTCS, incontinent of urine |
| 339 | 339201 | UK | CAE | GSA | 8 | | GSWD | 3 | Supports diagnosis of CAE |
| 340 | 340202 | UK | CAE | GSA | 6 | | GSWD | 0.5 | Several bursts of 3 second spike & wave |
| 343 | 343202 | UK | CAE | GSA | 7 | Y | GSWD | | |
| 345 | 345201 | UK | CAE | GSA | 9 | | GSWD | 3 | Compatible with primary generalised epilepsy |
| 349 | 349201 | UK | CAE | GSA | 6 | | GSWD | 2.5 | Prolonged runs of 2.5 spike& wave |
| 350 | 350202 | UK | CAE | GSA | 6 | Y | GSWD | 3 | Spike and slow wave activity confirm CAE |
| 352 | 352202 | UK | CAE | GSA | 12 | Y | GSWD | 3 | Number of typical absences associated with 3Hz generalised spike&wave |
| 354 | 354201 | UK | CAE | GSA | 3 | Y | GSWD | 2.5-3 | Definite abnormality /generalised spike&wave complexes at 3/s |
| 355 | 355201 | UK | CAE | GSA | 8 | | GSWD | 3 | Spike & wave complexes at 3/s accompanied by clinical absence |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | attack |
| 356 | 356201 | UK | CAE | GSA | 5 | Y | GSWD | 3 | Frontally dominated spike&wave activity associated with clinical signs.Suggest CAE |
| 357 | 357201 | UK | CAE | GSA | 1 | Y | GSWD | 3 | Compatible with petit mal 3/sec spike wave. Measles at 10/12 |
| 360 | 360202 | UK | CAE | GSA | 10 | | GSWD | 3-3.5 | 3-3.5 second spike and wave complexes generalised over the 2 hemispheres.some lateralised abnormalities. Typical petit mal absence |
| 361 | 361202 | UK | CAE | GSA | 8 | | GSWD | 3 | Definite abnormality with runs of 3 second spike wave |
| 362 | 362201 | GR | CAE | GSA | 5 | | GSWD | | |
| 364 | 364202 | UK | CAE | GSA | 6 | | GSWD | 3 | Confirms the clinical suspicion of childhood absence epilepsy. Language delay |
| 365 | 365202 | UK | CAE | GSA | 3 | Y | GSWD | 3 | Clinical attacks associated with approx 3/s spike & waves |
| 366 | 366202 | UK | CAE | GSA | 4 | | GSWD | | Support clinical diagnosis of primary generalised epilepsy in the form of absence attacks |
| 367 | 367202 | UK | CAE | GSA | 7 | | GSWD | | |
| 368 | 368201 | UK | CAE | GSA | 5 | Y | GSWD | 3 | Generalised epileptic discharges supporting absence type primary generalised epilepsy. Forceps delivery No SCBU |
| 369 | 369201 | UK | CAE | GSA | 6 | | GSWD | 3-3.5 | Vaginal delivery 36/40 mild jaundice |
| 371 | 371201 | UK | CAE | GSA | 7 | | GSWD | | Brief generalised discharges up to 0.5 second in duration of spike and slow waves.-typical absences |
| 380 | 380201 | DE | CAE | GSA | 8 | | GSWD | | |
| 385 | 385201 | UK | CAE | GSA | 2 | Y | GSWD | 3 | Emergency LSCS-delay in labour.Resus+Low apgar score at1 minute, @normal at 10minutes. Asbergers syndrome & fine motor dyspraxia, occasionally requires PR diazepam to stop fitting |
| 387 | 387202 | UK | CAE | GSA | 8 | | GSWD | | sharpened slow wave posteriorly |
| 388 | 388203 | UK | CAE | GSA | 10 | | GSWD | 3 | |
| 389 | 389201 | UK | CAE | GSA | 4 | | GSWD | 2.5-3 | EEG findings pathognomonic of childhood absence epilepsy |
| 390 | 390202 | UK | CAE | GSA | 11 | | GSWD | 3 | Petit mal epilepsy |
| 391 | 391202 | UK | CAE | GSA | 3 | | GSWD | | |
| 392 | 392202 | UK | CAE | GSA | 5 | | GSWD | 3 | Supports the diagnosis of CAE |
| 406 | 406202 | UK | CAE | GSA | 5 | Y | GSWD | 3.5 | Definite abnormality associated with overbreathing |
| 408 | 408201 | UK | CAE | GSA | 9 | | GSWD | 2-3 | generalised paroxysms from the occipital focus |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| 409 | 409201 | UK | CAE | GSA | 5 | | GSWD | 3 | Typical absence epilepsy |
| 410 | 410202 | UK | CAE | GSA | 5 | | GSWD | 3 | EEG changes confirm diagnosis of CAE |
| 412 | 412202 | UK | CAE | GSA | 11 | | GSWD | 2.5-3 | 2.5-3Hz spike & wave activity with clinical absences |
| 414 | 414201 | UK | CAE | GSA | 7 | | GSWD | | |
| 415 | 415201 | UK | CAE | GSA GTCS | 12 | | GSWD | 3 | Prolonged spike & wave discharges occur in response to hyperventilation. Support CAE |
| 417 | 417201 | UK | CAE | GSA | 6 | | GSWD | 3 | Suggest a potential for an idiopathic epilepsy syndrome with absence epilepsy |
| 418 | 418201 | UK | CAE | GSA | 4 | | GSWD | 3-4 | Diagnostic of childhood absence epilepsy, mild learning difficulties |
| 453 | 453201 | UK | CAE | GSA | 9 | | GSWD | 3 | Abnormal EEG with 2 attacks of petit mal absence consistent with primary generalised epileptic discharges |
| 456 | 456202 | UK | CAE | GSA | 4 | Y | GSWD | 3 | Record show clear changes of 3/second spike & wave with clinical changes. Other Provoking Factors: Going into bright lights and very cold things |
| 457 | 457201 | UK | CAE | GSA | 4 | Y | GSWD | | Hypoxic at birth no resus needed no SCBU Neurodevelopment: Requires Speech therapy and poor short term memory. Medication Regime: Started on Sodium Valproate with no response so was changed to Epilim Chrono |
| 459 | 459201 | SE | CAE | GSA | 2-12 | | GSWD | | |
| 462 | 462201 | SE | CAE | GSA | 2-12 | | GSWD | | |
| 463 | 463202 | SE | CAE | GSA | 2-12 | | GSWD | | |
| 464 | 464201 | SE | CAE | GSA | 2-12 | | GSWD | | |
| 465 | 465201 | SE | CAE | GSA | 2-12 | | GSWD | | |
| 468 | 468201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 469 | 469201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 470 | 470201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 471 | 471201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 472 | 472201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 473 | 473201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 474 | 474201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 475 | 475201 | AT | CAE | GSA | 2-12 | | GSWD | | |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| 476 | 476201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 477 | 477201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 478 | 478201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 479 | 479201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 480 | 480201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 481 | 481201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 482 | 482201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 483 | 483201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 484 | 484201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 485 | 485201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 486 | 486201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 487 | 487201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 488 | 488201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 489 | 489201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 490 | 490201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 491 | 491201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 492 | 492201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 493 | 493201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 494 | 494201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 495 | 495201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 496 | 496201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 497 | 497201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 498 | 498201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 499 | 499201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 500 | 500201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 501 | 501201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 502 | 502201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 503 | 503201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 504 | 504201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 505 | 505201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 506 | 506201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 507 | 507201 | AT | CAE | GSA | 2-12 | | GSWD | | |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| 508 | 508201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 509 | 509201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 510 | 510201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 511 | 511201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 512 | 512201 | AT | CAE | GSA | 2-12 | | GSWD | | |
| 519 | 519201 | SE | CAE | GSA | 2-12 | | GSWD | | |
| 520 | 520201 | SE | CAE | GSA | 2-12 | | GSWD | | |
| 522 | 522201 | SE | CAE | GSA | 2-12 | | GSWD | | |
| 533 | 533201 | UK | CAE | GSA | 2-12 | | GSWD | | |
| 573 | 573201 | GR | CAE | GSA | 2-12 | | GSWD | | |
| 574 | 574201 | GR | CAE | GSA | 2-12 | | GSWD | | |
| 576 | 576201 | GR | AE | GSA | 2-12 | | GSWD | | |
| 587 | 587202 | UK | CAE | GSA | 2-12 | | GSWD | 2-3 | |
| 589 | 589202 | UK | CAE | GSA | 6 | Y | GSWD | | Generalised spike and wave activity most during hyperventilation. Single slight burst during photic stimulation but this was probably coincidental. Full term baby, "went blue"-aspirated. In scub x 1/52 otherwise well |
| 595 | 595201 | UK | CAE | GSA GTCS | 8 | | GSWD | 3 | Findings consistent with primary generalised epilepsy. Nephrotic Syndrome |
| 596 | 596201 | UK | CAE | GSA GTCS | 7 | Y | GSWD | 3 | Abnormal record with 3 bilateral bursts of high voltage rhythmical spike and slow wave activity. Forceps delivery. ?Provoked by onset of puberty |
| 597 | 597201 | UK | CAE | GSA | 3 | Y | GSWD | 2.5 -4 | Primary generalised CAE with mildly atypical features |
| 598 | 598202 | UK | CAE | GSA | 6 | | GSWD | 3-6 | Treat for CAE. 41 weeks gestation Normal development |
| 599 | 599201 | UK | CAE | GSA | 5 | Y | GSWD | 3 | Abnormality with a run of 3 second spike wave complexes associated with clinical absences. Breech delivery |
| 600 | 600201 | UK | JAE | GSA GTCS | 12+ | | GPSWD | 3 | EEG confirms epileptic absences and now probably juvenile absence epilepsy spectrum. High Forceps Delivery for "long labour" Bruising to face and head. ? Provoked by watching movement-at race tract/watching TV |
| 601 | 601201 | UK | CAE | GSA | 7 | | GSWD | 3 | Definite abnormality with frequent generalised spike & slow wave discharges around 3 sec greatly facilitated by |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | hyperventilation. LSCS at 32 weeks gestation (4lbs at birth). SCBU for 1 week- no problems |
| 603 | 603202 | UK | CAE | GSA | 6 | | GSWD | 3 | Habitual absences during hyperventilation, associated with generalised regular3-4C/S spike & Wave activity |
| 604 | 604203 | UK | CAE | GSA | 8 | | GSWD | 3 | generalised regular 3-4c/ spike & wave activity, implies the potential for generalised seizures and is consistent with childhood absence epilepsy |
| 605 | 605201 | UK | CAE | GSA GTCS | 3 | Y | GSWD | 3 | Likely 'Petit mal' syndrome, several paroxysms of 3 sec discharges. Delayed Speech, Undergoing Speech Therapy. Mild global development delay. Fell down and unresponsive , PR Diazepam controlled fit. Total time 20 minutes |
| 606 | 606202 | UK | CAE | GSA | 7 | | GSWD | | Current EEG support the clinical picture of CAE responding well to sodium valproate |
| 608 | 608201 | UK | CAE | GSA | 6 | | GSWD | 3 | Generalised 3HZ Regular Spike & Wave. LSCS for Breech presentation |
| 609 | 609204 | UK | CAE | GSA FS | 6 | Y | GSWD | 2.5- 3 | Bursts of generalised 2.5-3Hz spike & Wave. Findings consistent with primary/secondary generalisation, former being most likely. Normal delivery. MRI/ arachnoid cyst. School work-slightly behind due to frequent absences. Febrile convulsion aged 9 months |
| 610 | 610201 | UK | CAE | GSA | 5 | Y | GSWD | 3 | EEG suggests a potential for a primary generalised epilepsy syndrome with epileptic absences and possible photosensitivity. 3 weeks premature, LBW-not in SCBU-normal development |
| 611 | 611203 | UK | JAE | GSA | 13 | | GSWD | | Generalised spike & Wave. Findings imply primary generalised epilepsy consistent with juvenile absence epilepsy |
| 612 | 612201 | UK | CAE | GSA | 10 | | GSWD | 4-5 | Pathological EEG. During drowsiness and several times during photic stimulation, short lasting irregular activity, mixed with epileptiform spike activity occurred lasting up to 2.5 seconds. |
| 614 | 614203 | UK | CAE | GSA | 6 | | GSWD | 2-3 | Bursts of spikes, one with intermittent slow waves. Generalised 2-3Hz spike/wave. Photic stimulation produced some following more left than right. Support a seizure disorder of primary generalised type. Delay in reaching milestones. Needing extra |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | help at school |
| 615 | 615202 | UK | CAE | GSA | 10 | | GSWD | 3 | Evidence for seizures of the primary generalised type, with ictal generalised 3 per second, at rest and during photic stimulation. 41 weeks gestation Normal development |
| 617 | 617203 | UK | CAE | GSA | 8 | | GSWD | ? | Generalised Frontally maximal spike and wave activity , consistent with Primary generalised type. Statemented due to epilepsy. Psychological report -15 extra hours help at school as behind peers/regular school |
| 618 | 618201 | UK | JAE | GSA | | | GSWD | 3 | EEG findings confirm epileptic absences consistent with absence epilepsy syndrome |
| 619 | 619202 | UK | CAE | GSA | 6 | | GSWD | 3 | EEG indicates primary generalised epilepsy with absence seizures and frequent subclinical discharges. 2 1/2, years behind peers, awaiting special help at school- regular school |
| 620 | 620201 | UK | CAE | GSA | 10 | | GSWD | 3 | An abnormal EEG diagnostic primary generalised epilepsy. |
| 621 | 621202 | UK | CAE | GSA | 4 | | GSWD | 2.5-4 | An abnormal EEG which is diagnostic of some form of absence epilepsy. 4 secs of generalised frontally predominant spike/polyspike and slow wave activity. In process of being statemented 2+ years behind peers |
| 622 | 622204 | UK | CAE | GSA | 6 | | GSWD | 3 | An EEG diagnostic of primary generalised epilepsy. No reproducible photosensitivity. . SCBU short term post delivery. Increased heart rate at birth. |
| 623 | 623201 | UK | CAE | GSA | 9 | | GSWD | 2-4 | Record continues to show frequent generalised runs of spikes waves complexes. LSCS/delay in labour, Dyslexia. |
| 624 | 624202 | UK | CAE | GSA | 3 | | GSWD | 3 | Three bursts of generalised 3Hz spike & wave activity. Overbreathing produced a brief episode of frontal polyspike and wave. Head injury aged 2 1/2 years. (banged into glass pane/unconscious briefly)Mum convinced this is a factor in development of CAE |
| 625 | 625202 | UK | CAE | GSA | 7 | | GSWD | 3 | Background show frequent generalised spikes. Bursts of regular generalised regular 3-4 C/S spike & wave occur throughout record, supports CAE. Scarlet fever at 36 months. Suffers from childhood migraines |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| 626 | 626201 | UK | CAE | GSA | 7 | | GSWD | 3 | EEG more unstable than usually seen in CAE |
| 632 | 632205 | DK | CAE | GSA GTCS | 10 | | GSWD | 3-5 | During hyperventilation and photic stimulation generalised paroxysms seen. |
| 635 | 635202 | UK | CAE | GSA | 5 | | GSWD | 3-5 | Features remain typical of childhood absence syndrome. Home birth NAD. Up with peers at school but poor to register information. onset 1999, poor response to ethosuximide, Sodium valproate initial good response, deteriorated-lamotrigine added-good response. Typical CAE. |
| 636 | 636201 | UK | CAE | GSA | 4 | | GSWD | 3-4 | Several brief bursts of regular 3-4 C/S spike & wave occur throughout the record. Imply the potential for generalised seizures and are consistent with CAE. No extra support at school but in last year mum notes school work deteriorating |
| 637 | 637202 | UK | CAE | GSA | 10 | | GSWD | 3 | Typical clinical events associated with 3Hz spike and waves, 2 spontaneously and 2 with overbreathing-consistent with CAE. Chicken- Pox quite severe but otherwise NAD |
| 638 | 638202 | UK | CAE | GSA GTCS | 5 | Y | GSWD | 3 | 2 bursts of 3/second spike & wave with a feature of there being 2 spikes for each wave. MRI -no structural abnormality in cerebral area- mild cerebellar ectopia noted. Speech therapy aged 3 years-couldn't pronounce "G's"- otherwise NAD. Lack of sleep- one GTCS, lasting 2 minutes approx, lips blue, LOC. |
| 639 | 639201 | UK | CAE | GSA | 5 | | GSWD | 2.5-3 | During hyperventilation typical 3 per second spike & wave discharges are seen as in petit mal syndrome. 42 weeks gestation-nil of note |
| 640 | 640202 | UK | CAE | GSA | 4 | | GSWD | 3-4 | Support clinical diagnosis of petit mal-"paroxysms are a little unusual" Induced 42 week-"Vernix ++", Aged 2:6 Language delay, Behind peers at school "dyscalculi"- needs extra help at school- Regular school. Ethosuximide controlled Absences x 2 years-seizures reoccurred-Ethosuximide restarted seizured now controlled once again on higher dose. |
| 642 | 642201 | UK | CAE | GSA | 4 | Y | SW | 3 | 3 second spike & wave seizure confirmed on EEG. MRI normal. 41weeks gestation. Well baby. |
| 644 | 644102 | UK | GTCS | GTCS | | | | | one GTCS aged 9 years. Never repeated. |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| 644 | 644201 | UK | CAE | GSA | 3 | | GSWD | 3 | Suggest primary generalised epilepsy. May have CAE but atypical features. Birth asphyxia-SCUB x 24 hours. Has extra help at school/ one-one sessions & speech therapy. vacant spells increased frequency. |
| 646 | 646202 | UK | CAE | GSA | 7 | | GSWD | 2.5-3 | 4 bursts of generalised spike & slow wave activity. 1 during overbreathing, 1 during photic stimulation- liability to primary generalised seizures fairly typical of CAE. 1 FS aged 18/12- associated with tonsillitis/ raised temp. |
| 647 | 647202 | UK | CAE | GSA FS | 3 | Y | GSWD | 3 | 3 typical 3Hz spike wave disturbance with a normal background-activity commensurate with the typical petit mal syndrome. Strep throat queried to have meningitis but transpired not to be so. Statemented - educationally has stayed static x 3 years. 1 FS aged 2 years associated with strep' throat infection. |
| 648 | 648201 | UK | CAE | GSA GTCS | | | GSWD | 2-3 | Features of primary generalised seizure present. Family history of petit mal on maternal side-great grandfather, great aunt and cousins. 4 GTCS ever. Provoked by being hot and sweaty. |
| 649 | 649202 | UK | CAE | GSA | 7 | | GSWD | 3-4 | Features typical of those seen in petit mal syndrome. "held breath at 3/12", Speech impediment corrected with speech therapy. |
| 650 | 650202 | UK | CAE | GSA | 10 | | GSWD | 2-4 | |
| 652 | 652201 | UK | CAE | GSA | 4 | | GSWD | 3 | Most of the features strongly suggest the clinical diagnosis of childhood absence epilepsy. Breech presentation and delivery/ Mother had LSCS ( Raised BP prior to delivery) |
| 653 | 653201 | UK | CAE | GSA | 4 | Y | GSWD | 3 | Abnormal Records. Multiple clinical attacks with 3Hz Spike and slow wave activity. One episode of polyspike activity. Consistent with clinical diagnosis of CAE. MRI normal. Slow at school. Was to be statemented but didn't happen. Behind peers |
| 657 | 657201 | UK | CAE | GSA | 12 | | GSWD | 2.5-4 | Record indicates an ongoing liability to primary seizure disorder and changes are consistent with absence seizures. Still having absences- nostrils flare/ appears absent for a short period. |
| 658 | 658101 | UK | CAE | GSA GTCS | 6 | | GSWD | | Diagnosed as having petit mal as a child. GTCS aged 14. 2-3 in total controlled with epilim and phenytoin. Presently being weaned off phenytoin. |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| 658 | 658202 | UK | CAE | GSA | 5 | | GSWD | 3-4 | Features suggest the presence of petit mal. SVD-Term baby. Nil of note |
| 659 | 659102 | UK | CAE | GSA | 8 | | GSWD | 3 | Diagnosed petit mal aged 8 years with 2-3 sec spike & wave. SVD. Well child |
| 659 | 659202 | UK | CAE | GSA | 3 | | GSWD | 3 | Multiple spike and slow waves at 3Hz. Support diagnosis of CAE. SVD .Term baby. Insignificant heart murmur diagnosed as neonate. |
| 660 | 660201 | UK | CAE | GSA GTCS | 8 | Y | GSWD | 3 | Episodes of generalised spike and slow waves in keeping with CAE. 42/40 gestation. Behind peers at school/ difficult to concentrate- extra help in classroom. GTCS , whole body would shake, mouth drop, put himself on the floor Some LOC |
| 661 | 661202 | UK | CAE | GSA | 7 | | GSWD | 3-4 | Few brief bursts of spike & slow wave discharges during hyperventilation accompanied with eyelid flicker. Compatible with CAE. SVD Term. Well in childhood. Absences- blank, continued in activity. Came around & was aware he had had an absence |
| 662 | 662201 | UK | CAE | GSA | 8 | | GSWD | 2.5-3 | Abnormal Record. Generalised predominant spike & wave activity typical of CAE. 41+3/40. Well child. School work slightly below average in past but now up with peers |
| 663 | 663201 | UK | CAE | GSA | 9 | | GSWD | 3 | 3 episodes of typical 3/s generalised spike & Wave activity consistent with primary generalised absence epilepsy. SVD at term. Well child. Separation anxiety aged 9/12 ( Mother spent time in Hosp) Very bright student. Absences stopped very soon having started medication. |
| 664 | 664203 | UK | CAE | GSA | 3 | | GSWD | 3 | Clinically and electrographically the EEG confirms typical childhood absence epilepsy. SVD at term. Nil of note |
| 666 | 666203 | UK | CAE | GSA | 6 | | GSWD | 3 | Features present are typical of those seen in CAE. Term Baby, SVD. Henoch-Schonlen Purpura aged 4 years-Resolved with no problems. Absences down from 40 a day to 10/15 since on epilim. |
| 667 | 667201 | UK | CAE | GSA | 9 | | GSWD | 3-4 | Typical of primary generalised epilepsy of the petit mal type. 42/40 gestation-induced. Forceps delivery. |
| 669 | 669201 | UK | CAE | GSA | 9 | | GSWD | 3 | Spike & wave 3c/s. Hyperventilation associated with clinical |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | absences. Consistent with CAE. SVD Term. School work slightly below average/ regular school. Absences controlled with epilim- Changing dose presently due to weight gain. |
| 670 | 670201 | UK | CAE | GSA GTCS | 2 | Y | GSWD | 3 | Whilst not absolutely typical of true CAE this should be viewed as very similar. SVD Term. Nil of note in early childhood. Cat Scan NAD. Absences- Mother says more like Day-Dreaming. GTCS x 5 from 11-15 years. LOC 2-3 MINUTES. Followed by headache. |
| 671 | 671202 | UK | CAE | GSA GTCS | 4 | Y | GSWD | 3 | The EEG and clinical changes indicate CAE. SVT-Term. Early childhood NAD. Normal for age. Special eds due to absences. LOC x 3. 2-3 minutes. Sept 2002 had an aura but no seizure. |
| 674 | 674203 | UK | CAE | GSA | 8 | Y | GSWD | 3-4 | Paroxysms exacerbated by hyperventilation and elicited by photic stimulation. Overall indicative of a primary seizure disorder but some unusual features for CAE. SVD Term. Asthmatic. Had extra help at school but now up with peers since D/C medication. Absences resolved with Rx. School work improved since Rx stopped. MRI NAD. Brother no Corpus Callosum on MRI. |
| 676 | 676202 | UK | CAE | GSA GTCS | 3 | Y | GSWD | 3-6 | Abnormal record.Changes are consistent with absence seizures, there are a few atypical features for CAE. SVD( Long labour) at term. Well child. LOC. While in bed sleeping, was shaking, foaming. Eyes opened staring. 1 GTCS ever. |
| 677 | 677203 | UK | CAE | GSA | 3 | Y | GSWD | 3 | Spike in mid-frontal-temporal regions. Findings consistent with clinical diagnosis of absence seizures. SVD, Term. Nil of note in early childhood. Sodium Valproate & Ethosuximide together controlled absences. |
| 678 | 678203 | UK | CAE | GSA | 11 | | GSWD | 3 | Hyperventilation evokes bursts of multispike & slow waves at 3 - 4 Hz. Viral meningitis as a baby. Well following same. Slightly below peers at school. Absences controlled well with epilim. |
| 679 | 679202 | UK | CAE | GSA | 7 | | GSWD | 3 | Several bursts of 3 per second spike & wave activity with an anterior predominance. In addition several bursts spike/waves especially during photostimulation. FS from aged 6/12- 5 years of age. ? May be statemented due to poor progress at school. FS |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | associated with acute illnesses. LOC with cynsis. Tired after same++. |
| 680 | 680203 | UK | CAE | GSA | 8 | | GSWD | 3 | Clinically & electronically typical CAE. SVD Term + 10/7. Held back at school . Was behind his peers x 18/12. Now catching up |
| 681 | 681201 | UK | CAE | GSA | 5 | | GSWD | 3-4 | Changes are compatible with typical CAE. SVD at 37/40 . Neonatal jaundice under UV lights.Left Dacryocystorhinosomy 12/96. Slightly below average at school due to inattentiveness. FH Maternal aunts GTCS X 3 maternal aunts. |
| 682 | 682101 | UK | CAE | GSA | | | GSWD | | Diagnosed as having petit mal in childhood |
| 682 | 682203 | UK | CAE | GSA | 7 | | GSWD | 3 | Clinical findings indicate CAE. Prolonged bursts of generalised spike & wave discharges. Isolated polyspike, spike & wave discharges over one or other hemisphere. SVD term. Generally lacks concentration. Slightly behind peers at school. |
| 683 | 683202 | UK | CAE | GSA | 8 | | GSWD | 3 | 3 per second spike & wave paroxysms typical of those seen in petit mal. SVD-Term. Nil of note in early development. Advanced -speaking at 8/12. Now up with peers with the exception of maths. Maternal Aunt had CAE in childhood. ? Maternal Grandmother CAE. |
| 686 | 686201 | UK | CAE | GSA FS | 5+ | | GSWD | 2-4 | Wakeful period synchronous spike & slow wave activity. During sleep frequent 1-3 second bursts of spike & slow wave. SVD term. Doing OK at school generally-did fall behind peers due to absences but now catching up. ? FS. With high temp appeared to be shaking++. ? |
| 687 | 687201 | UK | CAE | GSA | 9 | | GSWD | 3 | Abnormal record. Bursts of anterior predominant spike/wave activity at 3 c/s recorded. SVD Term. Above average at school, Got into grammar school |
| 688 | 688202 | UK | CAE | GSA FS | 6 | | GSWD | 3 | Abnormal record. Recurrent bursts of generalised spike & wave occur with HV. Compatible with CAE. SVD term. Well child. Below peers at school- reading particularly bad. With raised temp FC. Eyes roll " goes blue" |
| 689 | 689201 | UK | CAE | GSA | 6 | | GSWD | 3 | Absence seizures with 3Hz spike & wave were seen associated with lip smacking & blinking. This is primary generalised |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | epilepsy |
| 690 | 690202 | UK | CAE | GSA | 3 | | GSWD | 3-4 | Abnormal record, fairly brief generalised bursts of spike & wave discharges. In addition isolated polyspike/spike & wave appear over one or other hemispheres. Findings consistent with primary generalised epilepsy. SVD Term. Septic arthritis viral infection in knee. Special need at school ?ADHA Rx with dexamphetamine. |
| 691 | 691201 | UK | CAE | GSA | 3 | Y | GSWD | 3 | Frequent bursts of generalised 3C/S spike & slow wave discharges. Consistent with CAE. SVD-Term. Statemented Oro-motor dyspraxia. Occasional absences still when tired |
| 693 | 693201 | UK | CAE | GSA | 8 | | GSWD | 3-4 | Bursts of generalised spike/wave discharges at 3-4 secs. The changes indicate liability to generalised seizure disorder and is consistent with CAE. SVD 38/40. Intelligent child |
| 694 | 694204 | UK | CAE | GSA GTCS | 5 | Y | GSWD | 2-3 | Findings consistent with seizure disorder of primary generalised type with clinical diagnosis of CAE. C/S elective due to mums past history. Special Care x 3 days Oxygen x 3 days. Behind peers at school/Regular school. MRI -NAD. 2 episodes of GTCS aged 3 & 5 years.Became stiff and body shaking. No loss of bowel function. Post ictal tiredness++ |
| 695 | 695201 | UK | CAE | GSA | 4 | Y | GSWD | 3-4 | Abnormal Record. Frequent bursts of generalised spike & wave discharges at 3-4hZ. Some associated with clinical absences and jerks. These changes indicate CAE. Term +10/7. Fell aged 18/12 ?Convulsion. Aged 3 Years eyes twitching. Behind peers at school, extra help given-behavioural problems. Convulsion aged 18/12- never diagnosed but mother report fall for n apparent reason with shaking limbs. Eyes twitching aged 3 years |
| 696 | 696203 | UK | CAE | GSA GTCS | 6 | Y | GSWD | 3 | Prolonged 3 /sec spike /wave seen, accompanied by staring & chewing.EEG consistent with primary absence epilepsy. Twin delivery. SVD- Asthmatic. Doing well at school, top 20%. GTCS 1 every 2-4 weeks/ LOC , pale sweaty. Also has automatisms at other times.Hands in the air, mimics -unaware he is doing same. |
| 697 | 697201 | UK | CAE | GSA | 7 | | GSWD | 3 | During HV a paroxysm of bilateral 3/sec spike & wave occurred, typical of CAE except duration prolonged. SVD , term + 13/7. |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | Went into distress during labour. Suctioned, O2 at birth-No further Rx required. Above average at school. |
| 698 | 698201 | UK | CAE | GSA | 7 | | GSWD | 2-3.5 | Clinical absence , generalised spike & wave activity recorded,Compatible with CAE. SVD Term. Hospital letter says "some special needs at school". Mum says "at regular school but slow" |
| 699 | 699201 | UK | CAE | GSA | 7 | | GSWD | 3-4 | Significantly abnormal EEG with brief 'paroxysmal' complex discharges seen spontaneously (once) and during HV (once). Suggests a liability to epilepsy with absence. Moderate photosensitivity. While some features support a diagnosis of typical absence. SVD 38/40- well child |
| 700 | 700202 | UK | CAE | GSA | 9 | | GSWD | 3 | During the recording had a brief episode of staring and blinking associated with 3Hz spike & wave . This is CAE. LSCS Difficult delivery, (mother has small pelvis) Well baby. Speech Therapy aged 5 years. High IQ, struggles with reading. FH Grandmother ? Petit mal ? Fainting attacks, Grand-uncle GTCS. |
| 701 | 701201 | UK | CAE | GSA | 5 | | GSWD | 3-4 | Four bursts of generalised anterior predominant 3-4 cps spike & slow wave- provoked by hyperventilation. Highly suggested of CAE. SVD term, 10 lbs, 8oz baby- Needed resus at birth.Spent 3-4 days in SCBU. Well child following same. Normal schooling, up with peers. No FH |
| 703 | 703202 | UK | CAE | GSA GTCS | 8 | | GSWD | 3-4 | Abnormal record. Single brief bursts of generalised spik and wave activity over one or other hemisphere. Changes indicate liability to generalised seizure disorder. Normal background favours primary generalised epilepsy. Migraine aged 9 LOC, sleeping post episode +++. |
| 704 | 704201 | UK | CAE | GSA | 9 | | GSWD | 3 | EEG confirms the clinical diagnosis of absence seizures showing photosensitivity. Paternal Grandmother developed epilepsy aged 55 years. |
| 706 | 706201 | UK | CAE | GSA | 4 | | GSWD | 3 | EEG shows several episodes of high voltage spike/waves at 3c/s. Compatible with absence, petit mal epilepsy. SVD 38/40. Well Child. Up with peers at school/doing fine. Paternal Uncle has |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | GTCS. |
| 707 | 707202 | UK | CAE | GSA | 5 | Y | GSWD | - | On normal background generalised paroxysms of high voltage spike & wave on overbreathing. Compatible with petit mal. SVD 37+/40. Well baby. Struggles at school-Statemented. Due to absences. Recently added sodium valproate and no seizure in 3/52. Great Paternal Aunt ? CAE. |
| 708 | 708201 | UK | CAE | GSA GTCS | 11 | | GSWD | 3 | Abnormality compatible with primary generalised epilepsy with absences as well as GTCS. SVD term. Heart murmur-nil else. Up with peers at school. 1 GTCS ?? Associated with dehydration. Fell to floor, eyes rolling, shaking, unconscious x 1 hr |
| 709 | 709201 | UK | CAE | GSA GTCS | 12 | Y | GSWD | 3 | The abnormality in the EEG is that of primary epilepsy with absences, GTCS as well as averse effect on cognition and memory aside from overt absence seizures. Emergency section at 34/40.Well child despite premature delivery. Doing very well at school. Father says Loss of awareness and drops to floor x 1 minute - 6 episodes in total ever. |
| 710 | 710201 | UK | CAE | GSA | 8+ | | GSWD | 3 | Abnormal EEG, 4 brief typical absences evoked by hyperventilation consistent with primary generalised seizure disorder. SVD 41/40. Well Child. Doing well at school. No FH. Started Sodium Valproate 3/52 ago. No absences since beginning Rx. |
| 711 | 711201 | UK | CAE | GSA FS | 3 | | GSWD | | Previous abnormal with spike & wave. SVD Term. NAD. Up with peers at school. FS aged 18/12. 1 ever. Shaking/ sweating. Hospitalised x 24 hours . Nil of note. |
| 712 | 712202 | UK | CAE | GSA | 3 | | GSWD | 3 | EEG impressively abnormal with electro-clinical features of CAE. SVD at term. Well baby. Doing very well at school. ECG revealed monophasic ventricular ectopic beat every few seconds throughout recording. Absences more difficult to detect since beginning Rx. Still having as many per day but trying out different doses of medication. Maternal Uncle had GTCS in childhood. Now nil. |
| 713 | 713201 | UK | CAE | GSA | 7 | | GSWD | 3-3.3 | Generalised seizures associated with clinical changes, lasting 10- |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | 20 seconds. Findings suggest primary generalised Epilepsy. Emergency Caesarean Section assume had pre-eclampsia. Gestation 32/40. 5+ weeks in SCBU, ventilated x 3/7. Nil else- Doing OK at school with specific problem with poor spelling. Nil else. Absences with fixed stare, trembling. School noticed quite a few but mum says 1 or less per month, On 3 occasions "big ones" whereby there was episodes of urine incontinence . Never loss consciousness but was drowsy following episodes. |
| 714 | 714202 | UK | CAE | GSA | 9 | | GSWD | 3 | Background EEG normal. 3 episodes of 3 per second spike & wave consistent with primary generalised epilepsy. SVD Term. Jaundiced post delivery/ Rx under lights. Otherwise NAD. Absences with arrested movement/conversation. Flutter of eyelids. Lasting seconds- unaware. Paternal Grandfather ?? JAE in teenage years."Taboo subject" therefore not discussed in detail. |
| 838 | 838102 | UK | IGE | GTCS | 18 | | | | GTCS from 18 years. Still medicated with tegretol. Seizures resolved but RX continues. GTCS from 18 years. LOC Rx with tegretol some years now since last seizure. Mother also ? GTCS. |
| 838 | 838202 | UK | CAE | GSA | 6 | | GSWD | | Normal Background. Numerous short episodes of high voltage generalised spike, wave associated with eyelid flicker. Compatible with epilepsy with absences. SVD term. Cord around neck but no consequence. Doing OK at School. Absences resolved with epilim.FH GTCS Mother and maternal Grandmother. |
| 839 | 839102 | UK | CAE | GSA | 7 | | GSWD | | IGE. SVD term. NAD in early childhood. Idiopathic generalised epilepsy diagnosed aged 7 years. Absences since aged 7 years. Also Currently has occasional fit like seizures while sleeping particularly when unwell. Aware same is happening but unable to communicate. |
| 839 | 839202 | UK | CAE | GSA | 6 | Y | GSWD | 3-3.5 | Brief absence attacks with normal background, consistent with Primary generalised seizure. n evidence of abnormal photosensitivity. LSCS for placenta Previa. Mild developmental delay due to intrauterine exposure to valproate.( Fine at school). |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | Dairy intolerance . Congenital Dislocation hip. MRI normal. Absences slightly improved since starting Rx 6/52 ago. Blinks/ arrested movement and stops talking. Unaware |
| 840 | 840203 | UK | CAE | GSA | 5 | Y | GSWD | 2.5-3 | The record shows changes fully compatible with childhood absence epilepsy. Forceps Delivery at term. Cord around neck. Cystic Fibrosis diagnosed at 6/52. Absences worsen when problematic breathing. Does well at school generally. Absences, eyes roll, Stops talking, continues to walk. Has cystic fibrosis and when breathing deteriorated absences were worse. |
| 841 | 841202 | UK | CAE | GSA | 2 | Y | GSWD | 2.5-3.5 | Previous EEG shows typical discharges compatible with absence epilepsy. Now on Rx ? Change in EEG. SVD Term. Nil of note medically. MRI-Normal. Behind at school ?due to absences/awaiting statementing. Tiredness provokes absences. Eyes flutter, raised eyebrows slowly, conversation cessation. Currently titrating RX and some improvement obvious. |
| 842 | 842102 | UK | | FS | | Y | | | FS in childhood. |
| 842 | 842201 | UK | CAE | GSA FS | 5 | | GSWD | 3 | Irregular spike wave complexes around 3c/s provoked by overbreathing. From present record it would appear that this child is having more atypical absence attacks than are being recorded. SVD 41/40. Aged 14/12 had FC x 1 with septaemia. Admitted to GOS x 3/7.right arm jerks. LOC 1-2 minutes. X 2 episodes. More recently parents noticed brief episodes of unresponsiveness, staring and occasional eye blinking. EEG report purports seizures are more frequent than are being reported and are atypical in nature. Irregular spike-wave complexes around 3c/s provoked by overbreathing |
| 843 | 843201 | UK | CAE | GSA FS | 2 | | GSWD | 2.5-3 | normal Background with changes of mild drowsiness. £ episodes of generalised spike/wave discharges ( 2 provoked by hyperventilation) associated with vacant facial expression . One with additional body jerk. Consistent with Primary generalised epilepsy with abs. LSCS 33/40 for severe pre-eclampsia. SCBU needing oxygen for 5/7. In Scub x 2/52 only. Generally well baby |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | with Hernia ops at 6/12 & 9/12 ( inguanil and umbilical). Three episodes generalised spike/wave ( 2 provoked by Hyper-ventilation) associated with vacant facial expression. One with upper body jerk. Consistent with primary generalised epilepsy with absence attacks, |
| 844 | 844201 | UK | CAE | GSA | 9 | Y | GSWD | 3 | EEG demonstrates 3c/s spike /wave discharge during a clinical event. Normal background supporting primary generalised ( absence ) epilepsy. SVD term. U/S brain NAD aged 6 years. Above average at school. Paternal aunt 30+ years developed epilepsy following fall. |
| 845 | 845201 | UK | CAE | GSA | 5+ | | GSWD | 3 | Record shows abnormality with runs of irregular spike and slower components at 3c's associated with typical absence seizures induced by overbreathing. Interictal record - frequent isolated spikes with variable laterisation but n constant focal signs. Mosaic marker ring chromosome. LSCS for poor progress in labour failed to dilate. Term approx. Well baby . Average at school. FH nil. Clinical absence seizures . Mosaic marker ring chromosome. |
| 846 | 846201 | UK | CAE | GSA GTCS | 13 | Y | GSWD | 2-3 | During HV 2 episodes of generalised activity at 3c/s associated with clinical absences. In keeping with Primary generalised epilepsy. Forceps Delivery for foetal Distress. Term baby. Migraine aged 5 years. Bright child doing well at school but absences does disturb school achievements. Being referred to Queen square for investigations. Simple Partial and Complex partial since aged 14. LOC Violent seizures with biting of tongue and arm. In last year happens in sleep and awakening. |
| 847 | 847102 | UK | IGE | GTCS | 13 | | | | GTCS aged 13 years. Lasted x approx 1 year. Rx with Carbamazipine. Resolved aged 14 years. GTCS aged 13 years resolved 14 years. D/C medication. No further seizures. |
| 847 | 847201 | UK | CAE | GSA FS | 3 | Y | GSWD | 3 | Abnormality in the form of generalised approx 3/seconds. Occur spontaneously but most prominently with photic stimulation. Strongly supports an idiopathic generalised epilepsy such as typical Childhood absence. 39/40 gestation. Small for Dates |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | baby.. Global Delay-statemented. FC x 1 aged 5 years. LOC lasted seconds associated with pryrexia. |
| 848 | 848202 | UK | CAE | GSA | 9 | | GSWD | 3 | EEG consistent with CAE. 3 events 2 during HV.Staring, finger fiddling movements. SVD Term baby. No medical History. Doing well at school. Normal development with no regression of skills. FH nil. CAE from age 9-10 years. Sod Valproate titrated and Absences stopped. Arrest, unresponsive, non-purposeful hand movements and hand movements. |
| 849 | 849201 | UK | CAE | GSA | 10 | | GSWD | 3 | On Normal cerebral activity numerous paroxysms of generalised high -voltage spike & polyspike waves at about 3s/c compatible with epilepsy with absences. Vantouse Delivery for delay in 2nd stage. Well baby. Doing very well at school. Absences only. On Rx but no resolve. Goes blank, arrested conversation. Unaware come around and resumes activity |
| 850 | 850201 | UK | CAE | GSA | 2.6 | Y | GSWD | 3 | Record shows changes of childhood absence epilepsy. SVD term. Well young child with delayed milestones- speaking at 4 years.( ? Autistic features) Full statement. Assessed at Wolfden unit assessed to be 4.9 years ( actual age 7 years). EEG states CAE, Dr Letter july 03 ? Myoclonic jerks and complex epilepsy ( Mother denies jerks says absences only) |
| 851 | 851201 | UK | CAE | GSA | 1 | | GSWD | 3 | Abnormal with frequent unprovoked subclinical spike wave discharges. Largely generalised. Overall consistent with Primary generalised epilepsy. SVD 38/40. Mother says absences since aged 1 noticed. Statemented (3). Poor at school. FH of GTCS Maternal Uncle ( RIP aged 19 years with GTCS) and Paternal Aunt GTCS. |
| 854 | 854201 | UK | CAE | GSA | 8 | | GSWD | 3 | EEG accompaniment of clinical absences is typical of CAE. Photosensitivity is atypical as is rhythmic eyebrow movement. Vantouse Delivery delay in 2nd stage. Statemented at school, special needs at "normal School". FH Maternal Grandfather Seizures onset 35 years Rx with Phenytoin/ Phenbarbitoin. Paternal Great Grandmother GTCS. |

| *Fam ID* | *Full ID* | *Country of Origin* | *Epilepsy* | *Seizure* | *Age at onset* | *Imaging* | *Generalised EEG Abnormality.* | *Frequemcy (HZ)* | *Comments* |
|---|---|---|---|---|---|---|---|---|---|
| 857 | 857201 | UK | CAE | GSA | 7 | | GSWD | 3 | EEG typical of CAE except paroxysms relatively brief.1 burst coinciding with photic stimulation . Possibly may evolve into JAE. EEG would do for typical CAE. Vantouse Delivery 41/40.Average at school. FH 2nd cousin had absences. |
| 860 | 860201 | UK | CAE | GSA | 6 | | GSWD | 3 | Abnormal EEG. Confirms Typical CAE and is having frequent attacks. SVD -Term. Well child. Bright Child - doing very well at school. Absences now much lesser- occurring only when exceedingly tired! |
| 861 | 861201 | UK | CAE | GSA | 5 | Y | GSWD | 3 | Background normal. Generalised spike & wave both clinical and sub clinical consistent with primary generalised seizure disorder with absence attacks. No focal abnormalities. LSCS for Breech Presentation 37/40 gestation. Asthma- otherwise NAD. Top end of average at school. Absences worse if taking diet coke. ( Aspartame). Brother allergic to Aspartame |
| 862 | 862201 | UK | CAE | GSA FS | 4 | | GSWD | 3 | Frequent brief episodes in resting record with discharges sometimes limited to frontal cortex- Generalised during HV/regular 3HZ. EEG suggest a potential for JAE. Vantouse Delivery. Term + 10/7. Glandular fever aged 3 years. Top end of class at school. Absences occasional |
| 863 | 863201 | UK | CAE | GSA GTCS | 3 | | GSWD | 3 | Recurrence of Absences?-EEG show definite epileptiform abnormality. Discharges L& R also several bursts of 3 second spike & wave. There is no clear photosensitivity. Preterm 32/40 gestation, SCBU weight 4 1/2lbs. Respiratory problems. Well otherwise in early childhood. Normal development but school work affected by absences. GTCS x 4 in total Last episode 1996. LOC provokes by flashing lights.-Known diagnosis of absence epilepsy. Sodium Valproate stopped after 2 years fit free in 1999. |

## Appendix 14 Blood samples sent for EBV transformed lymphocyte cell lines

| Individual ID | Epilepsy | Country |
|---|---|---|
| 009201 | CAEatyp | UK |
| 027302 | CAE | UK |
| 357201 | CAE | UK |
| 363201 | CAE | UK |
| 364202 | CAE | UK |
| 366202 | CAE | UK |
| 371201 | CAE | UK |
| 388203 | CAE | UK |
| 390202 | CAE | UK |
| 392202 | CAE | UK |
| 572201 | CAE | Greece |
| 575102 | CAE | Greece |

## Appendix 15 Novel SSLP markers designed using Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)

| Novel SSLP | Gene | Repeat | Forward primer | Reverse primer | Size (bp) |
|---|---|---|---|---|---|
| *CACNG3EX1CA* | *CACNG3* | *CA* | aagcctaggcgttaagacga | ctctgggtttctttgccttg | 333 |
| *UCL10321* | *CACNG3* | *CA* | tgcaaacagactaggaccac | gagctcaggagtttgagacc | 213 |
| *A1GUCL* | *CACNA1G* | *GT* | ggtgttagggcggggtcgg | ggtattaagattaagaggggccaa | 268 |

## Appendix 16 Examples of input files required for GAS software

Appendix 16a GAS locus data file (.gas)

```
set logfile = chr16;
set locus cae affection
2     0.995       0.005
1
0.001 0.7   0.7
set locus D16S420 named
7     0.14 0.14 0.14 0.14 0.14 0.14 0.16
nodata;
set locus CACNG3EX1CA named
7     0.14 0.14 0.14 0.14 0.14 0.14 0.16
nodata;
set locus UCL10321 named
7     0.14 0.14 0.14 0.14 0.14 0.14 0.16
nodata;
read ( pedigree cae77.ped ) ;
read ( alsize chr16.siz locus D16S420 CACNG3EX1CA UCL10321 samesize
= 0.9 diffsize = 1.3 orderfirst global ) ;
write ( lpedigree chr16.ped ) ;
```

Appendix 16b GAS pedigree file (.ped). Data is shown for three families.

```
5     101   0     0     2     1
5     102   0     0     1     1
5     202   102   101   2     1
5     204   102   101   2     2
5     207   102   101   1     2
9     101   0     0     1     1
9     102   0     0     2     1
9     201   101   102   2     2
9     202   101   102   2     2
26    101   0     0     1     1
26    102   0     0     2     1
26    202   101   102   1     2
26    204   101   102   2     2
26    206   101   102   1     1
26    207   101   102   1     1
```

Appendix 16c GAS allele data file (.siz). The data in each row comprises the marker number in correct chromosomal order, the pedigree number and the allele sizes generated by the genescan and genotyper programmes.

```
D16S420     5.101 180   198
D16S420     5.102 184   192
D16S420     5.202 180   192
D16S420     5.204 180   184
D16S420     5.207 180   192
D16S420     9.101 190   192
D16S420     9.102 180   198
D16S420     9.201 192   198
D16S420     9.202 190   198
D16S420     26.101      188   198
D16S420     26.102      180   188
D16S420     26.202      180   198
D16S420     26.204      180   198
D16S420     26.206      180   188
D16S420     26.207      180   188
CACNG3EX1CA 5.101 x     x
CACNG3EX1CA 5.102 333   335
CACNG3EX1CA 5.202 333   335
CACNG3EX1CA 5.204 x     x
CACNG3EX1CA 5.207 333   335
CACNG3EX1CA 9.101 x     x
CACNG3EX1CA 9.102 x     x
CACNG3EX1CA 9.201 335   335
CACNG3EX1CA 9.202 x     x
CACNG3EX1CA 26.101      335   335
CACNG3EX1CA 26.102      333   335
CACNG3EX1CA 26.202      335   335
CACNG3EX1CA 26.204      335   335
CACNG3EX1CA 26.206      335   335
CACNG3EX1CA 26.207      335   335
UCL10321    5.101 210   214
UCL10321    5.102 212   214
UCL10321    5.202 212   214
UCL10321    5.204 x     x
UCL10321    5.207 214   214
UCL10321    9.101 208   214
UCL10321    9.102 208   214
UCL10321    9.201 208   214
UCL10321    9.202 208   214
UCL10321    26.101      208   214
UCL10321    26.102      208   214
UCL10321    26.202      208   214
UCL10321    26.204      208   214
UCL10321    26.206      208   208
UCL10321    26.207      208   208
```

**Appendix 17 Locus data file for GENEHUNTER. The penultimate line gives the recombination fractions between the four markers**

```
5 0 0 1
0 0.0 0.0 0
2 3 4 5
1 2
0.99 0.01
1
0.001 0.5 0.5
3 11  # D16S420
0.091 0.091 0.091 0.091 0.091 0.091 0.091 0.091 0.091 0.091 0.09
3 3  # CACNG3EX1CA
0.33 0.33 0.34
3 9  # UCL10321
0.11 0.11 0.11 0.11 0.11 0.11 0.11 0.11 0.12
3 8  # D16S401
0.125 0.125 0.125 0.125 0.125 0.125 0.125 0.125
0  0
0.0003 0.0025 0.0017
1  0.1  0.45
```

## Appendix 18 Additional data file required for GENEHUNTER-TWOLOCUS.

```
0.995 0.005 (w-t, mutant allele frequencies for disease locus 1)
0.995 0.005 (w-t, mutant allele frequencies for disease locus 2)
1           (liability classes)
0.00  0.00  0.00  0.00  (two-locus, parent-of-origin specific
0.00  0.90  0.90  0.90  penetrances matrix)
0.00  0.90  0.90  0.90
0.00  0.90  0.90  0.90
```

## Appendix 19 Individuals selected for *CACNG3* exonic sequencing

| Individual ID | Epilepsy | Family: CAE linkage | AE linkage | CAE Trio |
|---|---|---|---|---|
| 5204 | CAE | y | y | y |
| 5207 | CAE | y | y | y |
| 26202 | CAE | y | y | y |
| 26204 | CAE | y | y | y |
| 27301 | CAE | y | y | y |
| 92202 | CAE | y | y | y |
| 202201 | CAE | y | y | y |
| 317201 | CAE | y | y | y |
| 321201 | CAE | y | y | y |
| 346202 | CAE | y | y | y |
| 348206 | CAE | y | y | y |
| 353201 | CAE | y | y | y |
| 372201 | CAE | y | y | y |
| 381201 | CAE | y | y | y |
| 382201 | CAE | y | y | y |
| 394302 | CAE | y | y | y |
| 395201 | CAE | y | y | y |
| 397201 | CAE | y | y | y |
| 401201 | CAE | y | y | y |
| 402201 | CAE | y | y | y |
| 403201 | CAE | y | y | y |
| 404302 | CAE | y | y | y |
| 405201 | CAE | y | y | y |
| 407201 | CAE | y | y | y |
| 421201 | CAE | y | y | y |
| 422201 | CAE | y | y | y |
| 432201 | CAE | y | y | y |
| 433301 | CAE | y | y | y |
| 434201 | CAE | y | y | y |
| 435201 | CAE | y | y | y |
| 436201 | CAE | y | y | y |
| 436202 | CAE | y | y | y |
| 437201 | CAE | y | y | y |
| 442201 | CAE | y | y | y |
| 443202 | CAE | y | y | y |
| 445202 | CAE | y | y | y |
| 446201 | CAE | y | y | y |
| 447201 | CAE | y | y | y |
| 450201 | CAE | y | y | y |
| 461201 | CAE | y | y | y |
| 466201 | CAE | y | y | y |
| 527201 | CAE | y | y | y |
| 529301 | CAE | y | y | y |
| 530201 | CAE | y | y | y |
| 577201 | CAE | y | y | y |
| 586302 | CAE | y | y | y |

| | | Family: | | |
|---|---|---|---|---|
| Individual ID | Epilepsy | CAE linkage | AE linkage | CAE Trio |
| 173201 | CAE | | y | y |
| 358201 | CAE | | y | y |
| 393201 | CAE | | y | y |
| 426201 | CAE | | y | y |
| 440202 | CAE | | y | y |
| 441202 | CAE | | y | y |
| 525203 | CAE | | y | y |
| 431202 | JAE | | y | y |
| 438301 | JAE | | y | y |
| 439201 | JAE | | y | y |
| 449201 | JAE | | y | y |
| 451202 | JAE | | y | y |
| 9201 | AE | | y | y |
| 96301 | AE | | y | y |
| 137301 | AE | | y | y |
| 137302 | AE | | y | y |
| 342201 | AE | | y | y |
| 396201 | AE | | y | y |
| 400201 | AE | | y | y |
| 429203 | AE | | y | y |
| 430301 | AE | | y | y |
| 444201 | AE | | y | y |
| 448201 | AE | | y | y |
| 525202 | AE | | y | y |
| 528302 | AE | | y | y |
| 357201 | CAE | | | y |
| 364202 | CAE | | | y |

**Appendix 20 TDT results at *CACNG3* locus on chromosome 16p12-13.3 for 53 CAE pedigrees and 82 AE pedigrees, generated by GENEHUNTER 2.1.**

Appendix 20a 53 CAE pedigrees

```
*********************************************************************
****
*
*
*                          GENEHUNTER  -  Complete  Linkage  Analysis
*
*                                              (version  2.1_r5  beta)
*
*
*
*********************************************************************
*


npl:10> TDT Summary - (0 non-original affecteds):
Marker D16S420          trans untrans  Chi2  p-val
D16S420   - Allele 1       26     27     0.02  0.890746
D16S420   - Allele 2       30     42     2.00  0.157299
D16S420   - Allele 3       27     17     2.27  0.131668
D16S420   - Allele 4       14     21     1.40  0.236724
D16S420   - Allele 5       17     19     0.11  0.738883
D16S420   - Allele 6       16      7     3.52  0.060569
D16S420   - Allele 7        8      3     2.27  0.131668
D16S420   - Allele 8        2      5     1.29  0.256839
D16S420   - Allele 9        3      2     0.20  0.654721
D16S420   - Allele 10       1      1     0.00  1.000000
Marker CACNG3EX1CA      trans untrans  Chi2  p-val
CACNG3EX1CA- Allele 1      45     21     8.73  0.003135  ++
CACNG3EX1CA- Allele 2      22     44     7.33  0.006769  --
CACNG3EX1CA- Allele 3       8     10     0.22  0.637352
Marker UCL10321         trans untrans  Chi2  p-val
UCL10321  - Allele 1       44     39     0.30  0.583128
UCL10321  - Allele 2       37     27     1.56  0.211300
UCL10321  - Allele 3       27     49     6.37  0.011617  -
UCL10321  - Allele 4       10      8     0.22  0.637352
UCL10321  - Allele 5        9      2     4.45  0.034808  +
UCL10321  - Allele 6        5      6     0.09  0.763025
UCL10321  - Allele 7        0      1     1.00  0.317311

npl:11> 'photo' is off.

npl:12>
        ...goodbye...
```

```
**********************************************************************
****
*
*
*                            GENEHUNTER  -  Complete  Linkage  Analysis
*
*                                              (version  2.1_r5  beta)
*
*
*
**********************************************************************
*

npl:10> TDT Summary - (0 non-original affecteds):
Marker D16S420          trans untrans  Chi2  p-val
D16S420   - Allele 1       49     44     0.27  0.604126
D16S420   - Allele 2       37     54     3.18  0.074735
D16S420   - Allele 3       37     28     1.25  0.264288
D16S420   - Allele 4       26     32     0.62  0.430791
D16S420   - Allele 5       21     24     0.20  0.654721
D16S420   - Allele 6       24     13     3.27  0.070546
D16S420   - Allele 7       11      7     0.89  0.345779
D16S420   - Allele 8        3      6     1.00  0.317311
D16S420   - Allele 9        3      2     0.20  0.654721
D16S420   - Allele 10       1      1     0.00  1.000000
D16S420   - Allele 11       0      1     1.00  0.317311
Marker CACNG3EX1CA      trans untrans  Chi2  p-val
CACNG3EX1CA- Allele 1      63     36     7.36  0.006656  ++
CACNG3EX1CA- Allele 2      38     63     6.19  0.012861  -
CACNG3EX1CA- Allele 3      11     13     0.17  0.683091
Marker UCL10321         trans untrans  Chi2  p-val
UCL10321  - Allele 1       63     59     0.13  0.717245
UCL10321  - Allele 2       59     48     1.13  0.287596
UCL10321  - Allele 3       42     60     3.18  0.074706
UCL10321  - Allele 4       13     12     0.04  0.841481
UCL10321  - Allele 5       10      8     0.22  0.637352
UCL10321  - Allele 6        7      8     0.07  0.796253
UCL10321  - Allele 7        3      3     0.00  1.000000
UCL10321  - Allele 9        2      1     0.33  0.563703

npl:11> 'photo' is off.

npl:12>
        ...goodbye...
```

**Appendix 21 TDT results at $GABA_AR$ gene cluster, *GABRB3*, *GABRA5* and *GABRG3* on chromosome 15q11-13 for 53 CAE pedigrees and 82 AE pedigrees, generated by GENEHUNTER 2.1.**

Appendix 21a 53 CAE pedigrees

```
*************************************************************************
****
*
*
*                          GENEHUNTER  -  Complete  Linkage  Analysis
*
*                                              (version  2.1_r5  beta)
*
*
*
*************************************************************************
*

npl:10> TDT Summary - (0 non-original affecteds):
Marker GABRB3          trans untrans  Chi2  p-val
GABRB3    - Allele 1      46    43     0.10  0.750485
GABRB3    - Allele 2      23    27     0.32  0.571608
GABRB3    - Allele 3      19    14     0.76  0.384088
GABRB3    - Allele 4      16    15     0.03  0.857462
GABRB3    - Allele 5      14    18     0.50  0.479500
GABRB3    - Allele 6      11    12     0.04  0.834827
GABRB3    - Allele 7       2     6     2.00  0.157299
GABRB3    - Allele 8      12     5     2.88  0.089555
GABRB3    - Allele 9       4     6     0.40  0.527089
GABRB3    - Allele 10      4     6     0.40  0.527089
GABRB3    - Allele 11      2     1     0.33  0.563703
Marker D155            trans untrans  Chi2  p-val
D155      - Allele 1      48    41     0.55  0.458088
D155      - Allele 2      34    33     0.01  0.902765
D155      - Allele 3      27    16     2.81  0.093448
D155      - Allele 4       8    20     5.14  0.023342  -
D155      - Allele 5      13    10     0.39  0.531615
D155      - Allele 6       9     4     1.92  0.165518
D155      - Allele 7       1    10     7.36  0.006656  --
D155      - Allele 8       2     4     0.67  0.414216
D155      - Allele 9       2     5     1.29  0.256839
D155      - Allele 10      0     2     2.00  0.157299
D155      - Allele 11      2     3     0.20  0.654721
D155      - Allele 15      2     0     2.00  0.157299
Marker DA55            trans untrans  Chi2  p-val
DA55      - Allele 1      46    34     1.80  0.179713
DA55      - Allele 2      38    31     0.71  0.399396
DA55      - Allele 3      19    37     5.79  0.016157  -
DA55      - Allele 4       8    11     0.47  0.491297
DA55      - Allele 5       6     8     0.29  0.592980
DA55      - Allele 6       6    11     1.47  0.225253
DA55      - Allele 7      14     5     4.26  0.038947  +
DA55      - Allele 8       4     2     0.67  0.414216
```

```
DA55      - Allele 9       5     1     2.67  0.102470
DA55      - Allele 10      2     7     2.78  0.095581
DA55      - Allele 11      0     2     2.00  0.157299
DA55      - Allele 12      3     2     0.20  0.654721
DA55      - Allele 13      3     0     3.00  0.083264
DA55      - Allele 14      0     1     1.00  0.317311
DA55      - Allele 15      0     2     2.00  0.157299
DA55      - Allele 16      2     0     2.00  0.157299
DA55      - Allele 17      0     2     2.00  0.157299
Marker D15S156         trans untrans  Chi2  p-val
D15S156   - Allele 1      44    57     1.67  0.195821
D15S156   - Allele 2      50    41     0.89  0.345448
D15S156   - Allele 3       7     6     0.08  0.781511
D15S156   - Allele 4       2     2     0.00  1.000000
D15S156   - Allele 5       2     0     2.00  0.157299
D15S156   - Allele 7       1     1     0.00  1.000000
D15S156   - Allele 10      1     0     1.00  0.317311

npl:11> 'photo' is off.

npl:12>
        ...goodbye...
```

Appendix 21b 82 AE pedigrees

```
*************************************************************************
****
*
*
*                          GENEHUNTER  -  Complete  Linkage  Analysis
*
*                                              (version  2.1_r5  beta)
*
*
*
*************************************************************************
****

npl:10> TDT Summary - (0 non-original affecteds):
Marker GABRB3          trans untrans  Chi2  p-val
GABRB3    - Allele 1      62    66     0.12  0.723674
GABRB3    - Allele 2      34    35     0.01  0.904177
GABRB3    - Allele 3      30    30     0.00  1.000000
GABRB3    - Allele 4      23    18     0.61  0.434880
GABRB3    - Allele 5      23    23     0.00  1.000000
GABRB3    - Allele 6      17    17     0.00  1.000000
GABRB3    - Allele 7      11    11     0.00  1.000000
GABRB3    - Allele 8      14     5     4.26  0.038947  +
GABRB3    - Allele 9       6    10     1.00  0.317311
GABRB3    - Allele 10      5     9     1.14  0.285049
GABRB3    - Allele 11      2     1     0.33  0.563703
GABRB3    - Allele 12      0     2     2.00  0.157299
Marker D155            trans untrans  Chi2  p-val
D155      - Allele 1      67    67     0.00  1.000000
D155      - Allele 2      43    47     0.18  0.673290
D155      - Allele 3      41    28     2.45  0.117579
D155      - Allele 4      18    28     2.17  0.140369
D155      - Allele 5      24    23     0.02  0.884028
D155      - Allele 6      14     7     2.33  0.126630
D155      - Allele 7       6    10     1.00  0.317311
D155      - Allele 8       5     7     0.33  0.563703
D155      - Allele 9       4     5     0.11  0.738883
D155      - Allele 10      0     2     2.00  0.157299
D155      - Allele 11      2     3     0.20  0.654721
D155      - Allele 13      1     1     0.00  1.000000
D155      - Allele 15      3     0     3.00  0.083264
Marker DA55            trans untrans  Chi2  p-val
DA55      - Allele 1      68    48     3.45  0.063318
DA55      - Allele 2      53    49     0.16  0.692061
DA55      - Allele 3      26    46     5.56  0.018422  -
DA55      - Allele 4      15    13     0.14  0.705457
DA55      - Allele 5       9    15     1.50  0.220671
DA55      - Allele 6      10    17     1.81  0.177932
DA55      - Allele 7      18     9     3.00  0.083264
DA55      - Allele 8       9     4     1.92  0.165518
DA55      - Allele 9       5     3     0.50  0.479500
DA55      - Allele 10      5    10     1.67  0.196706
DA55      - Allele 11      2     4     0.67  0.414216
DA55      - Allele 12      3     2     0.20  0.654721
```

```
DA55      - Allele 13      3     2     0.20  0.654721
DA55      - Allele 14      0     2     2.00  0.157299
DA55      - Allele 15      0     2     2.00  0.157299
DA55      - Allele 16      2     0     2.00  0.157299
DA55      - Allele 17      0     2     2.00  0.157299
Marker D15S156         trans untrans  Chi2  p-val
D15S156   - Allele 1      67    68     0.01  0.931414
D15S156   - Allele 2      61    62     0.01  0.928155
D15S156   - Allele 3      11    14     0.36  0.548506
D15S156   - Allele 4       6     4     0.40  0.527089
D15S156   - Allele 5       2     0     2.00  0.157299
D15S156   - Allele 6       2     0     2.00  0.157299
D15S156   - Allele 7       2     2     0.00  1.000000
D15S156   - Allele 8       0     1     1.00  0.317311
D15S156   - Allele 9       0     2     2.00  0.157299
D15S156   - Allele 10      2     0     2.00  0.157299

npl:11> 'photo' is off.

npl:12>
        ...goodbye...
```

**Appendix 22 PDT results at *CACNG3* locus on chromosome 16p12-13.3 for 53 CAE pedigrees and 82 AE pedigrees, generated by PDT 5.1**

Appendix 22a 53 CAE pedigrees

**Marker *D16S420*:**

AVE PDT

| Marker | Allele/Band | Parental Contrib Trans | Parental Contrib Not trans | Discordant Sibs Aff | Discordant Sibs Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ 1 | 42 | 42 | 27 | 15 | 1 | 0.554 | 0.5797 | 06/08/06 | avePDT |
| | 2/ 2 | 31 | 43 | 18 | 20 | 1 | -1.746 | 0.0809 | 06/08/06 | avePDT |
| | 3/ 3 | 31 | 18 | 7 | 2 | 1 | 1.738 | 0.0823 | 06/08/06 | avePDT |
| | 4/ 4 | 10 | 22 | 6 | 10 | 1 | -1.919 | 0.0550 | 06/08/06 | avePDT |
| | 5/ 5 | 21 | 25 | 5 | 7 | 1 | -0.733 | 0.4638 | 06/08/06 | avePDT |
| | 6/ 6 | 19 | 10 | 9 | 5 | 1 | 1.750 | 0.0801 | 06/08/06 | avePDT |
| | 7/ 7 | 10 | 4 | 6 | 2 | 1 | 1.800 | 0.0719 | 06/08/06 | avePDT |
| | 8/ 8 | 2 | 3 | 1 | 1 | 1 | -1.236 | 0.2164 | 06/08/06 | avePDT |
| | 9/ 9 | 3 | 2 | 2 | 4 | 1 | 0.175 | 0.8608 | 06/08/06 | avePDT |
| | 10/ 10 | 1 | 1 | 1 | 0 | 1 | 1.000 | 0.3173 | 06/08/06 | avePDT |
| IND COUNTS: | | 170 | 170 | 41 | 33 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 9 | 17.507 | 0.0413 | 06/08/06 | GLOBAL_AVE_PDT |

**Marker *CACNG3EX1CA*:**

AVE PDT

| Marker | Allele/Band | Parental Contrib Trans | Parental Contrib Not trans | Discordant Sibs Aff | Discordant Sibs Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ 1 | 132 | 108 | 74 | 51 | 1 | 2.949 | 0.0032 | 06/08/06 | avePDT |
| | 2/ 2 | 28 | 50 | 11 | 17 | 1 | -2.496 | 0.0126 | 06/08/06 | avePDT |
| | 3/ 3 | 8 | 10 | 1 | 0 | 1 | -0.277 | 0.7815 | 06/08/06 | avePDT |
| IND COUNTS: | | 168 | 168 | 43 | 34 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 2 | 10.001 | 0.0067 | 06/08/06 | GLOBAL_AVE_PDT |

**Marker *UCL10321*:**

AVE PDT

| Marker | Allele/Band | | Parental Contrib Trans | Discordant Sibs Not trans | Aff | Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ | 1 | 70 | 63 | 40 | 23 | 1 | 0.745 | 0.4561 | 06/08/06 | avePDT |
| | 2/ | 2 | 35 | 29 | 18 | 10 | 1 | 1.296 | 0.1950 | 06/08/06 | avePDT |
| | 3/ | 3 | 33 | 54 | 11 | 23 | 1 | -2.807 | 0.0050 | 06/08/06 | avePDT |
| | 4/ | 4 | 10 | 8 | 2 | 2 | 1 | 0.473 | 0.6363 | 06/08/06 | avePDT |
| | 5/ | 5 | 9 | 2 | 8 | 2 | 1 | 1.504 | 0.1326 | 06/08/06 | avePDT |
| | 6/ | 6 | 5 | 5 | 3 | 6 | 1 | 0.493 | 0.6219 | 06/08/06 | avePDT |
| | 7/ | 7 | 0 | 1 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| IND COUNTS: | | | 162 | 162 | 41 | 33 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 6 | 11.866 | 0.0650 | 06/08/06 | GLOBAL_AVE_PDT |

Appendix 22b 82 AE pegrees

**Marker *D16S420*:**

AVE PDT

| Marker | Allele/Band | Parental Contrib Trans | Parental Contrib Not trans | Discordant Sibs Aff | Discordant Sibs Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ 1 | 77 | 71 | 36 | 20 | 1 | 1.115 | 0.2650 | 06/08/06 | avePDT |
| | 2/ 2 | 44 | 61 | 24 | 30 | 1 | -2.042 | 0.0411 | 06/08/06 | avePDT |
| | 3/ 3 | 44 | 32 | 13 | 5 | 1 | 1.544 | 0.1225 | 06/08/06 | avePDT |
| | 4/ 4 | 25 | 36 | 12 | 15 | 1 | -1.737 | 0.0823 | 06/08/06 | avePDT |
| | 5/ 5 | 27 | 32 | 7 | 10 | 1 | -0.868 | 0.3856 | 06/08/06 | avePDT |
| | 6/ 6 | 27 | 16 | 11 | 6 | 1 | 1.811 | 0.0701 | 06/08/06 | avePDT |
| | 7/ 7 | 13 | 8 | 9 | 5 | 1 | 1.050 | 0.2935 | 06/08/06 | avePDT |
| | 8/ 8 | 3 | 4 | 1 | 1 | 1 | -1.236 | 0.2164 | 06/08/06 | avePDT |
| | 9/ 9 | 3 | 2 | 2 | 4 | 1 | 0.175 | 0.8608 | 06/08/06 | avePDT |
| | 10/ 10 | 1 | 1 | 1 | 0 | 1 | 1.000 | 0.3173 | 06/08/06 | avePDT |
| | 11/ 11 | 0 | 1 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| IND COUNTS: | | 264 | 264 | 58 | 48 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 10 | 17.738 | 0.0596 | 06/08/06 | GLOBAL_AVE_PDT |

**Marker *CACNG3EX1CA***

AVE PDT

| Marker | Allele/Band | Parental Contrib Trans | Parental Contrib Not trans | Discordant Sibs Aff | Discordant Sibs Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ 1 | 195 | 168 | 99 | 76 | 1 | 2.598 | 0.0094 | 06/08/06 | avePDT |
| | 2/ 2 | 46 | 71 | 20 | 28 | 1 | -2.387 | 0.0170 | 06/08/06 | avePDT |
| | 3/ 3 | 11 | 13 | 3 | 0 | 1 | 0.034 | 0.9729 | 06/08/06 | avePDT |
| IND COUNTS: | | 252 | 252 | 61 | 52 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 2 | 8.300 | 0.0158 | 06/08/06 | GLOBAL_AVE_PDT |

**Marker *UCL10321***

AVE PDT

| Marker | Allele/Band | Parental Contrib Trans | Not trans | Discordant Sibs Aff | Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ 1 | 104 | 98 | 55 | 42 | 1 | 0.249 | 0.8036 | 06/08/06 | avePDT |
| | 2/ 2 | 67 | 60 | 26 | 18 | 1 | 0.867 | 0.3860 | 06/08/06 | avePDT |
| | 3/ 3 | 56 | 73 | 20 | 30 | 1 | -1.617 | 0.1059 | 06/08/06 | avePDT |
| | 4/ 4 | 13 | 12 | 2 | 2 | 1 | 0.212 | 0.8318 | 06/08/06 | avePDT |
| | 5/ 5 | 10 | 8 | 8 | 2 | 1 | 0.711 | 0.4768 | 06/08/06 | avePDT |
| | 6/ 6 | 7 | 7 | 5 | 6 | 1 | 0.361 | 0.7180 | 06/08/06 | avePDT |
| | 7/ 7 | 3 | 3 | 0 | 0 | 1 | 0.525 | 0.5994 | 06/08/06 | avePDT |
| | 9/ 9 | 2 | 1 | 0 | 0 | 1 | 1.000 | 0.3173 | 06/08/06 | avePDT |
| IND COUNTS: | | 262 | 262 | 58 | 50 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 7 | 4.712 | 0.6951 | 06/08/06 | GLOBAL_AVE_PDT |

**Appendix 23 PDT results at $GABA_AR$ gene cluster, *GABRB3*, *GABRA5* and *GABRG3* on chromosome 15q11-13 for 53 CAE pedigrees and 82 AE pedigrees, generated by PDT 5.1.**

Appendix 23a 53 CAE pedigrees

**Marker *GABRB3CA***

AVE PDT

| Marker | Allele/Band | Parental Contrib Trans | Parental Contrib Not trans | Discordant Sibs Aff | Discordant Sibs Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ 1 | 74 | 71 | 37 | 29 | 1 | 0.520 | 0.6031 | 06/08/06 | avePDT |
| | 2/ 2 | 25 | 29 | 15 | 10 | 1 | -0.688 | 0.4918 | 06/08/06 | avePDT |
| | 3/ 3 | 19 | 14 | 6 | 8 | 1 | 1.334 | 0.1821 | 06/08/06 | avePDT |
| | 4/ 4 | 16 | 16 | 7 | 3 | 1 | 0.506 | 0.6131 | 06/08/06 | avePDT |
| | 5/ 5 | 14 | 18 | 6 | 6 | 1 | -1.258 | 0.2083 | 06/08/06 | avePDT |
| | 6/ 6 | 11 | 12 | 7 | 3 | 1 | -0.385 | 0.7001 | 06/08/06 | avePDT |
| | 7/ 7 | 2 | 6 | 3 | 3 | 1 | -1.127 | 0.2596 | 06/08/06 | avePDT |
| | 8/ 8 | 11 | 5 | 4 | 4 | 1 | 1.271 | 0.2039 | 06/08/06 | avePDT |
| | 9/ 9 | 4 | 4 | 1 | 1 | 1 | -0.174 | 0.8618 | 06/08/06 | avePDT |
| | 10/ 10 | 4 | 6 | 2 | 1 | 1 | -0.756 | 0.4497 | 06/08/06 | avePDT |
| | 11/ 11 | 2 | 1 | 0 | 0 | 1 | 1.000 | 0.3173 | 06/08/06 | avePDT |
| IND COUNTS: | | 182 | 182 | 44 | 34 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 10 | 8.179 | 0.6114 | 06/08/06 | GLOBAL_AVE_PDT |

## Marker *155CA2*

AVE PDT

| Marker | Allele/Band | Parental Contrib Trans | Parental Contrib Not trans | Discordant Sibs Aff | Discordant Sibs Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ 1 | 76 | 71 | 37 | 31 | 1 | 0.474 | 0.6355 | 06/08/06 | avePDT |
| | 2/ 2 | 36 | 33 | 15 | 12 | 1 | -0.090 | 0.9283 | 06/08/06 | avePDT |
| | 3/ 3 | 30 | 18 | 15 | 11 | 1 | 1.896 | 0.0580 | 06/08/06 | avePDT |
| | 4/ 4 | 11 | 23 | 4 | 6 | 1 | -2.357 | 0.0184 | 06/08/06 | avePDT |
| | 5/ 5 | 15 | 12 | 9 | 2 | 1 | 0.636 | 0.5245 | 06/08/06 | avePDT |
| | 6/ 6 | 9 | 4 | 2 | 1 | 1 | 1.050 | 0.2935 | 06/08/06 | avePDT |
| | 7/ 7 | 3 | 12 | 0 | 1 | 1 | -2.137 | 0.0326 | 06/08/06 | avePDT |
| | 8/ 8 | 1 | 3 | 1 | 1 | 1 | -1.342 | 0.1797 | 06/08/06 | avePDT |
| | 9/ 9 | 4 | 7 | 1 | 1 | 1 | -1.633 | 0.1025 | 06/08/06 | avePDT |
| | 10/ 10 | 0 | 2 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 11/ 11 | 2 | 3 | 3 | 3 | 1 | -1.414 | 0.1573 | 06/08/06 | avePDT |
| | 13/ 13 | 0 | 0 | 0 | 1 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 15/ 15 | 1 | 0 | 1 | 0 | 1 | 1.000 | 0.3173 | 06/08/06 | avePDT |
| IND COUNTS: | | 188 | 188 | 44 | 35 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 12 | 23.009 | 0.0277 | 06/08/06 | GLOBAL_AVE_PDT |

## Marker *A55CA1*

AVE PDT

| Marker | Allele/Band | Parental Contrib Trans | Parental Contrib Not trans | Discordant Sibs Aff | Discordant Sibs Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ 1 | 72 | 58 | 39 | 26 | 1 | 0.983 | 0.3258 | 06/08/06 | avePDT |
| | 2/ 2 | 47 | 39 | 17 | 13 | 1 | 1.309 | 0.1905 | 06/08/06 | avePDT |
| | 3/ 3 | 20 | 37 | 7 | 10 | 1 | -2.429 | 0.0152 | 06/08/06 | avePDT |
| | 4/ 4 | 6 | 11 | 0 | 5 | 1 | -0.896 | 0.3701 | 06/08/06 | avePDT |
| | 5/ 5 | 6 | 10 | 5 | 3 | 1 | -0.588 | 0.5566 | 06/08/06 | avePDT |
| | 6/ 6 | 6 | 11 | 5 | 6 | 1 | -0.697 | 0.4858 | 06/08/06 | avePDT |
| | 7/ 7 | 14 | 5 | 9 | 5 | 1 | 1.541 | 0.1234 | 06/08/06 | avePDT |
| | 8/ 8 | 4 | 2 | 3 | 1 | 1 | 1.342 | 0.1797 | 06/08/06 | avePDT |
| | 9/ 9 | 5 | 1 | 2 | 1 | 1 | 1.342 | 0.1797 | 06/08/06 | avePDT |
| | 10/ 10 | 2 | 7 | 1 | 0 | 1 | -1.413 | 0.1576 | 06/08/06 | avePDT |
| | 11/ 11 | 0 | 2 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 12/ 12 | 3 | 2 | 0 | 0 | 1 | 0.632 | 0.5271 | 06/08/06 | avePDT |
| | 13/ 13 | 3 | 0 | 0 | 0 | 1 | 1.000 | 0.3173 | 06/08/06 | avePDT |
| | 14/ 14 | 0 | 1 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 15/ 15 | 0 | 2 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 16/ 16 | 2 | 0 | 0 | 0 | 1 | 1.000 | 0.3173 | 06/08/06 | avePDT |
| | 17/ 17 | 0 | 2 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| IND COUNTS: | | 190 | 190 | 44 | 35 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 16 | 23.136 | 0.1101 | 06/08/06 | GLOBAL_AVE_PDT |

**Marker *D15S156***

AVE PDT

| Marker | Allele/Band | Parental Contrib Trans | Parental Contrib Not trans | Discordant Sibs Aff | Discordant Sibs Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ 1 | 81 | 94 | 39 | 29 | 1 | -0.685 | 0.4930 | 06/08/06 | avePDT |
| | 2/ 2 | 74 | 65 | 40 | 36 | 1 | 0.172 | 0.8635 | 06/08/06 | avePDT |
| | 3/ 3 | 7 | 6 | 1 | 1 | 1 | 0.218 | 0.8273 | 06/08/06 | avePDT |
| | 4/ 4 | 2 | 2 | 0 | 0 | 1 | 0.000 | 1.0000 | 06/08/06 | avePDT |
| | 5/ 5 | 2 | 0 | 3 | 1 | 1 | 1.571 | 0.1161 | 06/08/06 | avePDT |
| | 7/ 7 | 1 | 1 | 1 | 1 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 10/ 10 | 1 | 0 | 0 | 0 | 1 | 1.000 | 0.3173 | 06/08/06 | avePDT |
| IND COUNTS: | | 168 | 168 | 42 | 34 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 6 | 4.300 | 0.6362 | 06/08/06 | GLOBAL_AVE_PDT |

## Appendix 23b 82 AE pedigrees

## Marker *GABRB3CA*

AVE PDT

| Marker | Allele/Band | Parental Contrib Trans | Parental Contrib Not trans | Discordant Sibs Aff | Discordant Sibs Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ 1 | 100 | 105 | 45 | 35 | 1 | -0.165 | 0.8690 | 06/08/06 | avePDT |
| | 2/ 2 | 36 | 38 | 21 | 21 | 1 | -0.498 | 0.6182 | 06/08/06 | avePDT |
| | 3/ 3 | 30 | 30 | 8 | 11 | 1 | 0.652 | 0.5142 | 06/08/06 | avePDT |
| | 4/ 4 | 23 | 18 | 10 | 10 | 1 | 0.747 | 0.4548 | 06/08/06 | avePDT |
| | 5/ 5 | 20 | 22 | 10 | 7 | 1 | -0.953 | 0.3407 | 06/08/06 | avePDT |
| | 6/ 6 | 17 | 16 | 11 | 4 | 1 | 0.292 | 0.7705 | 06/08/06 | avePDT |
| | 7/ 7 | 12 | 10 | 6 | 4 | 1 | 0.224 | 0.8224 | 06/08/06 | avePDT |
| | 8/ 8 | 13 | 5 | 6 | 4 | 1 | 1.590 | 0.1119 | 06/08/06 | avePDT |
| | 9/ 9 | 6 | 8 | 3 | 1 | 1 | -0.556 | 0.5785 | 06/08/06 | avePDT |
| | 10/ 10 | 5 | 9 | 2 | 1 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 11/ 11 | 2 | 1 | 0 | 0 | 1 | 1.000 | 0.3173 | 06/08/06 | avePDT |
| | 12/ 12 | 0 | 2 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| IND COUNTS: | | 264 | 264 | 61 | 49 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 11 | 7.461 | 0.7606 | 06/08/06 | GLOBAL_AVE_PDT |

**Marker *155CA2***

AVE PDT

| Marker | Allele/Band | | Parental Contrib Trans | Parental Contrib Not trans | Discordant Sibs Aff | Discordant Sibs Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ | 1 | 102 | 105 | 51 | 44 | 1 | -0.114 | 0.9089 | 06/08/06 | avePDT |
| | 2/ | 2 | 47 | 49 | 17 | 12 | 1 | -0.451 | 0.6523 | 06/08/06 | avePDT |
| | 3/ | 3 | 44 | 27 | 20 | 16 | 1 | 1.797 | 0.0723 | 06/08/06 | avePDT |
| | 4/ | 4 | 23 | 32 | 7 | 9 | 1 | -1.001 | 0.3166 | 06/08/06 | avePDT |
| | 5/ | 5 | 27 | 27 | 11 | 5 | 1 | -0.136 | 0.8917 | 06/08/06 | avePDT |
| | 6/ | 6 | 13 | 7 | 5 | 2 | 1 | 1.439 | 0.1500 | 06/08/06 | avePDT |
| | 7/ | 7 | 8 | 12 | 3 | 5 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 8/ | 8 | 4 | 6 | 3 | 2 | 1 | -1.265 | 0.2059 | 06/08/06 | avePDT |
| | 9/ | 9 | 6 | 7 | 1 | 1 | 1 | 0.000 | 1.0000 | 06/08/06 | avePDT |
| | 10/ | 10 | 0 | 2 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 11/ | 11 | 2 | 3 | 3 | 3 | 1 | -1.414 | 0.1573 | 06/08/06 | avePDT |
| | 13/ | 13 | 1 | 1 | 0 | 1 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 15/ | 15 | 1 | 0 | 1 | 0 | 1 | 1.000 | 0.3173 | 06/08/06 | avePDT |
| IND COUNTS: | | | 278 | 278 | 61 | 50 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 12 | 13.052 | 0.3652 | 06/08/06 | GLOBAL_AVE_PDT |

## Marker *A55CA1*

AVE PDT

| Marker | Allele/Band | Parental Contrib Trans | Parental Contrib Not trans | Discordant Sibs Aff | Discordant Sibs Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ 1 | 108 | 87 | 57 | 40 | 1 | 1.214 | 0.2248 | 06/08/06 | avePDT |
| | 2/ 2 | 68 | 62 | 26 | 23 | 1 | 0.945 | 0.3445 | 06/08/06 | avePDT |
| | 3/ 3 | 27 | 46 | 10 | 11 | 1 | -2.392 | 0.0167 | 06/08/06 | avePDT |
| | 4/ 4 | 13 | 13 | 1 | 8 | 1 | -0.028 | 0.9776 | 06/08/06 | avePDT |
| | 5/ 5 | 9 | 17 | 9 | 7 | 1 | -0.904 | 0.3661 | 06/08/06 | avePDT |
| | 6/ 6 | 10 | 17 | 5 | 6 | 1 | -0.985 | 0.3248 | 06/08/06 | avePDT |
| | 7/ 7 | 18 | 9 | 10 | 5 | 1 | 1.343 | 0.1792 | 06/08/06 | avePDT |
| | 8/ 8 | 9 | 4 | 3 | 2 | 1 | 1.406 | 0.1599 | 06/08/06 | avePDT |
| | 9/ 9 | 5 | 3 | 2 | 1 | 1 | 0.333 | 0.7389 | 06/08/06 | avePDT |
| | 10/ 10 | 5 | 10 | 1 | 0 | 1 | -1.020 | 0.3078 | 06/08/06 | avePDT |
| | 11/ 11 | 2 | 4 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 12/ 12 | 3 | 2 | 0 | 0 | 1 | 0.632 | 0.5271 | 06/08/06 | avePDT |
| | 13/ 13 | 3 | 2 | 0 | 1 | 1 | 0.068 | 0.9459 | 06/08/06 | avePDT |
| | 14/ 14 | 0 | 2 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 15/ 15 | 0 | 2 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 16/ 16 | 2 | 0 | 0 | 0 | 1 | 1.000 | 0.3173 | 06/08/06 | avePDT |
| | 17/ 17 | 0 | 2 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| IND COUNTS: | | 282 | 282 | 62 | 52 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 16 | 19.024 | 0.2674 | 06/08/06 | GLOBAL_AVE_PDT |

**Marker *D15S156***

AVE PDT

| Marker | Allele/Band | | Parental Contrib Trans | Parental Contrib Not trans | Discordant Sibs Aff | Discordant Sibs Unaff | DF | Z | Pvalue | Date | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1/ | 1 | 136 | 136 | 58 | 38 | 1 | 0.368 | 0.7130 | 06/08/06 | avePDT |
| | 2/ | 2 | 97 | 97 | 53 | 55 | 1 | -0.548 | 0.5834 | 06/08/06 | avePDT |
| | 3/ | 3 | 13 | 16 | 3 | 5 | 1 | -0.284 | 0.7761 | 06/08/06 | avePDT |
| | 4/ | 4 | 6 | 4 | 2 | 1 | 1 | 0.602 | 0.5472 | 06/08/06 | avePDT |
| | 5/ | 5 | 2 | 0 | 3 | 1 | 1 | 1.571 | 0.1161 | 06/08/06 | avePDT |
| | 6/ | 6 | 2 | 0 | 0 | 0 | 1 | 1.000 | 0.3173 | 06/08/06 | avePDT |
| | 7/ | 7 | 2 | 2 | 1 | 1 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 8/ | 8 | 0 | 1 | 0 | 1 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| | 9/ | 9 | 0 | 2 | 0 | 0 | 1 | -1.000 | 0.3173 | 06/08/06 | avePDT |
| IND COUNTS: | | | 258 | 258 | 60 | 51 | | | | | |

Global score:

| Marker | DF | ChiSq | Pvalue | Date | |
|---|---|---|---|---|---|
| | 8 | 6.532 | 0.5878 | 06/08/06 | GLOBAL_AVE_PDT |

## Appendix 24 Predicted tagging SNPs based on whole resource and subset of 31 trios, using two strategies incorporated in Haploview

### Appendix 24a Comparison of tag SNPs selected by two strategies

| Whole resource: | | | Subset of 31 trios: | | |
|---|---|---|---|---|---|
| **SNP number** | **Frequency threshold** | **Tagger** | **SNP number** | **Frequency threshold** | **Tagger** |
| 1 | ✓ | ✓ | 1 | ✓ | ✓ |
| 2 | ✓ | ✓ | 2 | ✓ | ✓ |
| 3 | ✓ | | 3 | ✓ | |
| 4 | | ✓ | 4 | | |
| 5 | ✓ | | 5 | ✓ | ✓ |
| 6 | ✓ | ✓ | 6 | ✓ | ✓ |
| 7 | | ✓ | 7 | ✓ | |
| 8 | | | 8 | | ✓ |
| 9 | ✓ | ✓ | 9 | ✓ | ✓ |
| 10 | ✓ | | 10 | ✓ | ✓ |
| 11 | ✓ | | 11 | ✓ | ✓ |
| 12 | | ✓ | 12 | | |
| 13 | ✓ | ✓ | 13 | ✓ | ✓ |
| 14 | ✓ | ✓ | 14 | ✓ | ✓ |
| 15 | ✓ | ✓ | 15 | ✓ | ✓ |
| 16 | ✓ | ✓ | 16 | ✓ | ✓ |
| 17 | ✓ | ✓ | 17 | ✓ | ✓ |
| 18 | ✓ | ✓ | 18 | ✓ | ✓ |
| 19 | ✓ | ✓ | 19 | ✓ | ✓ |
| 20 | ✓ | ✓ | 20 | ✓ | ✓ |
| 21 | ✓ | ✓ | 21 | ✓ | ✓ |
| 22 | ✓ | ✓ | 22 | ✓ | ✓ |
| 23 | ✓ | ✓ | 23 | ✓ | ✓ |
| Total tag SNPs | 19 | 18 | Total tag SNPs | 20 | 19 |

**Appendix 24b SNPs captured by tag SNPs based on whole resource**

| Frequency threshold: | |
|---|---|
| **Tag SNP** | **SNPs captured** |
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |
| 5 | 5 |
| 6 | 6 |
| 9 | 9 |
| 10 | 10 |
| 11 | 11 |
| 13 | 13 |
| 14 | 14 |
| 15 | 15 |
| 16 | 16 |
| 17 | 17 |
| 19 | 19 |
| 20 | 20 |
| 21 | 21 |
| 22 | 22 |
| 23 | 23 |
| 2,3,5,6 | 4,7,8 |
| 10,11,13 | 12 |

| Tagger: | |
|---|---|
| **Tag SNP** | **SNPs captured** |
| 1 | 1 |
| 2 | 2 |
| 4 | 4 |
| 6 | 6 |
| 7 | 3,8,7 |
| 9 | 9 |
| 12 | 11,12 |
| 13 | 13 |
| 14 | 14 |
| 15 | 15 |
| 16 | 16 |
| 17 | 17 |
| 18 | 18 |
| 19 | 19 |
| 20 | 20 |
| 21 | 21 |
| 22 | 22 |
| 23 | 23 |
| 7,4,9 | 5 |
| 13,7,12 | 10 |

**Appendix 25 Results from the PDT analysis of each *CACNG3* SNP independently.**

Appendix 25a Affectedness criteria CAE

| | | | | *SUM PDT* | | *AVE PDT* | |
|---|---|---|---|---|---|---|---|
| *SNP* | *Allele* | *Transmitted* | *Not transmitted* | *Z (1df)* | *p-value* | *Z (1df)* | *p-value* |
| 1 | 2(A) | 238 | 230 | 0.75 | 0.454 | 0.75 | 0.454 |
| 2 | 2(A) | 199 | 168 | 1.67 | 0.094 | 1.74 | 0.083 |
| 3 | 2(G) | 276 | 235 | 2.41 | 0.016 | 2.18 | 0.029 |
| 4 | 1(A) | 375 | 354 | 1.76 | 0.078 | 1.63 | 0.103 |
| 5 | 1(T) | 394 | 378 | 0.82 | 0.411 | 0.58 | 0.560 |
| 6 | 1(T) | 453 | 435 | 1.52 | 0.128 | 1.69 | 0.091 |
| 7 | 2(G) | 247 | 208 | 2.40 | 0.016 | 2.23 | 0.026 |
| 8 | 2(A) | 266 | 222 | 2.64 | 0.008 | 2.47 | 0.013 |
| 9 | 1(G) | 232 | 227 | 0.41 | 0.682 | 0.41 | 0.682 |
| 10 | 2(A) | 286 | 260 | 1.25 | 0.213 | 1.59 | 0.113 |
| 11 | 2(G) | 179 | 139 | 2.90 | 0.004 | 2.49 | 0.013 |
| 12 | 1(A) | 99 | 86 | 1.21 | 0.225 | 1.21 | 0.225 |
| 13 | 1(T) | 399 | 381 | 1.62 | 0.105 | 1.29 | 0.199 |
| 14 | 2(G) | 317 | 300 | 0.89 | 0.376 | 0.86 | 0.391 |
| 15 | 2(T) | 345 | 335 | 1.26 | 0.207 | 0.22 | 0.827 |
| 16 | 1(T) | 409 | 380 | 2.14 | 0.033 | 2.38 | 0.017 |
| 17 | 2(C) | 239 | 227 | 0.79 | 0.431 | 1.07 | 0.286 |
| 18 | 1(G) | 401 | 390 | 0.36 | 0.717 | 0.67 | 0.503 |
| 19 | 1(T) | 317 | 299 | 1.54 | 0.123 | 0.82 | 0.411 |
| 20 | 1(C) | 361 | 358 | 0.37 | 0.710 | 0.62 | 0.533 |
| 21 | 1(T) | 296 | 272 | 1.24 | 0.215 | 1.08 | 0.282 |
| 22 | 1(C) | 245 | 244 | 0.10 | 0.919 | 0.10 | 0.919 |
| 23 | 2(C) | 141 | 135 | 0.48 | 0.633 | 0.48 | 0.633 |

Appendix 25b Affectedness criteria CAE, JAE or AE

| | | | | *SUM PDT* | | *AVE PDT* | |
|---|---|---|---|---|---|---|---|
| *SNP* | *Allele* | *Transmitted* | *Not transmitted* | *Z (1df)* | *p-value* | *Z (1df)* | *p-value* |
| 1 | 2(A) | 252 | 242 | 0.92 | 0.357 | 0.92 | 0.357 |
| 2 | 2(A) | 240 | 200 | 2.17 | 0.030 | 1.86 | 0.063 |
| 3 | 2(G) | 331 | 279 | 2.68 | 0.007 | 2.46 | 0.014 |
| 4 | 1(A) | 454 | 428 | 2.05 | 0.040 | 1.80 | 0.072 |
| 5 | 1(T) | 465 | 448 | 0.63 | 0.530 | 0.36 | 0.720 |
| 6 | 1(T) | 535 | 520 | 1.50 | 0.134 | 1.74 | 0.081 |
| 7 | 2(G) | 298 | 248 | 2.67 | 0.008 | 2.52 | 0.012 |
| 8 | 2(A) | 312 | 256 | 2.98 | 0.003 | 2.75 | 0.006 |
| 9 | 1(G) | 244 | 237 | 0.56 | 0.576 | 0.56 | 0.576 |
| 10 | 2(A) | 335 | 310 | 0.89 | 0.372 | 1.21 | 0.227 |
| 11 | 2(G) | 209 | 167 | 2.37 | 0.018 | 2.34 | 0.019 |
| 12 | 1(A) | 106 | 92 | 1.23 | 0.220 | 1.23 | 0.220 |
| 13 | 1(T) | 479 | 462 | 1.40 | 0.162 | 1.20 | 0.230 |
| 14 | 2(G) | 367 | 359 | 0.15 | 0.885 | 0.38 | 0.703 |
| 15 | 2(T) | 408 | 392 | 1.29 | 0.198 | 0.31 | 0.755 |
| 16 | 1(T) | 481 | 452 | 1.62 | 0.105 | 2.18 | 0.030 |
| 17 | 2(C) | 292 | 276 | 1.28 | 0.202 | 1.46 | 0.145 |
| 18 | 1(G) | 483 | 463 | 1.15 | 0.250 | 1.19 | 0.233 |
| 19 | 1(T) | 377 | 359 | 1.52 | 0.129 | 1.22 | 0.223 |
| 20 | 1(C) | 443 | 445 | 0.14 | 0.893 | 0.22 | 0.827 |
| 21 | 1(T) | 358 | 320 | 1.85 | 0.065 | 1.63 | 0.102 |
| 22 | 2(G) | 66 | 62 | 0.39 | 0.695 | 0.39 | 0.695 |
| 23 | 2(C) | 150 | 141 | 0.68 | 0.494 | 0.68 | 0.494 |

**Appendix 26 SNP-based haplotypes in LD blocks 1 and 2 of *CACNG3* showing significant (p≤0.05) transmission disequilibrium in the entire resource when analysed using the PDT with affectedness criteria CAE**

a) LD block 1

| *SNPs* | | | | | | | Frequency | | | *SUM PDT* | | *GLOBAL* | | *AVE PDT* | | *GLOBAL* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *2* | *3* | *4* | *5* | *6* | *7* | *8* | in parents (%) | *Transmitted* | *Not Transmitted* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* |
| 1 | 1 | | | | | | 48.2 | 246 | 288 | -2.53 | 0.011 | | | -2.14 | 0.032 | | |
| 1 | 2 | | | | | | 16.2 | 88 | 92 | 0.33 | 0.741 | | | -0.82 | 0.413 | | |
| 2 | 1 | | | | | | 4.1 | 25 | 20 | 0.87 | 0.387 | | | 0.33 | 0.739 | | |
| 2 | 2 | | | | | | 31.5 | 195 | 154 | 2.18 | 0.029 | $9.00_{(3)}$ | 0.029 | 2.69 | 0.007 | $9.43_{(3)}$ | 0.024 |
| | 1 | 1 | | | | | 30.6 | 144 | 160 | -0.90 | 0.367 | | | -0.61 | 0.543 | | |
| | 1 | 2 | | | | | 21.2 | 93 | 117 | -1.92 | 0.056 | | | -1.80 | 0.072 | | |
| | 2 | 1 | | | | | 46.2 | 250 | 208 | 2.54 | 0.011 | | | 2.38 | 0.017 | | |
| | 2 | 2 | | | | | 2.0 | 9 | 11 | -0.58 | 0.564 | $8.44_{(3)}$ | 0.038 | -0.58 | 0.564 | $7.21_{(3)}$ | 0.066 |
| | | 1 | 1 | | | | 49.2 | 263 | 223 | 2.41 | 0.016 | | | 2.24 | 0.025 | | |
| | | 1 | 2 | | | | 27.8 | 129 | 146 | -0.91 | 0.364 | | | -0.71 | 0.481 | | |
| | | 2 | 1 | | | | 22.3 | 98 | 122 | -1.90 | 0.057 | | | -1.77 | 0.076 | | |
| | | 2 | 2 | | | | 0.7 | 4 | 3 | 0.38 | 0.706 | $7.80_{(3)}$ | 0.050 | 0.38 | 0.706 | $6.62_{(3)}$ | 0.085 |
| | | 1 | 1 | 1 | | | 41.9 | 215 | 167 | 2.90 | 0.004 | | | 2.83 | 0.005 | | |
| | | 1 | 1 | 2 | | | 7.0 | 27 | 37 | -0.84 | 0.401 | | | -1.05 | 0.293 | | |
| | | 1 | 2 | 1 | | | 27.1 | 117 | 130 | -0.72 | 0.472 | | | -0.47 | 0.640 | | |
| | | 1 | 2 | 2 | | | 0.5 | 0 | 5 | -1.89 | 0.059 | | | -2.00 | 0.046 | | |
| | | 2 | 1 | 1 | | | 22.4 | 94 | 110 | -1.48 | 0.139 | | | -1.22 | 0.222 | | |
| | | 2 | 1 | 2 | | | 0.4 | 0 | 4 | -1.63 | 0.103 | | | -1.73 | 0.083 | | |
| | | 2 | 2 | 1 | | | 0.7 | 3 | 3 | 0.00 | 1.000 | $15.46_{(6)}$ | 0.017 | 0.00 | 1.000 | $15.29_{(6)}$ | 0.018 |
| | | | 1 | 1 | 1 | | 26.2 | 112 | 133 | -1.84 | 0.066 | | | -1.59 | 0.111 | | |
| | | | 1 | 1 | 2 | | 38.6 | 206 | 155 | 3.09 | 0.002 | | | 3.08 | 0.002 | | |
| | | | 1 | 2 | 1 | | 0.5 | 0 | 5 | -1.89 | 0.059 | | | -2.00 | 0.046 | | |

| *SNPs* | | | | | | | Frequency | | | *SUM PDT* | | *GLOBAL* | | *AVE PDT* | | *GLOBAL* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *2* | *3* | *4* | *5* | *6* | *7* | *8* | in parents (%) | *Transmitted* | *Not Transmitted* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* |
| | | | 1 | 2 | 2 | | 6.7 | 27 | 36 | -0.76 | 0.450 | | | -1.05 | 0.293 | | |
| | | | 2 | 1 | 1 | | 24.8 | 112 | 120 | -0.44 | 0.659 | | | 0.05 | 0.957 | | |
| | | | 2 | 1 | 2 | | 2.4 | 10 | 12 | -0.63 | 0.527 | | | -0.63 | 0.527 | | |
| | | | 2 | 2 | 1 | | 0.7 | 1 | 6 | -1.89 | 0.059 | | | -2.00 | 0.046 | | |
| | | | 2 | 2 | 2 | | 0.1 | 0 | 1 | -1.00 | 0.317 | $19.47_{(7)}$ | 0.007 | -1.00 | 0.317 | $19.71_{(7)}$ | 0.006 |
| | | | 1 | 1 | 1 | 1 | 26.0 | 106 | 129 | -1.96 | 0.050 | | | -1.71 | 0.088 | | |
| | | | 1 | 1 | 1 | 2 | 0.2 | 1 | 1 | 0.00 | 1.000 | | | 0.00 | 1.000 | | |
| | | | 1 | 1 | 2 | 1 | 0.4 | 1 | 3 | -1.41 | 0.157 | | | -1.41 | 0.157 | | |
| | | | 1 | 1 | 2 | 2 | 37.9 | 198 | 145 | 3.26 | 0.001 | | | 3.35 | 0.001 | | |
| | | | 1 | 2 | 1 | 1 | 0.6 | 0 | 5 | -1.89 | 0.059 | | | -2.00 | 0.046 | | |
| | | | 1 | 2 | 2 | 1 | 0.1 | 0 | 1 | -1.00 | 0.317 | | | -1.00 | 0.317 | | |
| | | | 1 | 2 | 2 | 2 | 6.5 | 26 | 33 | -0.57 | 0.567 | | | -0.85 | 0.397 | | |
| | | | 2 | 1 | 1 | 1 | 24.7 | 107 | 116 | -0.56 | 0.579 | | | -0.16 | 0.875 | | |
| | | | 2 | 1 | 1 | 2 | 0.2 | 2 | 0 | 1.00 | 0.317 | | | 1.00 | 0.317 | | |
| | | | 2 | 1 | 2 | 1 | 0.1 | 0 | 1 | -1.00 | 0.317 | | | -1.00 | 0.317 | | |
| | | | 2 | 1 | 2 | 2 | 2.3 | 10 | 11 | -0.33 | 0.739 | | | -0.33 | 0.739 | | |
| | | | 2 | 2 | 1 | 1 | 0.8 | 1 | 6 | -1.89 | 0.059 | | | -2.00 | 0.046 | | |
| | | | 2 | 2 | 2 | 2 | 0.1 | 0 | 1 | -1.00 | 0.317 | $26.20_{(12)}$ | 0.010 | -1.00 | 0.317 | $26.76_{(12)}$ | 0.008 |
| | | | | 1 | 1 | | 51.0 | 225 | 256 | -2.04 | 0.042 | | | -1.66 | 0.098 | | |
| | | | | 1 | 2 | | 41.0 | 219 | 168 | 3.09 | 0.002 | | | 3.08 | 0.002 | | |
| | | | | 2 | 1 | | 1.3 | 1 | 11 | -2.67 | 0.008 | | | -2.83 | 0.005 | | |
| | | | | 2 | 2 | | 6.8 | 27 | 37 | -0.85 | 0.397 | $16.19_{(3)}$ | 0.001 | -1.20 | 0.230 | $16.25_{(3)}$ | 0.001 |
| | | | | 1 | 1 | 1 | 50.7 | 214 | 248 | -2.25 | 0.024 | | | -1.98 | 0.048 | | |
| | | | | 1 | 1 | 2 | 0.4 | 3 | 1 | 0.82 | 0.414 | | | 0.82 | 0.414 | | |
| | | | | 1 | 2 | 1 | 0.5 | 1 | 4 | -1.73 | 0.083 | | | -1.73 | 0.083 | | |
| | | | | 1 | 2 | 2 | 40.4 | 211 | 157 | 3.31 | 0.001 | | | 3.43 | 0.001 | | |
| | | | | 2 | 1 | 1 | 1.3 | 1 | 11 | -2.67 | 0.008 | | | -2.83 | 0.005 | | |
| | | | | 2 | 2 | 1 | 0.1 | 0 | 1 | -1.00 | 0.317 | | | -1.00 | 0.317 | | |
| | | | | 2 | 2 | 2 | 6.6 | 26 | 34 | -0.66 | 0.506 | $24.24_{(6)}$ | 0.001 | -1.00 | 0.318 | $25.12_{(6)}$ | 0.000 |

| *SNPs* | | | | | | | Frequency | | | *SUM PDT* | | *GLOBAL* | | *AVE PDT* | | *GLOBAL* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *2* | *3* | *4* | *5* | *6* | *7* | *8* | in parents (%) | *Transmitted* | *Not Transmitted* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* |
| | | | | | 1 | 1 | 52.1 | 237 | 284 | -2.86 | 0.004 | | | -2.76 | 0.006 | | |
| | | | | | 1 | 2 | 0.4 | 3 | 1 | 0.82 | 0.414 | | | 0.82 | 0.414 | | |
| | | | | | 2 | 1 | 0.6 | 1 | 5 | -1.63 | 0.103 | | | -1.63 | 0.103 | | |
| | | | | | 2 | 2 | 46.9 | 259 | 210 | 2.95 | 0.003 | $15.15_{(3)}$ | 0.002 | 2.91 | 0.004 | $14.58_{(3)}$ | 0.002 |

b) LD block 2

| *SNP* | | | | *Frequency* | | | *SUM PDT* | | *GLOBAL* | | *AVE PDT* | | *GLOBAL* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *10* | *11* | *12* | *13* | *in parents (%)* | *Transmitted* | *Not Transmitted* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* |
| 1 | 1 | | | 42.3 | 190 | 223 | -1.72 | 0.086 | | | -1.98 | 0.048 | | |
| 1 | 2 | | | 1.4 | 6 | 8 | -0.30 | 0.763 | | | -0.54 | 0.593 | | |
| 2 | 1 | | | 26.3 | 123 | 134 | -0.93 | 0.353 | | | -0.73 | 0.469 | | |
| 2 | 2 | | | 29.9 | 169 | 123 | 3.32 | 0.001 | $11.21_{(3)}$ | 0.011 | 3.09 | 0.002 | $10.72_{(3)}$ | 0.013 |
| | 1 | 1 | | 5.1 | 16 | 20 | -1.63 | 0.103 | | | -1.63 | 0.103 | | |
| | 1 | 2 | | 65.7 | 217 | 243 | -2.32 | 0.020 | | | -2.30 | 0.022 | | |
| | 2 | 1 | | 23.9 | 96 | 71 | 2.27 | 0.023 | | | 2.23 | 0.026 | | |
| | 2 | 2 | | 5.3 | 21 | 16 | 1.89 | 0.059 | $12.58_{(3)}$ | 0.006 | 1.89 | 0.059 | $12.37_{(3)}$ | 0.006 |
| 1 | 1 | 1 | | 1.5 | 4 | 6 | -1.00 | 0.317 | | | -1.00 | 0.317 | | |
| 1 | 1 | 2 | | 41.6 | 128 | 158 | -1.84 | 0.066 | | | -2.05 | 0.041 | | |
| 1 | 2 | 1 | | 0.7 | 2 | 3 | -0.45 | 0.655 | | | -0.45 | 0.655 | | |
| 1 | 2 | 2 | | 1.0 | 2 | 5 | -1.04 | 0.297 | | | -0.71 | 0.480 | | |
| 2 | 1 | 1 | | 3.8 | 12 | 14 | -0.71 | 0.480 | | | -0.71 | 0.480 | | |
| 2 | 1 | 2 | | 23.7 | 84 | 79 | 0.00 | 1.000 | | | -0.03 | 0.973 | | |
| 2 | 2 | 1 | | 23.4 | 93 | 68 | 2.30 | 0.021 | | | 2.30 | 0.021 | | |
| 2 | 2 | 2 | | 4.4 | 19 | 11 | 2.13 | 0.033 | $14.01_{(7)}$ | 0.051 | 2.65 | 0.008 | $16.36_{(7)}$ | 0.022 |
| | 1 | 1 | 1 | 4.6 | 14 | 17 | -1.00 | 0.317 | | | -1.00 | 0.317 | | |
| | 1 | 1 | 2 | 0.7 | 2 | 3 | -0.45 | 0.655 | | | -0.45 | 0.655 | | |
| | 1 | 2 | 1 | 49.0 | 149 | 180 | -1.63 | 0.103 | | | -2.11 | 0.035 | | |
| | 1 | 2 | 2 | 15.8 | 55 | 51 | -0.44 | 0.664 | | | -0.03 | 0.978 | | |
| | 2 | 1 | 1 | 24.4 | 95 | 69 | 2.35 | 0.019 | | | 2.35 | 0.019 | | |
| | 2 | 2 | 1 | 4.2 | 15 | 13 | 0.00 | 1.000 | | | 1.00 | 0.317 | | |
| | 2 | 2 | 2 | 1.3 | 6 | 3 | 1.21 | 0.225 | $9.45_{(6)}$ | 0.150 | 1.41 | 0.157 | $12.14_{(6)}$ | 0.059 |

**Appendix 27 SNP-based haplotypes in LD blocks 1 and 2 of *CACNG3* showing significant (p≤0.05) transmission disequilibrium in the entire resource when analysed using the PDT with affectedness criteria CAE, JAE or AE**

a) LD block 1

| *SNP* | | | | | | | *Frequency* | | | | *SUM PDT* | | *GLOBAL* | | *AVE PDT* | | *GLOBAL* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *2* | *3* | *4* | *5* | *6* | *7* | *8* | *in parents (%)* | *Transmitted* | *Not Transmitted* | | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* |
| 1 | 1 | | | | | | 296 | 346 | 48.3 | | -2.66 | 0.008 | | | -2.31 | 0.021 | | |
| 1 | 2 | | | | | | 104 | 108 | 16.0 | | 0.06 | 0.953 | | | -0.69 | 0.491 | | |
| 2 | 1 | | | | | | 28 | 23 | 3.8 | | 0.85 | 0.396 | | | 0.38 | 0.705 | | |
| 2 | 2 | | | | | | 236 | 187 | 31.9 | | 2.56 | 0.011 | $10.74_{(3)}$ | 0.013 | 2.67 | 0.008 | $9.83_{(3)}$ | 0.020 |
| 1 | 1 | 1 | | | | | 166 | 180 | 29.0 | | -0.55 | 0.582 | | | -0.11 | 0.913 | | |
| 1 | 1 | 2 | | | | | 98 | 131 | 19.2 | | -2.52 | 0.012 | | | -2.30 | 0.021 | | |
| 1 | 2 | 1 | | | | | 86 | 87 | 14.5 | | 0.24 | 0.809 | | | -0.44 | 0.663 | | |
| 1 | 2 | 2 | | | | | 4 | 9 | 1.1 | | -1.89 | 0.059 | | | -1.89 | 0.059 | | |
| 2 | 1 | 1 | | | | | 11 | 11 | 1.8 | | 0.00 | 1.000 | | | 0.00 | 1.000 | | |
| 2 | 1 | 2 | | | | | 13 | 10 | 1.9 | | 0.69 | 0.493 | | | 0.26 | 0.795 | | |
| 2 | 2 | 1 | | | | | 212 | 165 | 31.6 | | 2.55 | 0.011 | | | 2.67 | 0.008 | | |
| 2 | 2 | 2 | | | | | 6 | 3 | 0.8 | | 1.00 | 0.317 | $15.97_{(7)}$ | 0.025 | 1.00 | 0.317 | $15.10_{(7)}$ | 0.035 |
| | 1 | 1 | | | | | 177 | 195 | 30.9 | | -0.73 | 0.464 | | | -0.47 | 0.638 | | |
| | 1 | 2 | | | | | 112 | 142 | 21.1 | | -2.25 | 0.025 | | | -2.07 | 0.039 | | |
| | 2 | 1 | | | | | 303 | 252 | 46.1 | | 2.70 | 0.007 | | | 2.60 | 0.009 | | |
| | 2 | 2 | | | | | 10 | 13 | 1.9 | | -0.83 | 0.405 | $10.19_{(3)}$ | 0.017 | -0.83 | 0.405 | $8.96_{(3)}$ | 0.030 |
| | 1 | 1 | 1 | 1 | | | 29 | 29 | 5.3 | | -0.28 | 0.782 | | | 0.00 | 1.000 | | |
| | 1 | 1 | 2 | 1 | | | 133 | 142 | 25.3 | | -0.15 | 0.881 | | | 0.00 | 0.998 | | |
| | 1 | 1 | 2 | 2 | | | 0 | 4 | 0.4 | | -1.63 | 0.103 | | | -1.73 | 0.083 | | |
| | 1 | 2 | 1 | 1 | | | 102 | 121 | 20.5 | | -1.68 | 0.094 | | | -1.53 | 0.127 | | |

| SNP | | | | | | | Frequency | | | SUM PDT | | GLOBAL | | AVE PDT | | GLOBAL | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 3 | 4 | 5 | 6 | 7 | 8 | in parents (%) | Transmitted | Not Transmitted | Z (1df) | p-value | $\chi^2_{(df)}$ | p-value | Z (1df) | p-value | $\chi^2_{(df)}$ | p-value |
| | 1 | 2 | 1 | 2 | | | 0 | 3 | 0.3 | -1.34 | 0.180 | | | -1.41 | 0.157 | | |
| | 1 | 2 | 2 | 1 | | | 3 | 3 | 0.6 | 0.00 | 1.000 | | | 0.00 | 1.000 | | |
| | 2 | 1 | 1 | 1 | | | 224 | 172 | 36.4 | 3.22 | 0.001 | | | 3.15 | 0.002 | | |
| | 2 | 1 | 1 | 2 | | | 29 | 43 | 6.6 | -1.40 | 0.161 | | | -1.52 | 0.128 | | |
| | 2 | 1 | 2 | 1 | | | 12 | 14 | 2.4 | -0.71 | 0.480 | | | -0.71 | 0.480 | | |
| | 2 | 1 | 2 | 2 | | | 2 | 1 | 0.3 | 1.00 | 0.317 | | | 1.00 | 0.317 | | |
| | 2 | 2 | 1 | 1 | | | 10 | 10 | 1.8 | 0.00 | 1.000 | | | 0.00 | 1.000 | | |
| | 2 | 2 | 1 | 2 | | | 0 | 2 | 0.2 | -1.41 | 0.157 | $21.27_{(11)}$ | 0.031 | -1.41 | 0.157 | $21.12_{(11)}$ | 0.032 |
| | | 1 | 1 | | | | 313 | 269 | 48.8 | 2.42 | 0.016 | | | 2.26 | 0.024 | | |
| | | 1 | 2 | | | | 161 | 177 | 28.4 | -0.60 | 0.549 | | | -0.38 | 0.703 | | |
| | | 2 | 1 | | | | 118 | 147 | 22.2 | -2.20 | 0.028 | | | -2.04 | 0.042 | | |
| | | 2 | 2 | | | | 4 | 3 | 0.6 | 0.38 | 0.706 | $8.38_{(3)}$ | 0.039 | 0.38 | 0.706 | $7.17_{(3)}$ | 0.067 |
| | | 1 | 1 | 1 | | | 255 | 204 | 41.9 | 3.08 | 0.002 | | | 2.93 | 0.003 | | |
| | | 1 | 1 | 2 | | | 29 | 43 | 6.6 | -1.40 | 0.161 | | | -1.52 | 0.128 | | |
| | | 1 | 2 | 1 | | | 146 | 157 | 27.6 | -0.25 | 0.805 | | | -0.21 | 0.837 | | |
| | | 1 | 2 | 2 | | | 2 | 5 | 0.6 | -1.13 | 0.257 | | | -1.00 | 0.317 | | |
| | | 2 | 1 | 1 | | | 113 | 131 | 22.3 | -1.60 | 0.109 | | | -1.37 | 0.172 | | |
| | | 2 | 1 | 2 | | | 0 | 5 | 0.5 | -1.89 | 0.059 | | | -2.00 | 0.046 | | |
| | | 2 | 2 | 1 | | | 3 | 3 | 0.5 | 0.00 | 1.000 | $16.21_{(6)}$ | 0.013 | 0.00 | 1.000 | $15.25_{(6)}$ | 0.018 |
| | | 1 | 1 | 1 | 1 | | 30 | 29 | 5.5 | 0.00 | 1.000 | | | 0.30 | 0.763 | | |
| | | 1 | 1 | 1 | 2 | | 224 | 174 | 36.9 | 3.11 | 0.002 | | | 2.93 | 0.003 | | |
| | | 1 | 1 | 2 | 2 | | 29 | 41 | 6.5 | -1.25 | 0.210 | | | -1.38 | 0.169 | | |
| | | 1 | 2 | 1 | 1 | | 129 | 137 | 24.6 | -0.10 | 0.920 | | | 0.10 | 0.917 | | |
| | | 1 | 2 | 1 | 2 | | 12 | 15 | 2.5 | -1.00 | 0.317 | | | -1.00 | 0.317 | | |
| | | 1 | 2 | 2 | 1 | | 0 | 4 | 0.4 | -1.63 | 0.103 | | | -1.73 | 0.083 | | |
| | | 1 | 2 | 2 | 2 | | 2 | 1 | 0.3 | 1.00 | 0.317 | | | 1.00 | 0.317 | | |
| | | 2 | 1 | 1 | 1 | | 100 | 122 | 20.6 | -1.84 | 0.066 | | | -1.71 | 0.087 | | |
| | | 2 | 1 | 1 | 2 | | 11 | 9 | 1.9 | 0.58 | 0.564 | | | 0.58 | 0.564 | | |
| | | 2 | 1 | 2 | 1 | | 0 | 3 | 0.3 | -1.34 | 0.180 | | | -1.41 | 0.157 | | |

| SNP | | | | | | | Frequency | | | | SUM PDT | | GLOBAL | | AVE PDT | | GLOBAL | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 3 | 4 | 5 | 6 | 7 | 8 | in (%) parents | Transmitted | Not Transmitted | Z (1df) | p-value | $\chi^2_{(df)}$ | p-value | Z (1df) | p-value | $\chi^2_{(df)}$ | p-value |
| | | 2 | 1 | 2 | 2 | | 0 | 2 | 0.2 | -1.41 | 0.157 | | | -1.41 | 0.157 | | |
| | | 2 | 2 | 1 | 1 | | 3 | 3 | 0.6 | 0.00 | 1.000 | $21.49_{(11)}$ | 0.029 | 0.00 | 1.000 | $20.97_{(11)}$ | 0.034 |
| | | 1 | 1 | 1 | 1 | 1 | 30 | 29 | 5.6 | 0.00 | 1.000 | | | 0.30 | 0.763 | | |
| | | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 0.3 | -1.00 | 0.317 | | | -1.00 | 0.317 | | |
| | | 1 | 1 | 1 | 2 | 2 | 217 | 167 | 36.5 | 3.17 | 0.002 | | | 3.03 | 0.003 | | |
| | | 1 | 1 | 2 | 2 | 1 | 0 | 2 | 0.2 | -1.34 | 0.180 | | | -1.21 | 0.225 | | |
| | | 1 | 1 | 2 | 2 | 2 | 28 | 37 | 6.2 | -0.94 | 0.347 | | | -1.13 | 0.257 | | |
| | | 1 | 2 | 1 | 1 | 1 | 124 | 133 | 24.4 | -0.20 | 0.841 | | | -0.11 | 0.911 | | |
| | | 1 | 2 | 1 | 1 | 2 | 2 | 0 | 0.2 | 1.00 | 0.317 | | | 1.00 | 0.317 | | |
| | | 1 | 2 | 1 | 2 | 1 | 0 | 1 | 0.1 | -1.00 | 0.317 | | | -1.00 | 0.317 | | |
| | | 1 | 2 | 1 | 2 | 2 | 12 | 14 | 2.5 | -0.71 | 0.480 | | | -0.71 | 0.480 | | |
| | | 1 | 2 | 2 | 1 | 1 | 0 | 4 | 0.4 | -1.63 | 0.103 | | | -1.73 | 0.083 | | |
| | | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 0.3 | 1.00 | 0.317 | | | 1.00 | 0.317 | | |
| | | 2 | 1 | 1 | 1 | 1 | 96 | 118 | 20.3 | -1.85 | 0.065 | | | -1.64 | 0.101 | | |
| | | 2 | 1 | 1 | 1 | 2 | 0 | 1 | 0.1 | -1.00 | 0.317 | | | -1.00 | 0.317 | | |
| | | 2 | 1 | 1 | 2 | 2 | 11 | 9 | 1.9 | 0.58 | 0.564 | | | 0.58 | 0.564 | | |
| | | 2 | 1 | 2 | 1 | 1 | 0 | 3 | 0.3 | -1.34 | 0.180 | | | -1.41 | 0.157 | | |
| | | 2 | 1 | 2 | 2 | 2 | 0 | 2 | 0.2 | -1.41 | 0.157 | | | -1.41 | 0.157 | | |
| | | 2 | 2 | 1 | 1 | 1 | 3 | 3 | 0.6 | 0.00 | 1.000 | $26.82_{(16)}$ | 0.044 | 0.00 | 1.000 | $25.93_{(16)}$ | 0.055 |
| | | | 1 | 1 | 1 | | 136 | 159 | 26.4 | -1.91 | 0.056 | | | -1.73 | 0.085 | | |
| | | | 1 | 1 | 2 | | 242 | 188 | 38.5 | 3.27 | 0.001 | | | 3.17 | 0.002 | | |
| | | | 1 | 2 | 1 | | 0 | 5 | 0.4 | -1.89 | 0.059 | | | -2.00 | 0.046 | | |
| | | | 1 | 2 | 2 | | 29 | 43 | 6.5 | -1.41 | 0.158 | | | -1.66 | 0.097 | | |
| | | | 2 | 1 | 1 | | 136 | 141 | 24.8 | 0.05 | 0.961 | | | 0.40 | 0.690 | | |
| | | | 2 | 1 | 2 | | 12 | 15 | 2.4 | -1.00 | 0.317 | | | -1.00 | 0.317 | | |
| | | | 2 | 2 | 1 | | 1 | 6 | 0.6 | -1.89 | 0.059 | | | -2.00 | 0.046 | | |
| | | | 2 | 2 | 2 | | 2 | 1 | 0.3 | 1.00 | 0.317 | $22.30_{(7)}$ | 0.002 | 1.00 | 0.317 | $22.71_{(7)}$ | 0.002 |
| | | | 1 | 1 | 1 | 1 | 130 | 155 | 26.3 | -2.02 | 0.044 | | | -1.84 | 0.066 | | |
| | | | 1 | 1 | 1 | 2 | 1 | 1 | 0.2 | 0.00 | 1.000 | | | 0.00 | 1.000 | | |

| SNP | | | | | | | Frequency | | | SUM PDT | | GLOBAL | | AVE PDT | | GLOBAL | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 3 | 4 | 5 | 6 | 7 | 8 | in (%) parents | Transmitted | Not Transmitted | Z (1df) | p-value | $\chi^2_{(df)}$ | p-value | Z (1df) | p-value | $\chi^2_{(df)}$ | p-value |
| | | | 1 | 1 | 2 | 1 | 1 | 3 | 0.4 | -1.41 | 0.157 | | | -1.41 | 0.157 | | |
| | | | 1 | 1 | 2 | 2 | 234 | 178 | 38.0 | 3.44 | 0.001 | | | 3.44 | 0.001 | | |
| | | | 1 | 2 | 1 | 1 | 0 | 5 | 0.5 | -1.89 | 0.059 | | | -2.00 | 0.046 | | |
| | | | 1 | 2 | 2 | 1 | 0 | 2 | 0.2 | -1.34 | 0.180 | | | -1.21 | 0.225 | | |
| | | | 1 | 2 | 2 | 2 | 28 | 39 | 6.2 | -1.10 | 0.270 | | | -1.43 | 0.153 | | |
| | | | 2 | 1 | 1 | 1 | 131 | 137 | 24.7 | -0.05 | 0.961 | | | 0.20 | 0.840 | | |
| | | | 2 | 1 | 1 | 2 | 2 | 0 | 0.2 | 1.00 | 0.317 | | | 1.00 | 0.317 | | |
| | | | 2 | 1 | 2 | 1 | 0 | 1 | 0.1 | -1.00 | 0.317 | | | -1.00 | 0.317 | | |
| | | | 2 | 1 | 2 | 2 | 12 | 14 | 2.4 | -0.71 | 0.480 | | | -0.71 | 0.480 | | |
| | | | 2 | 2 | 1 | 1 | 1 | 6 | 0.6 | -1.89 | 0.059 | | | -2.00 | 0.046 | | |
| | | | 2 | 2 | 2 | 2 | 2 | 1 | 0.3 | 1.00 | 0.317 | $29.11_{(12)}$ | 0.004 | 1.00 | 0.317 | $29.78_{(12)}$ | 0.003 |
| | | | | 1 | 1 | | 273 | 307 | 51.2 | -1.95 | 0.052 | | | -1.67 | 0.094 | | |
| | | | | 1 | 2 | | 259 | 204 | 40.9 | 3.31 | 0.001 | | | 3.17 | 0.002 | | |
| | | | | 2 | 1 | | 1 | 11 | 1.1 | -2.67 | 0.008 | | | -2.83 | 0.005 | | |
| | | | | 2 | 2 | | 33 | 44 | 6.8 | -1.15 | 0.250 | $17.40_{(3)}$ | 0.001 | -1.34 | 0.180 | $16.99_{(3)}$ | 0.001 |
| | | | | 1 | 1 | 1 | 262 | 299 | 51.0 | -2.15 | 0.032 | | | -1.99 | 0.047 | | |
| | | | | 1 | 1 | 2 | 3 | 1 | 0.4 | 0.82 | 0.414 | | | 0.82 | 0.414 | | |
| | | | | 1 | 2 | 1 | 1 | 4 | 0.5 | -1.73 | 0.083 | | | -1.73 | 0.083 | | |
| | | | | 1 | 2 | 2 | 251 | 193 | 40.4 | 3.52 | 0.000 | | | 3.52 | 0.000 | | |
| | | | | 2 | 1 | 1 | 1 | 11 | 1.1 | -2.67 | 0.008 | | | -2.83 | 0.005 | | |
| | | | | 2 | 2 | 1 | 0 | 2 | 0.2 | -1.34 | 0.180 | | | -1.21 | 0.225 | | |
| | | | | 2 | 2 | 2 | 32 | 40 | 6.5 | -0.84 | 0.401 | $25.96_{(6)}$ | 0.000 | -1.10 | 0.270 | $26.28_{(6)}$ | 0.000 |
| | | | | | 1 | 1 | 287 | 344 | 52.2 | -3.01 | 0.003 | | | -3.03 | 0.003 | | |
| | | | | | 1 | 2 | 3 | 1 | 0.3 | 0.82 | 0.414 | | | 0.82 | 0.414 | | |
| | | | | | 2 | 1 | 1 | 6 | 0.6 | -1.90 | 0.058 | | | -1.73 | 0.084 | | |
| | | | | | 2 | 2 | 313 | 253 | 46.9 | 3.19 | 0.001 | $17.64_{(3)}$ | 0.001 | 3.19 | 0.001 | $17.23_{(3)}$ | 0.001 |

b) LD block 2

| *SNP* | | | | *Frequency* | | | *SUM PDT* | | *GLOBAL* | | *AVE PDT* | | *GLOBAL* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *10* | *11* | *12* | *13* | *in parents (%)* | *Transmitted* | *Not Transmitted* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* | *Z (1df)* | *p-value* | $\chi^2_{(df)}$ | *p-value* |
| 1 | 1 | | | 42.7 | 237 | 269 | -1.31 | 0.191 | | | -1.68 | 0.093 | | |
| 1 | 2 | | | 1.4 | 8 | 9 | -0.15 | 0.882 | | | -0.26 | 0.796 | | |
| 2 | 1 | | | 26.3 | 148 | 163 | -0.93 | 0.354 | | | -0.95 | 0.343 | | |
| 2 | 2 | | | 29.6 | 199 | 151 | 2.72 | 0.007 | $7.50_{(3)}$ | 0.058 | 2.87 | 0.004 | $9.04_{(3)}$ | 0.029 |
| | 1 | 1 | | 5.3 | 17 | 23 | -1.90 | 0.058 | | | -1.90 | 0.058 | | |
| | 1 | 2 | | 65.7 | 234 | 263 | -2.51 | 0.012 | | | -2.41 | 0.016 | | |
| | 2 | 1 | | 23.5 | 103 | 75 | 2.45 | 0.014 | | | 2.35 | 0.019 | | |
| | 2 | 2 | | 5.4 | 24 | 17 | 2.11 | 0.035 | $15.27_{(3)}$ | 0.002 | 2.11 | 0.035 | $14.52_{(3)}$ | 0.002 |
| 1 | 1 | 1 | | 1.3 | 4 | 6 | -1.00 | 0.317 | | | -1.00 | 0.317 | | |
| 1 | 1 | 2 | | 41.7 | 140 | 170 | -1.81 | 0.070 | | | -1.91 | 0.056 | | |
| 1 | 2 | 1 | | 0.7 | 2 | 3 | -0.45 | 0.655 | | | -0.45 | 0.655 | | |
| 1 | 2 | 2 | | 1.1 | 3 | 5 | -0.82 | 0.414 | | | -0.33 | 0.739 | | |
| 2 | 1 | 1 | | 4.0 | 13 | 17 | -1.16 | 0.248 | | | -1.16 | 0.248 | | |
| 2 | 1 | 2 | | 23.7 | 89 | 87 | -0.21 | 0.838 | | | -0.25 | 0.805 | | |
| 2 | 2 | 1 | | 23.1 | 100 | 72 | 2.49 | 0.013 | | | 2.42 | 0.015 | | |
| 2 | 2 | 2 | | 4.4 | 21 | 12 | 2.12 | 0.034 | $15.07_{(7)}$ | 0.035 | 2.31 | 0.021 | $15.36_{(7)}$ | 0.032 |
| | 1 | 1 | 1 | 4.8 | 15 | 20 | -1.39 | 0.166 | | | -1.39 | 0.166 | | |
| | 1 | 1 | 2 | 0.7 | 2 | 3 | -0.45 | 0.655 | | | -0.45 | 0.655 | | |
| | 1 | 2 | 1 | 48.9 | 161 | 195 | -1.77 | 0.077 | | | -2.13 | 0.033 | | |
| | 1 | 2 | 2 | 15.9 | 60 | 56 | -0.43 | 0.667 | | | -0.03 | 0.979 | | |
| | 2 | 1 | 1 | 24.0 | 102 | 73 | 2.53 | 0.011 | | | 2.46 | 0.014 | | |
| | 2 | 2 | 1 | 4.0 | 16 | 13 | 0.20 | 0.842 | | | 1.27 | 0.206 | | |
| | 2 | 2 | 2 | 1.6 | 8 | 4 | 1.28 | 0.201 | $11.59_{(6)}$ | 0.072 | 1.13 | 0.257 | $13.36_{(6)}$ | 0.038 |

## Appendix 28 Molecular Biology Solutions and reagents

**ACD solution:**

| | | |
|---|---|---|
| Trisodium citrate | | 13.2g/L |
| Citric acid | | 4.8g/L |
| Dextrose | 1 | 4.7g/L |

**Agarose:**

GibcoBRL electrophoresis grade 10% w/v

**10% APS:** Ammonium persulphate

**Blue dextran loading buffer:**

| | |
|---|---|
| De-ionised formamide | 10ml |
| 50mM EDTA (pH8.0) | 2ml |
| Blue Dextran | 360mg |

**0.5M EDTA:**

EDTA.Na2.2H2O (pH8.0)
186.12g disodium ethylenediaminetetra-acetate/litre of water.
Adjust pH to 8.0 with NaOH to allow disodium salt to go into solution

**Ethidium bromide (10mg/ml):**

1g of ethidium bromide to 100ml of water. Stir for several hours. Place in dark bottle or wrap in aluminium foil and store at 4$^{o}$C.

**Proteinase K:** 20mg/ml in water.

**Reaction buffer:**

0.01M Tris,
0.005M EDTA,
0.5% SDS.
Stored at -20$^{o}$C

**10% SDS:** **Sodium dodecyl sulphate**

100g of electrophoresis grade SDS in 900 ml of water. Heat to 68$^{o}$C. Add HCL to adjust pH to 7.2. Make up to 1 litre with water.

**TE buffer:** 10mM Tris-HCl (pH7.6)
1.0mM EDTA (pH8.0)

**10X TBE:** **Tris-borate-EDTA electrophoresis buffer**

| | |
|---|---|
| Tris base | 54g |
| Boric acid | 27.5g |
| 0.5M EDTA | 20ml |

## Published papers

### *i. Research Papers*

Everett K, Chioza B et al, Robinson R, Taske N, Rees M, Gardiner RM. **Linkage and association analysis of CACNG3 in childhood absence epilepsy.** Eur J Hum Genet. 2007;15(4):463-72.

Hempelmann A, Taylor K, Heils A, Lorenz S, Prud'Homme J, Nabbout R, Dulac O, Rudolf G, Zara F, Bianchi A, Robinson R, Gardiner RM, Covanis A, Lindhout D, Stephani U,. Elger C, Weber Y, Lerche H, Nürnberg P, Kron KL, Scheffer I, Mulley J, Berkovic S, Sander T. **Exploration of the Genetic Architecture of Idiopathic Generalized Epilepsies.** Epilepsia. 2006;47(10):1682-90.

Chioza B, Everett K, Aschauer H, Brouwer O, Callenbach P, Covanis A, Dulac O, Durner M, Eeg-Olofsson O, Feucht M, Friis M, Heils A, Kjeldsen A, Nashef L, Olsson I, Sander T, Siren A, Robinson R, Rees M, Gardiner RM. **Evaluation of *CACNA1H* in European patients with childhood absence epilepsy.** Epilepsy Research;69(2):177-81.

Imbrici P, Jaffe SL, Eunson LH, Davies NP, Herd C, Robinson R, Kullmann DM, Hanna MG. **Dysfunction of the brain calcium channel [$Ca_V$ 2.1] in absence epilepsy .** Brain 2004;127(Pt 12):2682-92.

**Robinson R, Taske N, Sander T, Heils A, Whitehouse W, Goutières G, Aicardi J, Lehesjoki A, Siren A, Friis M, Kjeldsen M, Panayiotopoulos C, Kennedy C, Ferrie C, Rees M, Gardiner RM.** Linkage analysis between childhood absence epilepsy and genes encoding $GABA_A$ and $GABA_B$ receptors, voltage-dependent calcium channels, and the *ECA1* region on chromosome 8q. **Epilepsy Res. 2002;48:p169-179**

ii. Reviews

Robinson R, Gardiner RM. **Molecular Genetics of the Epilepsies**. In: Epilepsy 2005: from neuron to NICE - a practical guide: National Society for Epilepsy (www.e-epilepsy.org.uk), 2005

Robinson R, Gardiner RM. **Genetics**. In: Wallace S, editor. Epilepsy in Children 2nd Edition: Chapman and Hall, 2004.

Robinson R, Gardiner RM. **Molecular basis of mendelian idiopathic epilepsies**. Ann Med. 2004;36(2):89-97

Robinson R, Gardiner RM. **Genetics of childhood epilepsy**. Arch Dis Child. 2000; 82:121-125

### *iv. Published Abstracts*

Everett K, Chioza B, Sharp R, McKeigue P, Robinson R, Taske N, Rees M, Gardiner RM. **Absence epilepsy: linkage and association analysis of *CACNG3***. British Paediatric Neurology Association Abstracts 2006. Dev Med & Child Neurol. 2006;S104:p34

Robinson R, Taske N, Boland P, Rees M, Gardiner RM. **Evidence for *CACNG3* as a susceptibility locus for childhood absence epilepsy.** British Paediatric Neurology Association Abstracts 2003. Dev Med & Child Neurol. 2003;S93:p35

Robinson R, Taske N, Rees M, Gardiner RM. **Linkage analysis between childhood absence epilepsy and *GABRA5, GABRB3, GABRG3* on chromosome 15q11-13.** Abstracts of the American Epilepsy Society Annual Meeting. Epilepsia 2002;43(S7):p277

Robinson R, Taske N, Heils A, Sander T, Rees M, Gardiner RM. **Linkage analysis between childhood absence epilepsy and genes encoding $GABA_A$ and $GABA_B$ receptors**. British Paediatric Neurology Association Abstracts 2002. Dev Med & Child Neurol. 2001;S90:p35

Robinson R, Taske N, Heils A, Sander T, Rees M, Gardiner RM. **Investigation of voltage dependent calcium channels as candidate genes for childhood absence epilepsy**. British Paediatric Neurology Association Abstracts 2001. Dev Med & Child Neurol. 2000;S85:p36

Robinson R, Taske N, Heils A, Rees M, Gardiner RM. **Investigation of voltage dependent calcium channels as candidate genes for childhood absence epilepsy.** Abstracts of the American Epilepsy Society Annual Meeting 2000. Epilepsia 2000;41(S7):p70